



Dissertation

DESIGN, APPLICATION, AND OPTIMIZATION OF SYNTHETIC ENZYME CASCADES IN *ESCHERICHIA COLI*

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften unter der Leitung von

Dr. Florian Rudroff

Prof. Dr. Marko D. Mihovilovic

Institut für Angewandte Synthesechemie (E163)

eingereicht an der Technischen Universität Wien

Fakultät für Technische Chemie

von

Mag. Thomas Bayer

Wien, 2017

*To my sister, Iris,
who is very brave.*

*To my parents,
who supported me no matter what.*

*To my grandmother,
who I miss heaps.
(21.01.1928 – 17.11.2014)*

*To Martin,
who will always have a place in my heart.
(18.07.1987 – 21.06.2016)*

Front Matter

Table of Contents

Front Matter	i
Table of Contents	i
Acknowledgements	vii
Abstract	ix
Kurzfassung	xi
A Synthetic enzyme cascade schemes	13
A I Artificial metabolic pathways producing polyhydroxylated compounds <i>in vivo</i>	13
B Compound library	18
B I Cascade substrates and intermediates: Aromatic esters, primary alcohols, aldehydes, and carboxylic acids	18
B II Cascade products: Aromatic aldol adducts	19
B III Metabolic background studies in <i>E. coli</i> : ERED substrates and products	20
C Introduction	21
C I Out of a test tube: The evolution of biocatalysis	21
C II Tools for cellular pathway engineering	24
C II.1 The challenges of multiple recombinant protein expression	25
C II.2 Advanced cloning techniques	28
C II.3 Genomic modification tools	31
C III Optimization of synthetic pathway elements	38
C III.1 Optimization on the levels of transcription and translation	39
C III.2 Engineering enzyme performance	43
C IV Improving synthetic pathway performance in the cellular context	48
C IV.1 Substrate uptake and product release	49
C IV.2 The intracellular environment and metabolic background	50
C IV.3 Cofactor balancing and recycling	53
C IV.4 Synthetic biology tools for carbon flux enhancement	59
C V State-of-the-art microbial cell factories	62
C V.1 The synthesis of aldehydes as products and valuable intermediates	63
C V.2 Carbon framework expansion by (engineered) aldolases	65
C VI Scope and aim of this thesis	68
D Biocatalyst selection and <i>de novo</i> pathway assembly	71
D I Biocatalytic retrosynthesis: From sugar derivatives to simple, achiral substrates	71
D II Identification and characterization of suitable biocatalysts	73
D II.1 Getting started: Expression and characterization of esterases	73
D II.2 Enzymes for the oxidation of primary alcohols to aldehydes <i>in vivo</i>	75
D II.2.1 Unsuccessful substrate acceptance screening of <i>ADH_{Lk}</i> , <i>ADH_{Rt}</i> , <i>ADH-A</i> , and <i>ADH-ht</i>	75
D II.2.2 Successful cloning and expression of <i>AlkJ</i> and substrate scope expansion toward non-native compound classes	78
D II.3 Investigation of carboxylic acid reductases (CARs) to access the intracellular sink of carboxylate byproducts	83
D II.4 Carbon framework expansion from aldehydes: Characterization and selection of aldolases	88
D II.4.1 Expression of <i>Fsa1</i> and <i>Fsa1-A129S</i> and <i>in vitro</i> characterization of the engineered aldolase	89

D II.4.2	Cloning, expression, and <i>in vitro</i> characterization of DHAP-dependent aldolases	90
D II.5	Last but not least: Characterization and selection of wild type and engineered phosphatases	93
D III	Assembly of <i>de novo</i> pathways	98
D III.1	Assembly of a synthetic mini-pathway consisting of <i>AlkJ</i> and the mutant aldolase <i>Fsa1-A129S</i>	98
D III.1.1	<i>De novo</i> mini-pathway assembly and characterization	98
D III.1.2	Successful production of polyhydroxylated compounds <i>in vivo</i>	102
D III.2	Assembly of a <i>de novo</i> pathway featuring <i>AlkJ</i> , the DHAP-dependent aldolase <i>FucA</i> , and <i>PhoN-Sf</i>	105
D III.2.1	Synthetic DHAP-dependent aldolase pathway assembly and characterization	105
D III.2.2	Unsuccessful production of (3 <i>R</i> ,4 <i>R</i>)-1,3,4-trihydroxy-5-phenylpentan-2-one (2e _{FucA}) <i>in vivo</i>	110
D III.2.3	Interim summary I	115
E	Optimization strategies for synthetic pathways <i>in vivo</i>	117
E I	Optimization of pathway performance by complementing flux enhancement strategies	117
E I.1	Synthetic pathway contextuality and interactions between <i>de novo</i> genetic elements	118
E I.2	Cofactor recycling for <i>in vitro</i> (and <i>in vivo</i>) cascades by the engineered <i>GDH_{2xBs}</i>	122
E I.2.1	Purification and characterization of two thermostable GDH mutants	122
E I.2.2	Cofactor dependent reductions of double bonds and carboxylic acids <i>in vitro</i>	124
E I.3	Utilization of host strains with engineered metabolic backgrounds	126
E I.4	A 'hidden reservoir' for reactive aldehyde species: Addressing a competing side reaction and the issue of aldehyde toxicity <i>in vivo</i>	130
E I.5	Once the microbial cell factory is done: Downstream processing by solid phase extraction (SPE) to maximize aldol product yields	137
E I.6	A difficult case: The optimization of <i>de novo</i> pathways employing DHAP-dependent aldolases	138
E I.6.1	Consideration of potential bottlenecks	138
E I.6.2	<i>DhaK</i> and <i>YqaB</i> : An enzymatic 'push and pull' module for pathway optimization	139
E I.6.3	There is always room for improvement: Unsuccessful production of 2e _{FucA}	141
E I.6.4	Interim summary II	144
F	Conclusion and outlook	146
G	Experimental part	150
G I	Materials and methods: Standard molecular biology techniques	150
G I.1	General stock solutions	150
G I.2	Standard media preparations	150
G I.3	Strain cultivation on agar plates	152
G I.4	Permanent culture preparation	152
G I.5	Preparation of <i>E. coli</i> cell-free extracts (CFEs)	152
G I.6	Preparation of <i>E. coli</i> resting cells (RCs)	153
G I.7	Total amount of protein determination by Bradford assay	153
G I.8	Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE)	154
G I.8.1	Buffer and reagent preparations	154
G I.8.2	Microwave staining	155
G I.9	Transformation of competent <i>E. coli</i> cells	155
G I.9.1	Chemical transformation using CaCl ₂	155
G I.9.2	Bulk preparation of RbCl competent <i>E. coli</i> cells	156
G I.9.3	Chemical transformation using RbCl	156
G I.10	Plasmid DNA isolation and quantification	157
G I.11	Restriction enzyme control digestion of plasmid DNA	157
G I.12	Ethanol (EtOH) precipitation of DNA	158
G I.13	Gradient PCR	158
G I.14	Colony PCR	159
G I.15	Protein purification methods	160
G I.15.1	Semi-purification of temperature stable proteins by heat shock (HS)	160
G I.15.2	Purification of 6xHis-tagged enzymes by IMAC	160
G II	Materials and methods: (Advanced) cloning techniques	161
G II.1	Optimized 'Florida' cloning	161

G II.1.1	PCR amplification and purification of target insert DNA	161
G II.1.2	Trimming of target insert DNA	162
G II.1.3	Restriction enzyme (double) digestion of target vector and gel purification	162
G II.1.4	Sticky end ligation of target insert and vector DNA	162
G II.1.5	Chemical transformation	163
G II.1.6	Plasmid assembly verification by colony PCR	163
G II.2	Sequence- and ligation-independent cloning (SLIC) methods	163
G II.2.1	FastCloning (FC)	163
G II.2.1.1	Primer design	163
G II.2.1.2	PCR amplification and processing of target DNA fragments	163
G II.2.1.3	<i>In vivo</i> assembly of DNA fragments	164
G II.2.2	Seamless and ligation-independent cloning extract (SLiCE) approach	164
G II.2.2.1	SLiCE preparation	164
G II.2.2.2	Primer design	164
G II.2.2.3	Preparation of DNA fragments for assembly by SLiCE	165
G II.2.2.4	SLiCE reaction	165
G III	Materials and methods: Biotransformations	166
G III.1	Biotransformations employing CFEs	166
G III.2	Standard screening conditions: Biotransformations using RCs	166
G IV	Strain and enzyme library	167
G V	'Florida' cloning trouble shooting:	
	The establishment of a reliable molecular cloning procedure	170
G V.1	The purification of PCR amplified target inserts by commercial kits	170
G V.2	Gel purification of target vector DNA as a crucial step in molecular cloning	171
G V.2.1	Commercially available kits	171
G V.2.2	Agarose gel disruption and EtOH precipitation of vector DNA	171
G V.3	Testing of ligases from different suppliers	172
G VI	Single enzyme characterizations	173
G VI.1	Characterization of esterases to prepare primary alcohols <i>in situ</i>	173
G VI.1.1	Cloning and expression of <i>BS2</i> , <i>Pfel</i> , and <i>PEST</i>	173
G VI.1.1.1	Cloning and single enzyme expression	173
G VI.1.1.2	Functional screening by pNPA assay and substrate acceptance screening	173
G VI.2	Characterization of ADHs for the oxidation of primary alcohols to aldehydes	174
G VI.2.1	Old friends letting one down: <i>ADH_{Lk}</i> and <i>ADH_{Rr}</i>	174
G VI.2.1.1	Cloning and single enzyme expression	174
G VI.2.1.2	Substrate acceptance screening	175
G VI.2.1.3	Cloning of the <i>adh_{Lk}</i> gene into pTYB21 and expression of N-terminally intein-tagged enzyme	175
G VI.2.2	Being thermostable is not enough: <i>ADH-ht</i>	178
G VI.2.2.1	Cloning, single enzyme expression, and heat shock purification	178
G VI.2.2.2	Substrate acceptance screening	182
G VI.2.3	A robust biocatalyst for the oxidation of secondary alcohols: <i>ADH-A</i>	183
G VI.2.3.1	Cloning and optimization of <i>ADH-A</i> production	183
G VI.2.3.2	Preparation of whole cell lyophilisates	184
G VI.2.3.3	Substrate acceptance screening	184
G VI.2.4	The winner takes it (almost) all: <i>AlkJ</i>	185
G VI.2.4.1	FC of the <i>alkJ</i> gene and single enzyme expression	185
G VI.2.4.2	Substrate scope and substrate profile expansion toward nonnative compound classes	187
G VI.2.4.3	Truncation of the <i>alkJ</i> gene by utilization of encoded <i>PstI</i> restriction sites	187
G VI.3	Characterization of CARs for the reduction of carboxylic acids to aldehydes	188
G VI.3.1	<i>CAR_{Ni}</i> and <i>CAR_{Mm}</i>	188
G VI.3.1.1	Cloning and enzyme expression by autoinduction	188
G VI.3.1.2	Purification of <i>CAR_{Ni}</i> by IMAC	189
G VI.3.1.3	Substrate acceptance screening of <i>CAR_{Ni}</i>	190
G VI.4	Characterization of different aldolases for C–C bond formation	191
G VI.4.1	DHAP-dependent aldolases: <i>FruA</i> , <i>RhuA</i> , and <i>FucA</i>	191
G VI.4.1.1	Cloning and expression of <i>FruA</i> , <i>RhuA</i> , and <i>FucA</i>	191
G VI.4.1.2	Construction of pKA1_ <i>fucA</i> by FC	191
G VI.4.1.3	Assaying <i>FucA</i> activity <i>in vitro</i>	193
G VI.4.2	The dihydroxyacetone (DHA)-utilizing aldolases <i>Fsa1</i> and <i>Fsa1-A129S</i>	194
G VI.4.2.1	Cloning and enzyme expression	194

G VI.4.2.2	Purification of <i>Fsa1-A129S</i> by HS and preparation of lyophilisates	194
G VI.4.2.3	Assaying <i>Fsa1-A129S</i> activity <i>in vitro</i>	195
G VI.5	Characterization of acidic phosphatases and engineered variants	195
G VI.5.1	The mutant acidic phosphatase <i>PhoN-Se V78L</i>	195
G VI.5.1.1	Cloning and expression of <i>PhoN-Se V78L</i> and variants	195
G VI.5.1.2	Release of periplasmic <i>PhoN-Se V78L</i> variants by osmotic shock	197
G VI.5.2	The wild type acidic phosphatase <i>PhoN-Sf</i>	197
G VI.5.2.1	Cloning and expression of <i>PhoN-Sf</i>	197
G VI.5.3	The <i>E. coli</i> phosphatase <i>YqaB</i>	198
G VI.5.3.1	Cloning and expression of <i>YqaB</i>	198
G VI.5.4	Phosphatase activity assays	198
G VI.5.4.1	Photometric pNPP assay	198
G VI.5.4.2	Colorimetric BCIP agar plate assay	198
G VI.6	Characterization of the DHA kinase (<i>DhaK</i>) from <i>C. freundii</i>	199
G VI.6.1	Cloning and enzyme expression	199
G VI.6.1.1	Insoluble enzyme production with pET22b(+) _{<i>dhaK</i>} according to Gottschalk	199
G VI.6.1.2	Soluble enzyme production with pRSETa _{<i>dhaK</i>} according to García-Junceda	200
G VI.6.2	Functional testing by metabolomic analysis	200
G VI.7	Characterization of mutant GDHs from <i>B. subtilis</i>	201
G VI.7.1	Cloning and enzyme expression	201
G VI.7.2	Purification of GDH mutants	201
G VI.7.2.1	Purification by HS	201
G VI.7.2.2	Purification by IMAC	201
G VI.7.3	Photometric NADP ⁺ consumption assay	202
G VI.7.4	Preliminary evaluation of storage conditions	202
G VI.7.5	<i>GDH_{2xBS}</i> as cofactor regenerating enzyme <i>in vitro</i>	202
G VI.8	Characterization of EREDs	202
G VI.8.1	Expression of <i>YqjM</i> from different vectors	202
G VI.8.2	Purification of N-terminally 6xHis-tagged <i>YqjM</i> by IMAC	203
G VI.8.3	(Assymmetric) reductions <i>in vitro</i>	203
G VI.9	Preliminary expression of enzymes for future modular pathway extension	204
G VI.9.1	From aldehyde intermediates to terminal amines: ω-transaminases	204
G VI.9.2	Alternative carboligations by a pyruvate decarboxylase (PDC) from <i>A. pasteurianus</i>	204
G VI.10	Expression of miscellaneous proteins	205
G VI.10.1	Oleate hydratase <i>OhyA2</i>	205
G VI.10.2	Outer-membrane protein <i>AlkL</i>	205
G VII	Construction of vectors for the coexpression of pathway enzymes	206
G VII.1	Plasmids harboring the <i>alkJ</i> and the <i>fsa1-A129S</i> gene in different genetic configurations	206
G VII.1.1	Construction of pKA1_alkJ: <i>fsa1-A129S</i> in operon configuration (pOPE)	206
G VII.1.1.1	Successful pOPE assembly from two linear DNA fragments by SLICE	208
G VII.1.2	Construction of pKA1_alkJ: <i>fsa1-A129S</i> in pseudo-operon configuration (pPOP)	210
G VII.1.2.1	Successful pPOP assembly by FC	211
G VII.1.2.2	Successful POP plasmid assembly from two DNA fragments by SLICE	212
G VII.1.3	Construction of pKA1_alkJ::B001x: <i>fsa1-A129S</i> in monocistronic configurations (pMONx)	213
G VII.1.3.1	Design of bidirectional terminator sequences	213
G VII.1.3.2	Introduction of terminator sequences by Florida cloning utilizing a unique <i>BamHI</i> restriction site	214
G VII.1.4	Validation of vector functionality by protein production analysis and biotransformations	216
G VII.1.4.1	Coexpression studies of <i>AlkJ</i> and <i>Fsa1-A129S</i>	216
G VII.1.4.2	Pathway validation with pPOP and DHA concentration screening	217
G VII.1.5	Changing context of pMON4: Increasing the spacer sequences in pMON5 and pMON6	217
G VII.1.5.1	Construction of pMON5 and pMON6	218
G VII.1.5.2	Testing for the restoration of <i>AlkJ</i> activity in biotransformations	219
G VII.2	Vector construction for the coexpression of <i>AlkJ</i> and the DHAP-dependent aldolase <i>FucA</i>	219
G VII.2.1	Construction of pKA1_alkJ: <i>fucA</i> in pseudo-operon configuration	219
G VII.2.2	Construction of pKA1_fucA::alkJ in pseudo-operon configuration	221
G VIII	Expression of multiple pathway enzymes in a single host cell	223
G VIII.1	Coproduction of <i>AlkJ</i> , <i>Fsa1-A129S</i> , <i>CAR_{Ni}</i> , and <i>Pptase_{Ec}</i>	223
G VIII.2	Coproduction of <i>AlkJ</i> and DHAP-dependent aldolases	223
G VIII.3	Preparation of LEPs coexpressing <i>FucA</i> with other pathway enzymes	224
G VIII.4	Coproduction of <i>AlkJ</i> , <i>FucA</i> , and <i>PhoN-Sf</i>	224

G VIII.5	Coproduction of <i>Fuca</i> , <i>YqaB</i> , and <i>DhaK</i>	225
G IX	Downstream processing by SPE	226
H	Appendix	227
H I	Gene sequences and plasmid maps	227
H I.1.1	Maps of empty vectors routinely used in this thesis	227
H I.1.1.1	pCDFDuet-1	227
H I.1.1.2	pET16b	228
H I.1.1.3	pET21b(+)	228
H I.1.1.4	pET22b(+)	229
H I.1.1.5	pET26b(+)	229
H I.1.1.6	pET28a(+)	230
H I.1.1.7	pETDuet-1	230
H I.1.1.8	pKA1	231
H I.1.1.9	pKK223-3	231
H I.1.2	Esterases	231
H I.1.2.1	<i>B2S</i> : pET28a(+)_ <i>bs2</i>	231
H I.1.2.2	<i>PEST</i> : pET28a(+)_ <i>pest</i>	232
H I.1.2.3	<i>Pfel</i> : pGASTON_ <i>pfel</i>	232
H I.1.3	ADHs	233
H I.1.3.1	<i>ADH-A</i> : pET22b(+)_ <i>adh-a</i>	233
H I.1.3.2	<i>ADH-ht</i> : pET26b(+)_ <i>adh-ht</i>	233
H I.1.3.3	<i>ADH_{LK}</i> : pET21b(+)_ <i>adh_{LK}</i>	233
H I.1.3.4	<i>LK-ADH_{N-int}</i> : pTYB21_ <i>intein-CBD:adh_{LK}</i>	234
H I.1.3.5	<i>ADH_{Rr}</i> : pRR_ <i>adh_{Rr}</i>	235
H I.1.3.6	<i>AlkJ</i> : pKA1_ <i>alkJ</i>	235
H I.1.4	CARs and <i>Pptase_{Ec}</i>	236
H I.1.4.1	<i>CAR_{Ni}</i> : pETDuet-1_ <i>pptase_{Ec}::car_{Ni}</i>	236
H I.1.4.2	<i>CAR_{Mm}</i> : pETDuet-1_ <i>pptase_{Ec}::car_{Mm}</i>	237
H I.1.5	Aldolases	237
H I.1.5.1	<i>Fsa1</i> : pET16b_ <i>fsa1</i>	237
H I.1.5.2	<i>Fsa1-A129S</i> : pET16b_ <i>fsa1-A129S</i>	238
H I.1.5.3	<i>FucA</i> : pKA1_ <i>fucA</i>	238
H I.1.5.4	<i>FucA</i> : pKK223-3_ <i>fucA</i>	239
H I.1.5.5	<i>FruA</i> : pKK223-3_ <i>fruA</i>	239
H I.1.5.6	<i>RhuA</i> : pKK223-3_ <i>rhuA</i>	239
H I.1.6	Phosphatases	239
H I.1.6.1	<i>PhoN-Se V78L_{fl}</i> : pCVD_ <i>phoN-V78L_{fl}</i>	239
H I.1.6.2	<i>PhoN-Se V78L_{fl}</i> : pET26b(+)_ <i>phoN-V78L_{fl}</i>	240
H I.1.6.3	<i>PhoN-Se V78L_{SP}</i> : pET26b(+)_ <i>phoN-V78L_{SP}</i>	240
H I.1.6.4	<i>PhoN-Se V78L_{3xFLAG}</i> : pET26b(+)_ <i>phoN-V78L_{3xFLAG}</i>	240
H I.1.6.5	<i>PhoN-Sf</i> : pET26b(+)_ <i>phoN-Sf</i>	241
H I.1.6.6	<i>YqaB</i> : pCDFDuet-1_ <i>yqaB</i>	241
H I.1.7	Mutant GDHs	241
H I.1.7.1	<i>GDH_{2xBS}</i> : pET28a(+)_ <i>gdh_{2xBS}</i>	241
H I.1.7.2	<i>GDH_{7xBS}</i> : pET28a(+)_ <i>gdh_{7xBS}</i>	242
H I.1.8	Transaminases	242
H I.1.8.1	<i>3FCR</i> : pET22b(+)_ <i>3fcr</i>	242
H I.1.8.2	<i>3GJU</i> : pET22b(+)_ <i>3gju</i>	242
H I.1.8.3	<i>3HMU</i> : pET22b(+)_ <i>3hmu</i>	243
H I.1.8.4	<i>3i5T</i> : pET22b(+)_ <i>3i5T</i>	243
H I.1.8.5	<i>AspFum</i> : pET22b(+)_ <i>aspfum</i>	244
H I.1.8.6	<i>AspTer</i> : pGASTON_ <i>aspter</i>	244
H I.1.8.7	<i>NeoFis</i> : pET22b(+)_ <i>neofis</i>	245
H I.1.8.8	<i>VflH6</i> : pET24a_ <i>vflH6</i>	245
H I.1.9	'Core modules' constructed for the coexpression of pathway enzymes	246
H I.1.9.1	pOPE: pKA1_ <i>alkJ:fsa1-A129S</i>	246
H I.1.9.2	pPOP: pKA1_ <i>alkJ::fsa1-A129S</i>	246
H I.1.9.3	pMON1: pKA1_ <i>alkJ::B0011::fsa1-A129S</i>	247
H I.1.9.4	pMON4: pKA1_ <i>alkJ::B0014::fsa1-A129S</i>	248
H I.1.9.5	pMON5: pKA1_ <i>alkJ::B0014_{so}::fsa1-A129S</i>	248

H I.1.9.6	pMON6: pKA1_alkJ::B0014 ₁₀₀ ::fsa1-A129S	248
H I.1.9.7	pJRE: pKA1_fucA::alkJ	249
H I.1.10	Other enzyme-coding genes and gene modifications	250
H I.1.10.1	<i>DhaK</i> : pRSETa_dhaK	250
H I.1.10.2	BLAST analysis and reannotation of pET22b(+)_dhaK	250
H I.1.10.3	<i>PDC_{Ap}</i> : pET22b(+)_pdc _{Ap}	251
H I.1.10.4	<i>AlkJ_{trnc}</i> : pKA1_alkJ _{trnc}	252
H I.1.10.5	<i>YqjM</i> : pHT_yqjM	252
H I.1.10.6	<i>YqjM</i> : pET22b(+)_yqjM	253
H I.1.11	Cosmids	253
H I.1.11.1	pGEc47	253
H II	List of primer pairs	255
H III	Publications resulting from this thesis	257
H IV	Curriculum vitae	258
H V	List of abbreviations	261
H VI	References	265

Acknowledgements

Since conducting and writing this thesis would not have been possible without the many wonderful people near, far, in, and outside the lab, I would like to fill the next few paragraphs with my heartfelt thanks to all of them.

First, I express my gratitude to Marko D. Mihovilovic for giving me the opportunity to conduct my dissertation in his group at the Institute for Applied Synthetic Chemistry at the TU Wien and Florian Rudroff for initially getting me on board of his first FWF project (P 24483 B21) and his supervision. To both my special thanks for allowing me to develop and follow my ideas and grow as scientist.

I thank the COST Action Systems Biocatalysis (CM 1303) and the ÖFG – Österreichische Forschungsgemeinschaft for financial support.

I would like to thank the examiners of my thesis, Roland Ludwig and Uwe T. Bornscheuer (UTB), for their commitment and their expertise. To Uwe my sincere thanks for having me for a short scientific stay in his group at the Ernst-Moritz-Arndt University Greifswald. Uwe, I really enjoyed my time in Greifswald and consider it as vital for the success of my dissertation. So, once again, thank you. In this regard, the biggest thanks to Christin Peters (CPE), who not only picked me up at the Greifswald train station and showed me around but taught me many of the cloning tricks used in this work. Many thanks to Lilly, Jan, Andy, Lisa, and all former and present members of the Bornscheuer group who made me feel very much at home in Greifswald and their support.

Many thanks to all the great scientists who kindly donated plasmids, engineered strains, and knowledge that proved to be more than helpful: Bruno Bühler (pGEc47), Uwe T. Bornscheuer (the ADHs from *L. kefir* and *R. ruber*, esterases, and ω -TAs), Wolf-Dieter Fessner and Michael Kickstein (all the different aldolases), Kristala L. Prather (*E. coli* K12 RARE), Dörte Rother (pET22b(+)_{pdCAp}), and Margit Winkler (CARs). Margit and Daniel 'travelling light' Schwendenwein, thank you for problem-solving discussions, your help, the one and the other beer, and for being amazing cooperation partners. Sandy Schmidt, thanks for meeting up in Amsterdam and your friendship.

To Michael 'Michl' Schnürch my sincere thanks for his open-door policy, support, and keeping cool and Christian Stanetty for his good spirit, Sabine Stiedry, Flo Untersteiner, Emilia, Irena, Thomas Seebauer, and all the other good guys of the IAS for their daily (organisational) support. My special thanks to Mrs Halasz for autoclaving tons of bio-waste over the years.

My wholehearted thanks go to Christian 'the generous from BH' Derntl and the Robert Mach group for working together across the different institutes. Christian, thanks for lending me equipment BI was lacking, your words of scientific advice, and the (occasionally needed) words of comfort.

Niko, thanks for working together even before I joined the Mihovilovic group and for making nightshifts in the lab most enjoyable. – I truly missed you since you have been gone. Many thanks to my superb colleagues and friends Anna-banana-fo-fana (ARE), who will be a married woman soon, David 'Südpazifik' Siebert, who will be a married woman, too, soon, Marcello 'the stylish and strong' Entner (MAE), Dani 'Sway with me!' Cintulová, Maria, Blanca, Hamid, Sebastian, and Thomas 'Thomsibomsi' Kremsmayr for the many hugs to start the days. I thank Elise and Magda and the whole PostSV beach volleyball group for ending (difficult) Mondays. Drasi, thanks for supplying coffee, and all past and present coworkers of the Mihovilovic group for the nice working atmosphere.

I would like to express my gratitude to my students Jakob Reiterlehner (JRE), Sophie Knoll (SCK), Birgit Walder (BWA), Carina Wokurek (CWO), Patricia Wolf (PWO), Katja Eibensteiner (KAE), Madlen Mollik (MMO), Kathi Figl (KFI), Franziska Karner (FRK), Katharina Lasic (KKL), and Alexandra Wanka (AWA). Had it not been for their dedication and enthusiasm, the complexity of this thesis could not have been achieved and working on it would not have been as much fun.

I am happy that I got the chance to work together with Thomas ‘Babysteps’ Wiesinger (TWI) and Sofia ‘Katze!’ Milker (SMI) on this challenging project. My sincere thanks to Sofia, who not only measured what felt like myriads of HPLC samples but stuck with me until the end and shared the joys and sorrows of the ‘Endstation Sehnsucht’ office with me. I also would like to thank my dear friend Leticia for her support, her scientific brilliancy, and all the sweet treats.

Patricia ‘Mon cher, Trish!’ Schaaf (PSC), my significant laboratory other from the very first day, I actually cannot thank you enough for your company and support in the last years. Thank you! <3

On the other side of the lab door, beyond the walls of the institute building, and the fences of the campus, there were always people waiting for me after long days (and nights) to pick, back, and lift me up: my dear friends.

A big shoutout to DI René ‘René René!’ Ziegler for being there for me, Dr Alex ‘Superfreund’ Hartmann for creating precious memories, Peter ‘die Insel der Freude’ Hillebrandt for excessive popcultural crossreferencing and endless laughter, Peter ‘Beinknospe’ Kraus for being the *other* nose, Firas for dancing with me through the nights, Nielo and Romu for their great taste, Albert for walking with me in *that* shoes, and all charter members of the Salon Limoncello. Lisa ‘lil-el-bow-wow’ Holzinger, my sincere thanks for reaching out for me.

I thank Lorenz for *gigantimous* gestures, Martin Walanka for tender moments, Willy ‘Michael’ Mutzenpachner for exchanging heart emojis, Gerhard ‘Gellert’ Butter for literally lifting me up, Christopher for his wit and the wonders of ‘Ulli Sima holding things’. My special thanks go to the Neigungsgruppe Vergnügungspark: Matthias ‘Aston’ Matters, Romana ‘Superfreundin. Romance inclusive.’ Pachner-Schmidthaler, and Martin ‘Schütze. Aszendent: Stier.’ Finnland. You became indispensable friends, everyone for a different but exceptionally good reason, and I would like to thank you. Martin, my extra thanks for not only being a beloved friend but taking me on an amazing journey through the Nesterval universe. In this regard, my thanks to the whole Nesterval cast for the most incredible time.

To Florian Glinserer my *cakefelt* thanks for spending hours with me at Gregor’s. As speaking of: I would like to thank Gregor, Dominik, and Gerlie for making Gregor’s my favorite café. I also thank the Neigungsgruppe Tortenschlacht: Thomas, Patrick, Mathias, Sven, Sebatsian, and Volkmar.

More than many thanks to Scott and Nina for travelling with me far east, Johann and Alex for travelling with me even further east, Erwin for the best times in Paris, Bordeaux, Moliets-et-Maa, and helping me to find Palermo on the map, Rita for always picking up her phone and helping me to keep my sanity, Claudia ‘CJ’ Juno and Andrea ‘Abrakadabra!’ Achleitner for ‘Eh-eh-eeeeh!’, Claudia and Christoph Six for their hospitality and sharing most memorable evenings together, Christoph Strauß and Josef ‘the charmer’ Wehninger for cheering me up.

I thank Mag (FH) Thomas Horvath and Dr Thomas Reiter – the major Ts in TnTnT – for having breakfast, dinner, or any other meal with me, being good listeners, and backing me up. I thank Christoph Hofbauer and Larissa Wiens for providing safe harbors when I was at a personal low.

To all of you: Thank you for being superb friends.

I thank Hansjörg, Georg, and Alexis – all ‘fire’ – and Maria ‘Wolfsspaßbergersøn’ Wolfsberger for being around for more than 10 years now.

To the Kleeblatt: Leni, Sandy, and Kathi thanks for being constant friends for more than 20 years. I hope you know how much I cherish our friendship.

Finally, I would like to thank Andi for everything he has done and Judith and Ludwig for keeping in touch. I thank my family, my sister Iris and my parents Veronika and Reinhold for their strength and their unconditional support, my grandmother Maria, who I miss heaps, and Martin, who will always have a place in my heart.

Abstract

Artificial enzyme cascades are constructed by the combination of biocatalysts that are metabolically unrelated in nature. Rapidly increasing numbers of available biocatalysts and cross-disciplinary efforts such as biocatalytic retrosynthesis accelerated the design of (bio)synthetic routes with increasing complexity to produce value-added chemicals. Well-established genetic regulatory elements (e.g., promoters) for balancing enzyme production, complementing substrate channeling approaches, and the engineering of enzymatic host backgrounds by gene knock-out target different molecular levels to implement and optimize pathways in whole cells.

The introduction of nonnative enzymes may interfere with the metabolic environment in hosts like *Escherichia coli* (*E. coli*), which might impair the carbon flux through the synthetic pathway *in vivo*. Particularly, unexpected interactions between different synthetic genetic elements are often underestimated contextual issues in pathway design and a current challenge. Consequently, this thesis aims at the application of synthetic enzyme cascades *in vivo* and the resolution of both compositional and host context dependencies by complementing flux enhancement strategies to maximize product titers.

The biocatalytic retrosynthetic approach pursued in this thesis, revealed two distinct pathway designs to produce polyhydroxylated compounds. Both involve the oxidation of primary alcohols to the corresponding aldehydes and subsequent carbonylation catalyzed by an aldolase. Regarding the second cascade step, aldols can be either produced from extracellularly added aldol donor molecules such as (di)hydroxyacetone [(D)HA] or by hijacking glycolytic DHA phosphate (DHAP). The latter will tightly interconnect the *de novo* pathway and the central carbon metabolism of *E. coli* and circumvent the lability of DHAP *in vitro*. The implementation of a phosphatase and stereocomplementary DHAP-dependent aldolases dephosphorylates the phosphorylated intermediate adduct to shift the reaction equilibrium and provides access to aldol products in different configurations, respectively.

Subsequent screening of the biocatalytic toolbox identified *AlkJ*, an alcohol dehydrogenase from *Pseudomonas putida*, as an efficient biocatalyst for the *in situ* production of reactive aldehyde acceptors. For the first cascade design, the (D)HA-dependent aldolase *Fsa1-A129S* from *E. coli* was most suitable yielding aldol adducts with (3*S*,4*R*) configuration, whereas, for the second pathway, the DHAP-dependent *FucA* (*E. coli*) was selected to produce (3*R*,4*R*) polyhydroxylated compounds and phosphatases from different microbial hosts studied.

Up-to-date sequence- and ligation-independent cloning techniques were successfully applied to assemble pathway modules in different genetic architectures. Compositional context was improved by the integration of multiple genetic (transcriptional) regulators including (synthetic) terminators to balance pathway enzyme production *in vivo*. The cellular host environment severely interfered with the *in situ* preparation of (cytotoxic) aldehyde intermediates by the rapid reduction to the alcoholic substrates and the irreversible metabolization to the corresponding carboxylic acids. These context issues were addressed by the utilization of highly engineered strains such as *E. coli* RARE and simply by the introduction of a reversing enzymatic activity, respectively, to reroute the carbon flux from the carboxylate sink toward cascade aldehyde intermediates. The latter employed a carboxylic acid reductase from *Nocardia iowensis* (*CAR_{Ni}*). In the presence of *AlkJ* and *CAR_{Ni}*, reactive aldehyde species were quickly interconverted between the corresponding alcohols and carboxylates. Consequently, aldehydes were equilibrated below nonviable, yet freely available concentrations for subsequent aldol reaction. This so far neglected strategy establishing a 'hidden reservoir for reactive aldehyde species' increased cell viability and could address the issue of aldehyde toxicity and persistence *in vivo*.

Although aldols could not be synthesized *via* the DHAP-dependent cascade, crucial bottlenecks such as insufficient intracellular DHAP concentrations were identified by metabolomic analysis. To compensate this

'parasitic' interaction, a DHA kinase from *Citrobacter freundii* was successfully studied in this thesis and offers an optimization strategy for future applications.

The research conducted in this thesis not only designed, assembled, implemented, and optimized an artificial biosynthetic pathway consisting of up to three metabolically unrelated enzymes (*AlkJ*, *CAR_{Ni}*, *Fsa1-A129S*). The 'hidden aldehyde reservoir' approach in combination with a refined solid phase extraction purification, tackling the issue of notoriously low yielding aldol reactions *in vitro*, demonstrated the applicability of the developed system and synthesized structurally different aldols from the donor molecules HA and DHA in up to 91% isolated yields in short reaction times in living cells.

Kurzfassung

Die Kombination von Enzymen aus verschiedenen metabolischen Kontexten führte zur Etablierung künstlicher Stoffwechselwege. Das sich ständig erweiternde Portfolio zur Verfügungen stehender Biokatalysatoren ermöglichte die Entwicklung komplexer, enzymatischer Syntheserouten zur Herstellung von Fein- und Bulkchemikalien. Für die Optimierung von Ganzzellsystemen wie *Escherichia coli* (*E. coli*) kommen genetische Regulatoren wie Promotoren zur Verbesserung der Enzymproduktion oder die gentechnische Veränderung von Bakterienstämmen durch das gezielte Ausschalten von Genen zum Einsatz.

Die Einbringung von nicht-nativen Enzymkaskaden kann nicht nur die Viabilität von *E. coli* negativ beeinflussen. Unterwünschte Interaktionen zwischen genetischen Elementen des künstlichen Stoffwechselweges oder des Stoffwechselweges mit dem metabolischen Kontext von *E. coli* minimieren den Kohlenstofffluss und – folglich – Produktausbeuten. Die Behebung dieser oft unterschätzten Kontextprobleme bei der Etablierung von Enzymkaskaden *in vivo* ist Hauptziel dieser Arbeit.

Die (bio)retrosynthetische Analyse pharmazeutisch relevanter Zuckermoleküle ergab zwei unterschiedliche Wege, polyhydroxylierte Verbindungen zu synthetisieren. Beide oxidieren primäre (aromatische) Alkohole zu den entsprechenden Aldehyden, die durch Aldolase-katalysierte Aldolreaktion zu den Zielverbindungen umgesetzt werden. Für die C–C-Knüpfung können extrazelluläre Donormoleküle wie (Di)hydroxyacetone [(D)HA] oder intrazelluläres DHA-Phosphat (DHAP) verwendet werden. Die Implementierung DHAP-abhängiger Aldolasen mit unterschiedlicher Stereospezifität verknüpft die künstliche Enzymkaskade nicht nur mit dem zentralen Kohlenstoffmetabolismus der Zelle über die Glykolyse; die Labilität von DHAP *in vitro* kann so umgangen werden und ermöglicht die Synthese von Aldolen mit unterschiedlicher Konfiguration. Die irreversible Dephosphorylierung von phosphorylierten Zwischenprodukten durch eine Phosphatase verschiebt schließlich das Reaktionsgleichgewicht in Richtung der gewünschten Aldole.

Die Evaluierung verschiedener Biokatalysatoren führte zu *AlkJ*, einer Alkoholdehydrogenase aus *Pseudomonas putida*, für die *in situ* Produktion von reaktiven Aldehyden, und zwei Aldolasen aus *E. coli*, der (D)HA-abhängigen *Fsa1-A129S* und der DHAP-abhängigen *FucA*. Erstere liefert polyhydroxylierte Verbindungen mit (3*S*,4*R*) und zweitere mit (3*R*,4*R*) Konfiguration. Im Zuge dieser Arbeit wurden Phosphatasen aus unterschiedlichen Mikroorganismen charakterisiert und implementiert.

Moderne sequenz- und ligation-unabhängige Klonierungsmethoden ermöglichten die Assemblierung von Enzymmodulen in verschiedenen genetischen Konfigurationen. Durch die Implementierung transkriptioneller Regulatoren wie (synthetischer) Terminatoren konnte die Produktion von *AlkJ* und *Fsa1-A129S* *in vivo* deutlich verbessert werden. In der zellulären Umgebung von *E. coli* jedoch wurden die *in situ* produzierten Aldehyde rasch verstoffwechselt. *E. coli* exprimiert eine Vielzahl von Enzymen, um sich vor oxidativem Stress zu schützen, und wandelt deshalb Aldehyde zu den entsprechenden Alkoholen um oder oxidiert sie irreversibel zu Carbonsäuren. Zur Lösung dieses Kontextproblems führte unter anderem der Transfer der künstlichen Enzymkaskade in einen geeigneten, genetisch optimierten Stamm wie *E. coli* RARE oder aber die Erweiterung der Kaskade um ein Enzym, das die Bildung von Carbonsäuren umkehren kann. Die Implementierung einer Carbonsäurereduktase von *Nocardia iowensis* (*CAR_{Ni}*) konnte die irreversible Nebenproduktbildung erfolgreich umkehren und führte, zusammen mit der enzymatischen Aktivität von *AlkJ*, zu einer Gleichgewichtseinstellung zwischen primären Alkoholen, Carbonsäuren und reaktiven Aldehyden. Aldehydkonzentrationen wurden so unter das Toxizitätslevel gesenkt, standen jedoch weiterhin dem nächsten Kaskadenschritt, der Aldolreaktion, zur Verfügung. Die Anwendung dieses Zweienzymsystems erhöhte die Viabilität von *E. coli* und stellt eine komplementäre Strategie für das „In-Schach-Halten“ reaktiver Intermediate wie Aldehyde dar.

Obwohl die Produktion von Aldolen über die DHAP-abhängige *FucA* und den entsprechenden künstlichen Stoffwechselweg nicht erfolgreich war, wurden wichtige, *E. coli*-abhängige Faktoren identifiziert. So reicht der intrazelluläre Pool an DHAP nicht aus, um das Gleichgewicht erfolgreich in Richtung der Aldolbildung zu verschieben. Für die *in situ* Produktion von DHAP wurde die DHA-Kinase aus *Citrobacter freundii* evaluiert und verspricht Potential für eine Optimierung dieses synthetischen Stoffwechselweges.

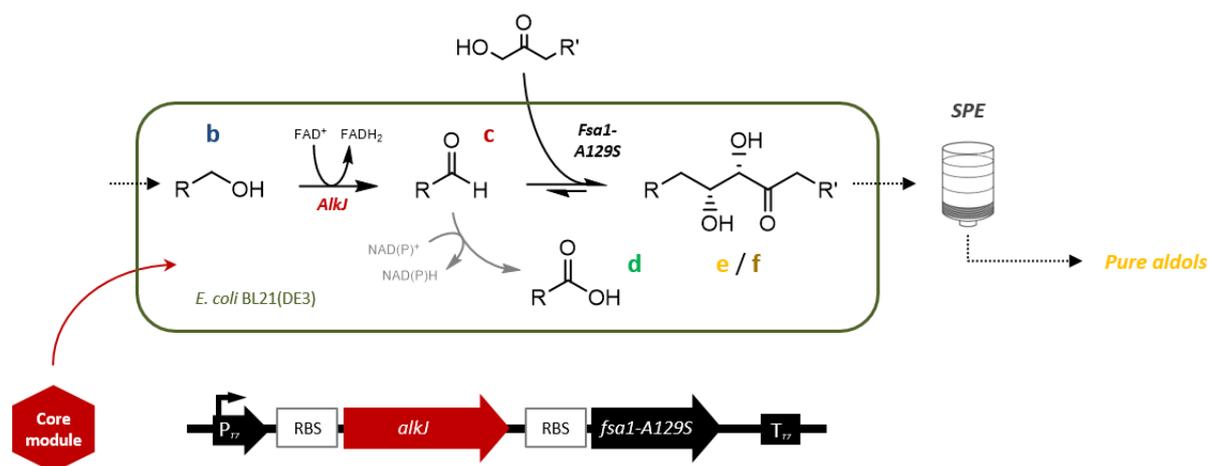
Mit der Kombination dreier Enzyme (*AlkJ*, *CAR_{Ni}*, *Fsa1-A129S*) aus verschiedenen metabolischen Kontexten und der Etablierung einer synthetischen Enzymkaskade in *E. coli* wurde das Ziel dieser Dissertation erreicht. In Kombination mit einer verfeinerten Festphasenextraktion zur Aufreinigung polyhydroxylierter Verbindungen, konnten die isolierten Ausbeuten strukturell unterschiedlicher Aldole mit den Donormolekülen HA und DHA in kurzen Reaktionszeiten auf bis zu 91% gesteigert werden und übertrifft dabei etablierte *in vitro* Systeme.

A Synthetic enzyme cascade schemes

All compounds prepared or used as starting materials in this thesis are numbered in bold Arabic numerals. Since this work aims at the development of artificial cascade reactions, substrates entering the cascade will successively converted *via* intermediate cascade steps to the target compounds. To emphasize this, the cascade step is designated with letters after the corresponding substrate compound number (e.g., **1a** for the 1st cascade step, **1b** for the 2nd cascade step, and so forth).

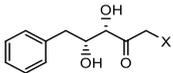
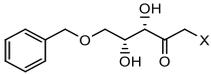
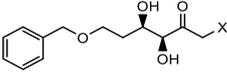
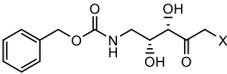
Noteworthy, *in vivo* cascades start from the primary alcohols (**b**) and not the more lipophilic esters (**a**) since the alcoholic substrates can freely pass the cellular membrane of *E. coli*.

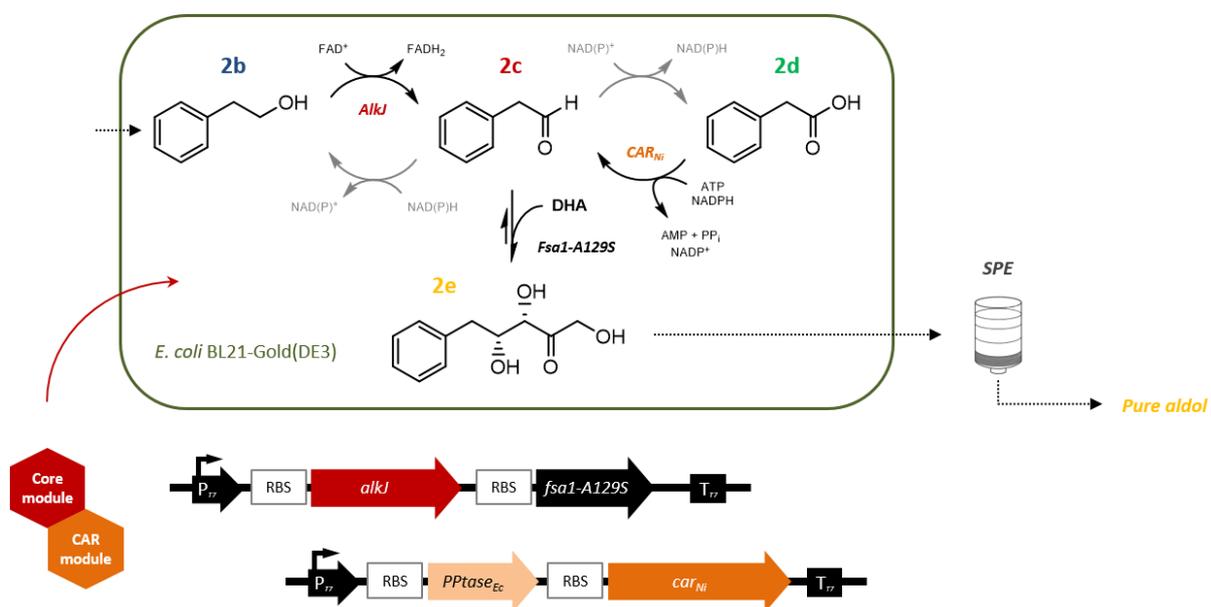
A I Artificial metabolic pathways producing polyhydroxylated compounds *in vivo*



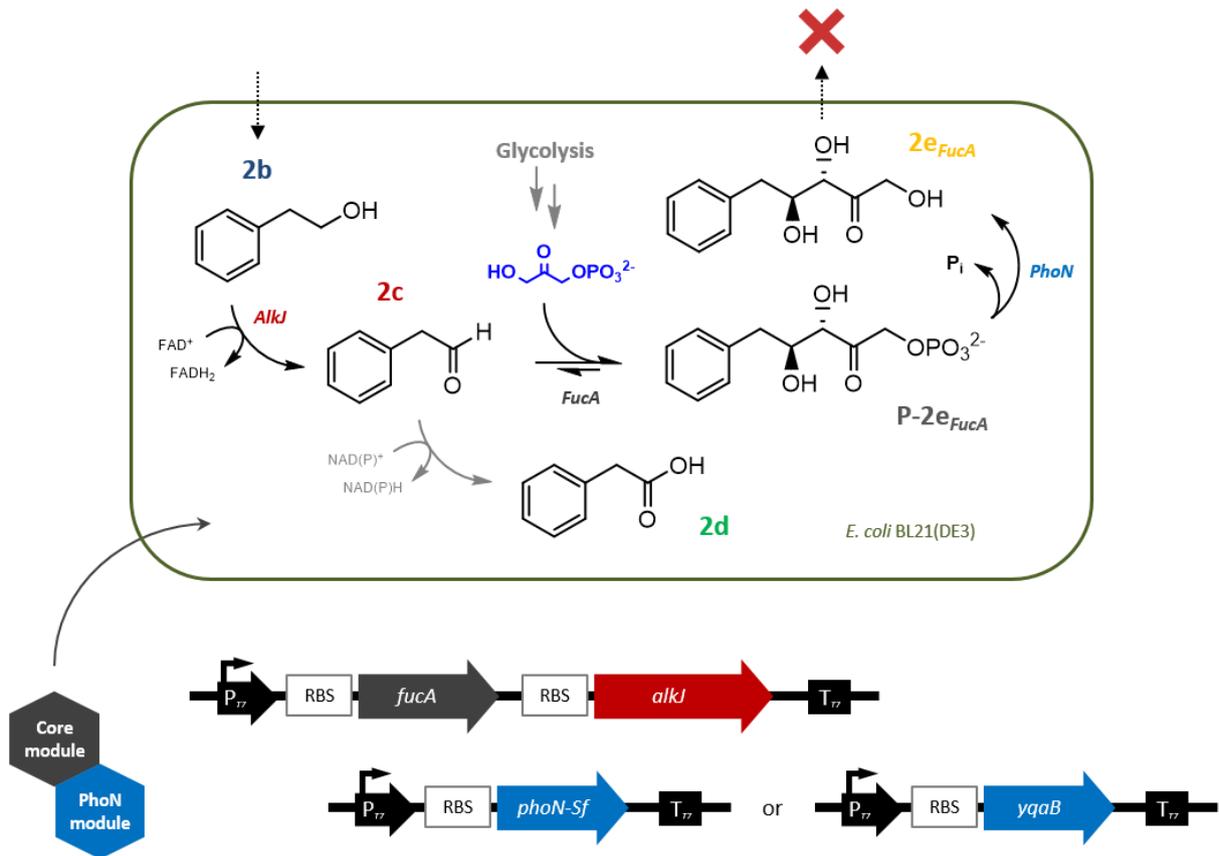
Scheme A-1 Mini-pathway for the successful production of target polyhydroxylated compounds. Pathway flux in black; competing endogenous activities in grey. Enzymatic core module containing pKA1_alkj::fsa1-A129S in pseudo-operon configuration (pPOP). Synthesis: Stationary phase *E. coli* BL21(DE3) whole cell biocatalyst coexpressing *AlkJ* and *Fsa1-A129S*, 1% (w/v) glucose, 5 mM substrate (**b**), 5% (v/v) ACN, 20 eq aldol donor (DHA, R' = OH, compounds **e**; HA, R' = H, compounds **f**), 25°C (250 rpm). Downstream aldol (**e** or **f**) purification by SPE.

Table A-1. Isolated yields of target aldol compounds

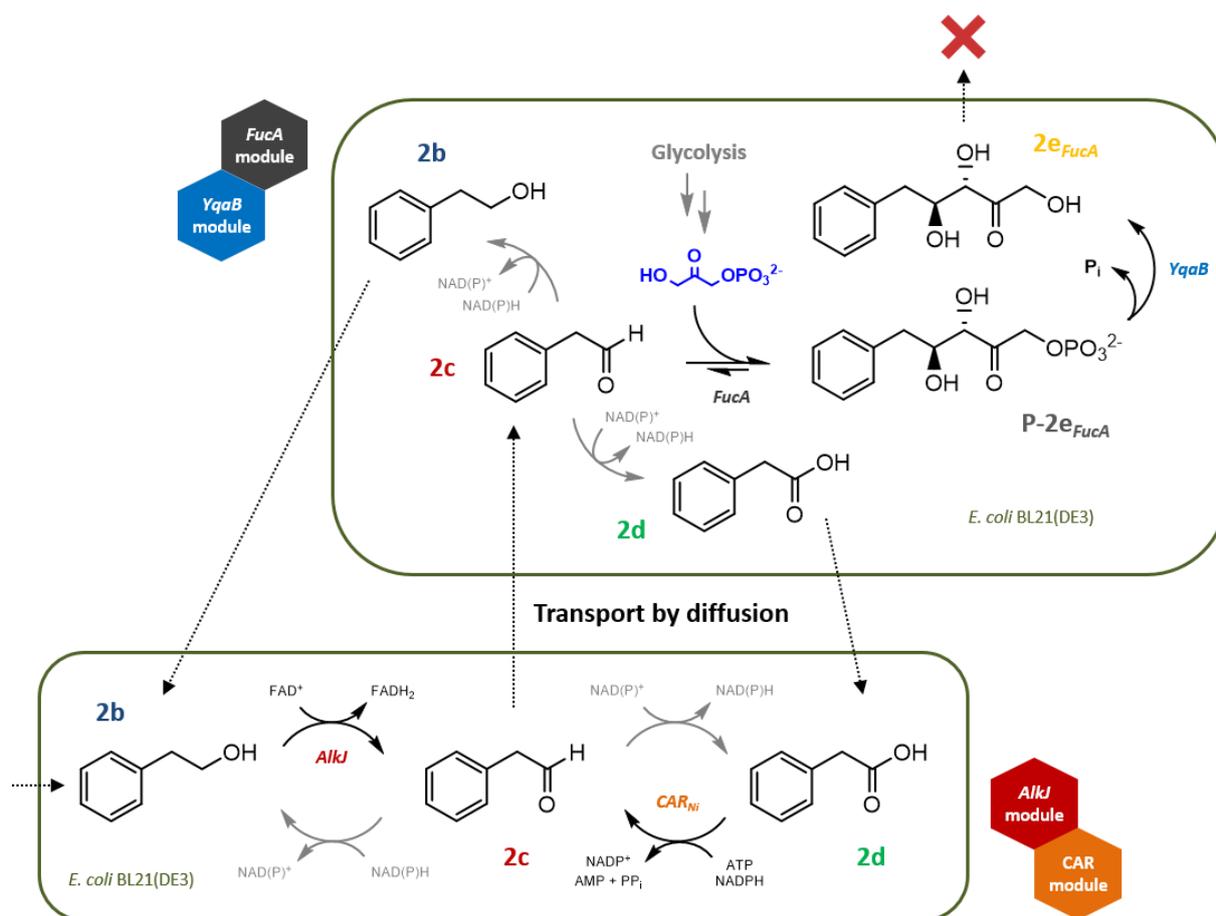
Products	Substrate	Donor	Isolated yields [%]
 (2e-f)	2b	DHA (e)	78
		HA (f)	70
 (4e-f)	4b	DHA (e)	60
		HA (f)	89
 (5e-f)	5b	DHA (e)	64
		HA (f)	61
 (6e-f)	6b	DHA (e)	91
		HA (f)	83



Scheme A-2 'Hidden aldehyde reservoir' approach. Pathway for the production of **2e**. Converged pathway flux in black; endogenous enzyme activities in grey. Enzymatic modules containing pKA1_alkJ::fsa1-A129S in pseudo-operon configuration (pPOP) and pETDuet-1_Pptase_{Ec}::car_{Ni} in pseudo-operon configuration. Synthesis: Stationary phase *E. coli* BL21(DE3) whole cell biocatalyst coexpressing *AlkJ*, *Fsa1-A129S*, *CAR_{Ni}*, *Pptase_{Ec}* (omitted for clarity), 1% (w/v) glucose, 5 mM **2b**, 5% (v/v) ACN, 20 eq DHA, 25°C (250 rpm). Downstream purification of **2e** by SPE: 78% isolated yield in 2 h.



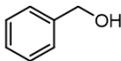
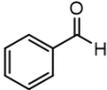
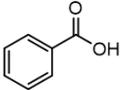
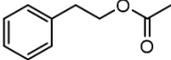
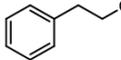
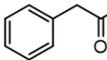
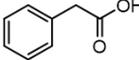
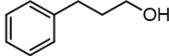
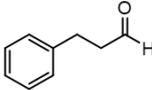
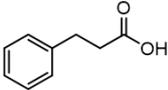
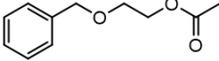
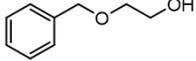
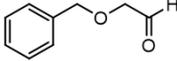
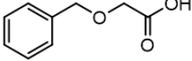
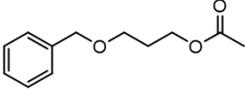
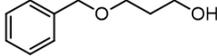
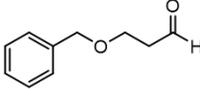
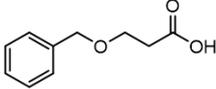
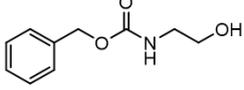
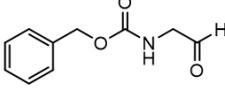
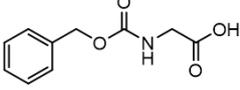
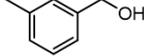
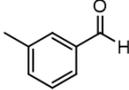
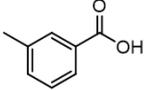
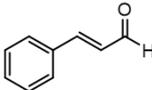
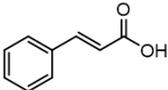
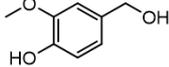
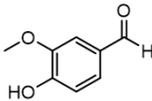
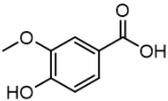
Scheme A-3 Unsuccessful production of $2e_{FucA}$ via a DHAP-dependent pathway. Pathway flux in black; endogenous fluxes in grey. Enzymatic modules containing $pKA1_{fucA}::alkJ$ in pseudo-operon configuration and a phosphatase (*PhoN-Sf* or *pCDFDuet-1_yqaB*). *E. coli* BL21(DE3) whole cell biocatalyst coexpressing *AlkJ*, *FucA*, *PhoN-Sf* or *YqaB* tested under various conditions and module combinations.



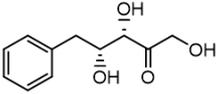
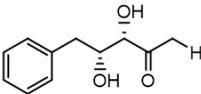
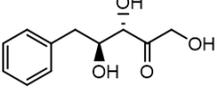
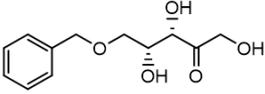
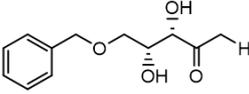
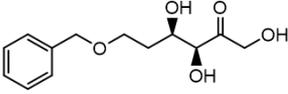
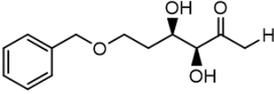
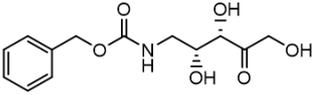
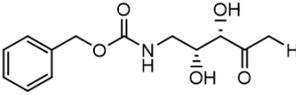
Scheme A-4 Unsuccessful 'mixed culture approach' for the production of $2e_{FucA}$. Pathway flux in black; endogenous fluxes in grey. Separated enzymatic modules consisting of pKA1_alkJ and pETDuet-1_PPtase_{Ec}::CAR_{Ni} and pKA1_fucA and pCDFDuet-1_yqaB. Biotransformations: Late exponential phase *E. coli* whole cell biocatalysts expressing the corresponding modules, 1% (w/v) glucose, 5 mM **2b**, 5% (v/v) ACN, 25°C (250 rpm).

B Compound library

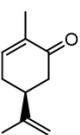
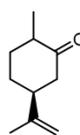
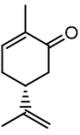
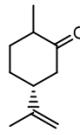
B I Cascade substrates and intermediates: Aromatic esters, primary alcohols, aldehydes, and carboxylic acids

#	Esters a	Alcohols b	Aldehydes c	Carboxylic acids d
1				
2				
3				
4				
5				
6				
7				
8				
9				

B II Cascade products: Aromatic aldol adducts

#	DHA adducts e	HA adducts f	Dephosphorylated DHAP adduct e _{FucA}
2			
4			
5			
6			

B III Metabolic background studies in *E. coli*: ERED substrates and products

#	Substrates	#	Products
10		11	
12		13	
14		15	
16		17	

C Introduction

C I Out of a test tube: The evolution of biocatalysis

Enzymes and natural metabolic pathways had been used by mankind in fermentation processes including baking, brewing, and the production of vinegar long before cells, proteins, and their functions were explored (**Figure C-1**). The anaerobic fermentation in yeast breaks down starch into monosaccharides and, ultimately, ethanol (EtOH) and carbon dioxide (CO₂) and was first scientifically described in the early 19th century. In 1896, Eduard Buchner performed the first cell-free fermentation with an extract from *Saccharomyces cerevisiae* (*S. cerevisiae*) converting sugars into alcohol in a test tube. His conclusion that the reduction was carried out by enzymes, marked the beginning of modern biocatalysis.^[1-3] Amongst the first nonnative substrates converted biocatalytically was benzaldehyde in the presence of hydrogen cyanide and a plant extract by Ludwig Rosenthaler yielding (*R*)-mandelonitrile (**Figure C-1**).^[4-5] Since then, the concept of biocatalysis has developed from converting simple molecules by (isolated) native enzymes in single-step reactions to engineered metabolic routes producing value-added (nonnatural) compounds by the application of multi-enzyme cascades (**Figure C-1** and **Figure C-3**).^[6-7]

Cascade type reactions benefit from the cooperative effect of multiple biocatalysts with their inherently high chemo-, regio- and stereoselectivity and the omission of intermediate isolation (**Figure C-2**).^[7-9] Although *in vitro* applications can be optimized by simply changing reaction parameters (e.g., substrate and enzyme concentrations, temperature, or pH), they consist only of a few enzymes and cannot compete with the efficiency of natural metabolic pathways.^[6, 10] Cell-free systems require the preparation of biocatalysts, which involves the (heterologous) expression, isolation, and purification of enzymes from host cells. Many biocatalysts, especially redox enzymes, depend on cofactors, which must be either added in stoichiometric amounts or regenerated by recycling systems.^[10-12] Both render *in vitro* cascades uneconomical and add complexity to a system that started out being facile (**Figure C-2A**). In comparison, living cells produce amino acids from simple nutrients like glucose and ammonium and synthesize enzymes and other proteins *via* the coordinated processes of transcription and translation.^[13] The host metabolism supplies and recycles cofactors and biocatalysts usually are more stable in the cellular environment of the host cell (**Figure C-2B**).^[6, 10-11]

In the beginning 1980s, the limited stability of enzymes *in vitro* had been overcome by immobilization techniques^[14-15] and emerging structure-based protein engineering tools were employed to substantially increase the tolerance of biocatalysts toward elevated temperatures or organic solvent concentrations. Soon, protein engineering was applied to extend the substrate scope of enzymes, which enabled the synthesis of unusual synthetic intermediates (e.g., ethyl-(*R*)-4-cyano-3-hydroxy butyrate for the commercial production of atorvastatin,^[16] **Figure C-1** and **Figure C-22**).^[17] Starting in the 1990s, pathway engineering tools were still in their infancy, Frances Arnold, Pim Stemmer, and others pioneered protein engineering methods mimicking an *in vitro* version of Darwinian evolution to efficiently modify biocatalysts. These techniques, now commonly known as directed evolution, equipped enzymes with remarkably new features to biocatalytically produce pharmaceutical intermediates and fine chemicals (**Figure C-23–Figure C-26**).^[6, 17]

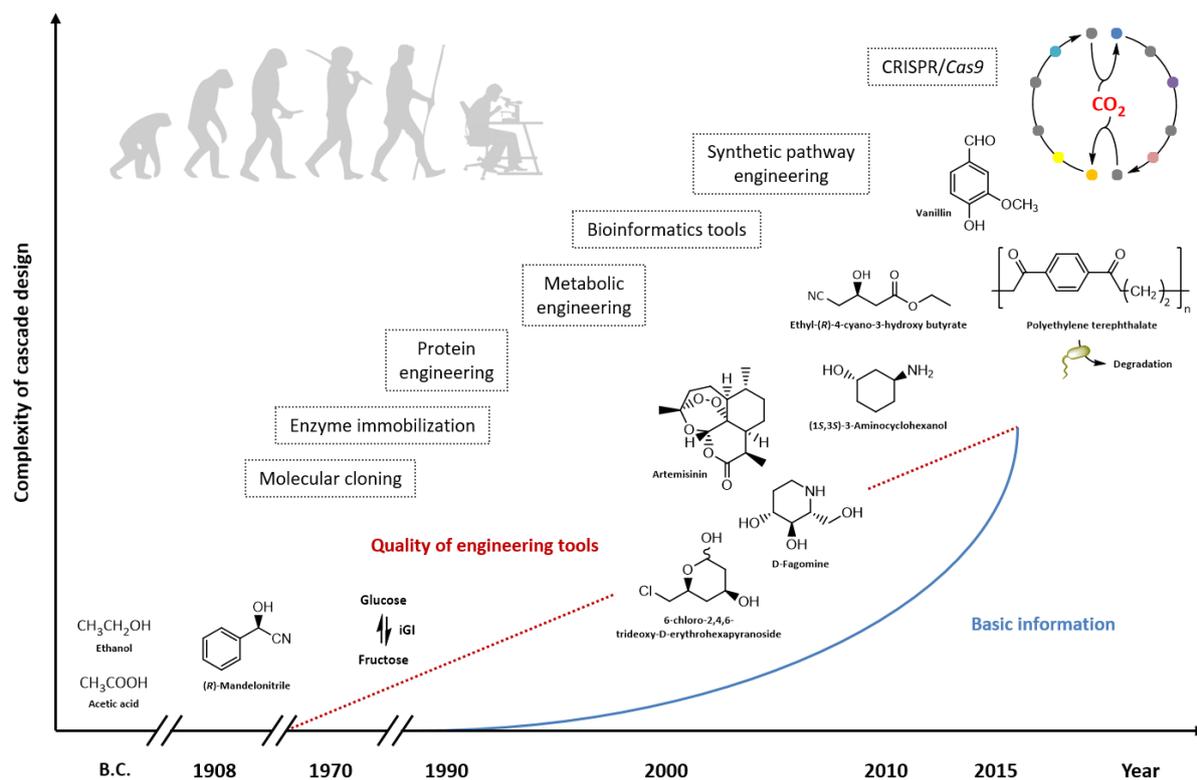


Figure C-1. The evolution of biocatalysis. From fermentations used for millennia and the start of modern biocatalysis with the synthesis of (*R*)-mandelonitrile,^[4] the field of biocatalysis has evolved to today's application of multi-enzyme cascades, extensively engineered enzymes, and whole cell biocatalysts in industrial applications.^[6, 17] This evolution was accompanied by an almost exponential gain of knowledge and constant technological innovations (dotted boxes): Molecular cloning marked the beginning of recombinant DNA technology.^[18-19] Immobilization of enzymes increased stability *in vitro* and was employed in the production of fructose from glucose by an immobilized glucose isomerase (iGI), for example.^[15] Protein engineering techniques highly accelerated the improvement of enzyme properties and led to the realization of synthetic routes to pharmaceutical intermediates as well as products (e.g., 6-chloro-2,4,6-trideoxy-D-erythrohexapyranoside,^[20] artemisinin,^[21] D-fagomine,^[22] ethyl-(*R*)-3-cyano-3-hydroxy butyrate,^[16] or (1*S*,3*S*)-3-aminocyclohexanol^[23]). Metabolic and synthetic pathway engineering strategies are supported by bioinformatics tools and enhanced, for example, the production of vanillin from ferulic acid in whole cells.^[24] Current research progress projects toward the efficient reprogramming of genomes, for example, by CRISPR/*Cas9*,^[25] the (computational) *de novo* design of enzymes,^[17] the assembly of sophisticated metabolic circuits accessing CO₂ as valuable carbon feedstock,^[26] and the identification and optimization of novel (degradation) pathways from metagenomes.^[27] The evolution of scientists' (top left) was adapted from G. R. Scott (McMaster University, CA, 2017).

Advances in DNA technologies and in bioinformatics have provided essential support to the rapidly growing field of biocatalysis. They have promoted the discovery of novel enzymes from natural resources such as metagenomes^[28] and accelerated the redesign of existing biocatalysts. Reduced costs for DNA synthesis and high-throughput screenings speeded up the development of tools for synthetic pathway design and engineering.^[17] The elucidation of whole genome sequences of organisms including *Bacillus subtilis* (*B. subtilis*), *Escherichia coli* (*E. coli*), and *S. cerevisiae* has not only led to the evolution of the *omics* disciplines transcriptomics, proteomics, metabolomics, and fluxomics; it offered comprehensive (microbial) metabolic network information that enabled the genetic remodeling of biological systems toward customized microbial cell factories for industrial uses (**Figure C-3**).^[6, 29-30] Bioinformatics tools complement experimental advances, have facilitated the understanding of protein sequence-function relationships, and become an integral part of modern protein engineering and pathway design.^[6, 31-33]

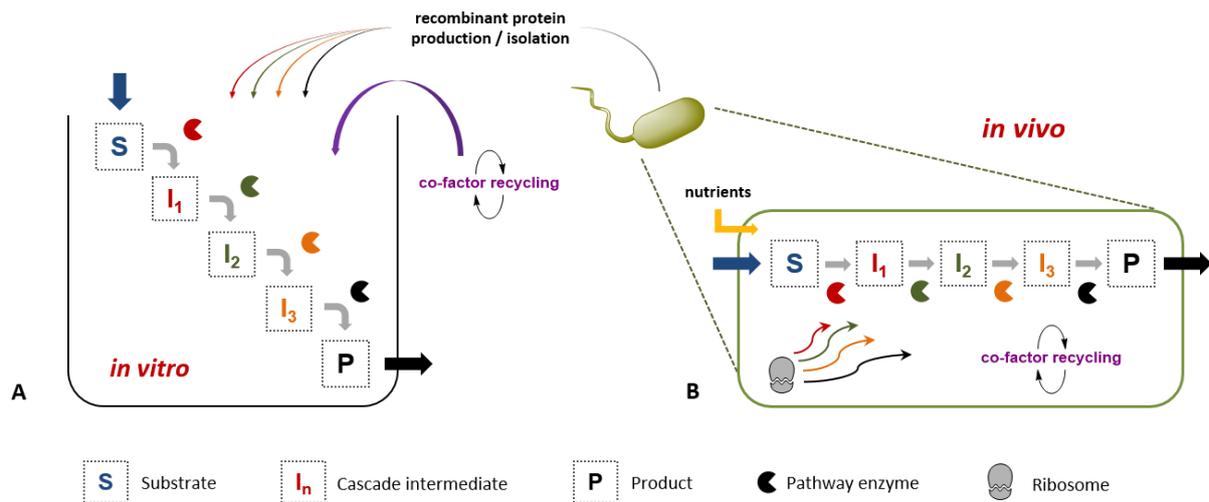


Figure C-2. Features of *in vitro* and *in vivo* multi-enzyme cascades. (A) *In vitro* cascades can be easily optimized by changing process parameters (e.g., substrate and enzyme concentrations), but their assembly requires the preparation of enzymes and cofactors must be artificially added or regenerated. (B) *In vivo* cascades benefit from the simultaneous production of all pathway enzymes and the recycling of cofactors by the host cell metabolism.^[6]

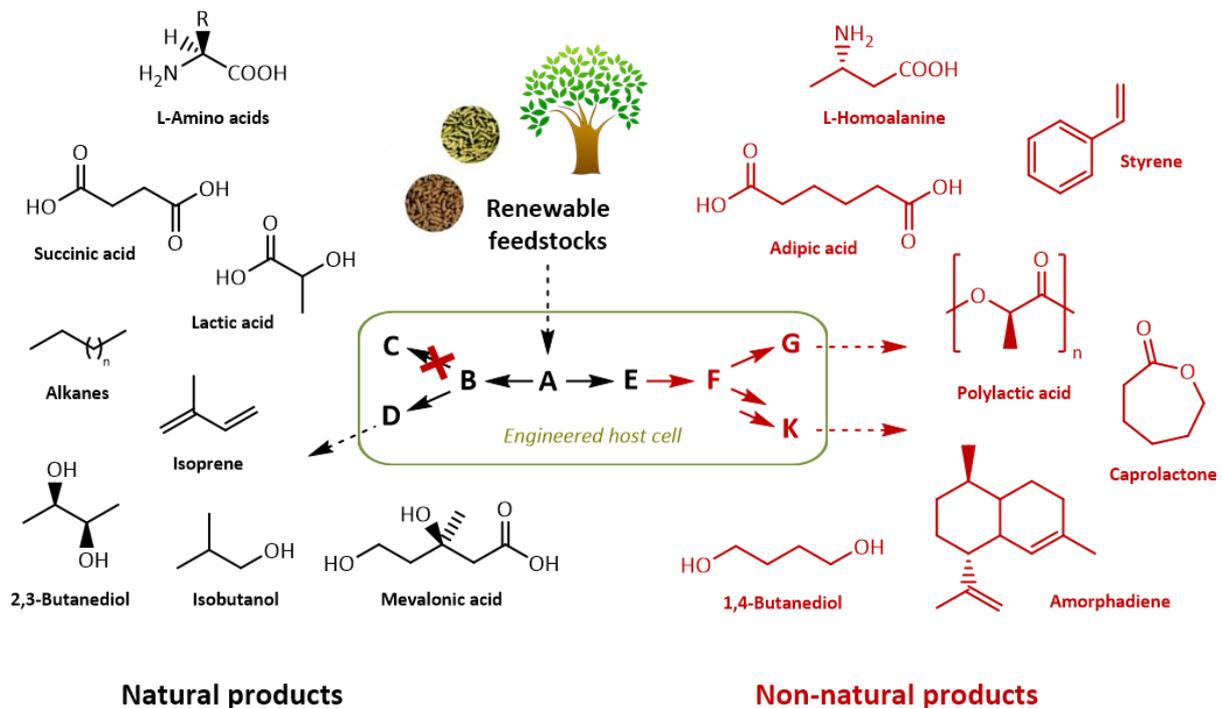


Figure C-3. Development of engineered metabolic routes. Metabolic engineering redirects native metabolic pathways (black arrows), for example, by the knock-out (KO) of endogenous enzyme activities (red X) to accumulate natural products (in black on the left). The engineering of microbial cells and the introduction of non-native pathways (red arrows) produced non-natural compounds (in red on the right) from renewable feedstocks.^[34]

In context of the increasing environmental impact of the chemical and pharmaceutical industry, the stewardship of available resources and the reduction of waste are key challenges of the present and the future. Biocatalysts are made from renewable sources, they are biodegradable and nontoxic. Since they are perfectly combinable in

cascade type reactions (**Figure C-2**), characterized by high selectivities that simplify downstream purification processes and result in higher product yields, biocatalysis provides an attractive alternative to classic organic syntheses, which is in full compliance with the 12 principles of green chemistry.^[35] With the recent advent of systems metabolic engineering, combining metabolic and genetic engineering with systems biology and synthetic biology, scientists have started to efficiently remove bottlenecks and to maximize product titers of artificial enzyme cascades *in vivo*.^[6, 36-37] Whereas the emerging CRISPR/*Cas9* (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9) genome editing tool for eukaryotic cells^[25, 38-39] is a yet underrepresented gene manipulation method in bacteria,^[40-41] scientists routinely reorganize the metabolic background of microbial host cells by gene knock-out (KO) strategies.^[42] Current research successfully assembled *de novo* pathway modules for manufacturing a broad variety of nonnatural compounds^[43-44] (**Figure C-4**) and projects toward the construction of highly engineered (*de novo*) biocatalysts – enzymes and whole cells – that offer potential to tackle today's environmental and industrial challenges, including electrogenic reactors based on 'microbial fuel cells' for harvesting electricity from biomass,^[45] the degradation of plastics to reduce waste,^[27] or the fixation of CO₂ as an important carbon feedstock for a green economy in the future (**Figure C-1**).^[26]

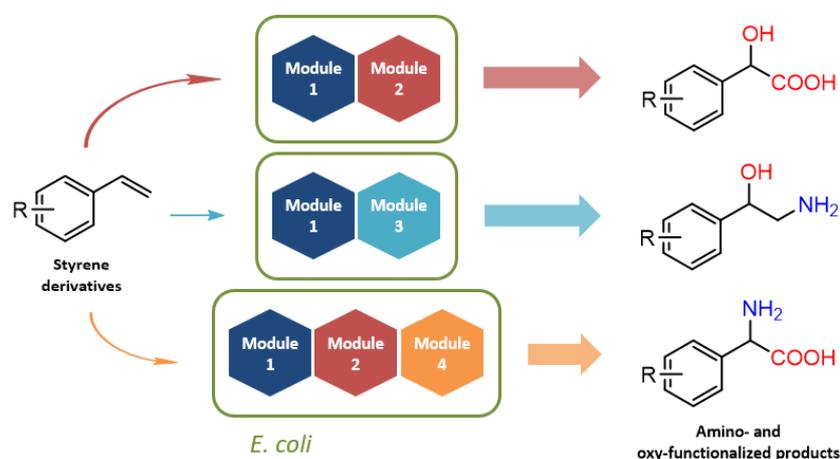


Figure C-4. Example of a modular pathway design for the functionalization of terminal alkenes. Production of α-hydroxy carboxylic acids (top), 1,2-amino alcohols (center), and α-amino acids (bottom) by *E. coli* cells harboring multiple enzyme modules. M1: epoxidase, epoxide hydrolase; M2: alcohol dehydrogenase (ADH), aldehyde dehydrogenase (AldH); M3: ADH, ω-transaminase (ω-TA), alanine dehydrogenase (AlaDH); M4: hydroxy acid oxidase, α-transaminase, catalase, glutamate dehydrogenase. The figure was adapted from S. Wu *et al.* (2016).^[43]

C II Tools for cellular pathway engineering

A suitable host proves essential for *in vivo* pathway design, introduction, and optimization. The gram-negative bacterium *E. coli* is widely used in pathway engineering because of its rapid growth at high density on inexpensive nutrients and its well-understood genetics.^[6, 46-47] The access to comprehensive genome sequences from databases enabled the application of rational gene KOs and the construction of mutant libraries (e.g., the Keio collection).^[48] In living cells, many different proteins and enzymes are simultaneously produced to ensure vital functions such as reproduction and growth. The introduction of metabolically nonrelated enzymes potentially disturbs the finely tuned metabolic host network, which, in turn, might primarily interfere with the production of pathway enzymes *in vivo*.^[6, 10, 47] To maximize the carbon flux through synthetic pathways and to meet performance metrics such as high product titers, both *in vitro* and *in vivo* applications require a balanced pathway enzyme stoichiometry. Therefore, target enzymes must be individually produced or coexpressed in *E. coli* in

sufficient amounts balanced for individual activities. For decades, enzymes have been overexpressed from plasmids, which can be constructed by standard molecular cloning techniques involving restriction enzymes and ligases.^[49] For the (heterologous) production of enzymes and other pathway proteins, the use of plasmids is still the method of choice since an increasing number of different vectors and standardized guidelines for construction and subsequent protein production are available.^[50] However, the expression of whole pathways is challenging, not only because of the innately high metabolic burden of artificial enzyme cascades bestowed upon the host.^[6]

C II.1 The challenges of multiple recombinant protein expression

Before 1970, the characterization of individual genes from microorganisms was impaired due to the inability to isolate them. This changed dramatically with the advent of molecular cloning methods after the identification and isolation of restriction endonucleases in bacteria, restricting the growth of bacteriophages (**Figure C-1**).^[18-19, 49] Since then, recombinant DNA molecules could be propagated in *E. coli*, target genes expressed, and their functions studied.

Today, synthetic pathways comprise of several pathway elements – usually enzymes – but the expression of a larger number of different enzymes from individual plasmids is not feasible. Not only does this approach involve separate rounds of molecular cloning; major obstacles are encountered after the introduction of plasmids. The replication of foreign DNA^[51] and the overproduction of recombinant proteins^[52] impose a high metabolic load on the host.^[6] The drain of resources from the host cell induces stress responses including amino acid depletion, starvation, and heat shock responses.^[50] Housekeeping genes involved in transcription, translation, and amino acid biosynthesis are down-regulated, reducing growth rates and cell viability.^[52-56] As a result, the flux through *de novo* pathways can be strongly impaired with low productivities as only one consequence.^[6, 57-59]

An immediate solution to reduce the metabolic burden^[60] is the coexpression of two (or more) genes from a single plasmid. One common tool for protein coproduction is the Duet expression system from Novagen. It offers plasmids with compatible origins of replication (ORIs) and different antibiotic markers including ampicillin (Amp; pETDuet-1), chloramphenicol (Cam; pACYCDuet-1), kanamycin (Kan; pRSFDuet-1 and pCOLADuet-1), and streptomycin (Str; pCDFDuet-1) (**Table C-1**).^[6] Single Duet vectors are useful to coexpress enzyme complexes^[61] or metabolically related enzymes, for example, carboxylic acid reductases (CARs) and a phosphopantetheinyl transferase (PPTase). The PPTase posttranslationally attaches a phosphopantetheinyl residue to the apo-CAR to form the functional holo-CAR enzyme (**Figure C-36**).^[62-64]

By employing four Duet vectors with compatible replicons and complementing antibiotic resistances, up to eight proteins can be produced in a single cell.^[43, 65] Hence, the Duet expression system has been successfully applied to enhance the production of various natural compounds in *E. coli* and to introduce synthetic pathway modules to access nonnatural products.^[6] Scientist in the Prather group assembled pathway enzymes in modules that allowed to rationally test different combinations to increase product titers of odd-chain fuels and chemicals including pentanol^[66] and 4-methyl-1-pentanol.^[67] The latter was formed *via* a ten-step *de novo* pathway that was designed by a retro-biosynthetic approach^[68] with enzymes taken from nine different microorganisms (**Figure C-5**). Wu *et al.* also utilized a modular approach to convert nonnatural (substituted) styrenes into amino- and oxyfunctionalized products in *E. coli* cells, heterologously coexpressing up to eight different enzymes from Duet vectors (**Figure C-4**).^[43]

Table C-1. Plasmid replicons, copy numbers, and compatibility

Plasmid(s)	ORI	Copy number ^[a]	Compatible ORIs
pUC (all), pGEM (all)	pMB1 ^[b]	300–500	RSF1030, Mini-F/RK2, CloDF13, ColA, P15A
pRSF, pRSFDuet-1	RSF1030	>100	ColE1, CloDF13, P15A, pSC101
pET (all), pETDuet-1	ColE1 (pBR322)	~ 40	RSF1030, Mini-F/RK2, CloDF13, ColA, P15A
pETcoco™	Mini-F/RK2	~ 40 ^[c]	ColE1, ColA, P15A
pCDF, pCDFDuet-1	CloDF13	20–40	RSF1030, ColE1, ColA, P15A, pSC101
pCOLA, pCOLADuet-1	ColA	20–40	ColE1, Mini-F/RK2, CloDF13, P15A, pSC101
pACYC, pACYCDuet-1, pRARE	P15A	10–12	RSF1030, ColE1, Mini- F/RK2, CloDF13, ColA
pSC101	pSC101	~ 5	RSF1030, ColE1, CloDF13, ColA

^[a] Copy numbers may vary depending on the size/type of DNA insert, growth conditions, and the *E. coli* strain used. Copy numbers were adapted from T. Bayer *et al.*^[6] and references therein. ^[b] pMB1 derivatives are related to the ColE1 replicon, which originates from the pBR322 ORI. ^[c] Amplifiable up to 40 copies per cell.

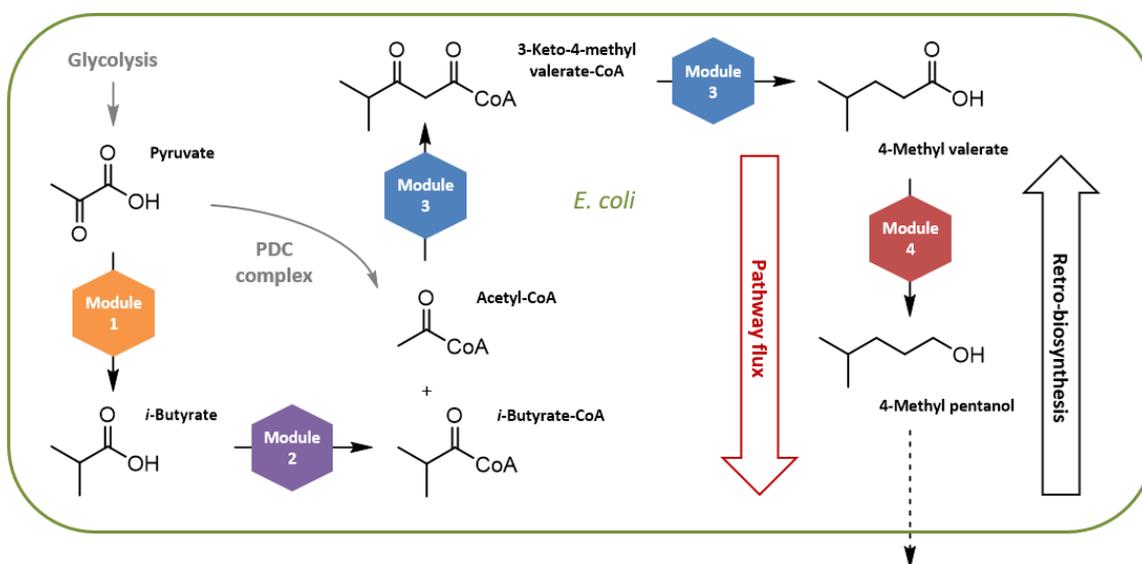


Figure C-5. Modular 4-methyl pentanol pathway designed by retro-biosynthesis. Modules M1 and M2 utilize the glycolytic product pyruvate to form the precursor *i*-butyrate-CoA. M3 is the coenzyme A (CoA)-dependent module to elongate the carbon chain. The condensation of *i*-butyrate-CoA and acetyl-CoA, which is formed by the pyruvate decarboxylase (PDC) complex, yields 3-keto-4-methyl valerate-CoA. After three consecutive enzymatic transformations, 4-methyl valerate is formed by endogenous thioesterase activity.^[69] M4 is the alcohol-terminating module yielding 4-methyl pentanol. Retro-biosynthetic analysis and *de novo* pathway construction started from the target compound (in opposite direction of the final carbon flux). The formation of *n*-butanol from acetyl-CoA as a pathway byproduct was omitted for clarity. M1: acetolactate synthase (*B. subtilis*), acetohydroxy acid isomeroreductase, dihydroxy acid dehydrolase (*E. coli*), decarboxylase (*Lactobacillus lactis*), ALDH (*Flavobacterium johnsoniae*); M2: isobutyryl-CoA ligase (*Rhodospseudomonas palustris*); M3: thiolase, acetoacetyl-CoA reductase, enoyl-CoA hydratase (*Cupriavidus necator*), enoyl-CoA reductase (*Treponema denticola*); M4: CAR (*Nocardia iowensis*), ADH (*Leifsonia* sp. strain S749). The figure was adapted from M. J. Sheppard *et al.* (2014).^[67]

Despite the inspiring synthetic pathway applications based on combinations of different Duet vectors and the assembly of enzyme modules (i.e., enzymatic toolboxes), there are limitations. Considering the metabolic burden, the Duet expression system exclusively features medium- and high-copy number plasmids (**Table C-1**).^[6] The plasmid maintenance by host cells can provoke stress responses that are related to the copy number.^[70] Contrary, low-copy number plasmids confer a lower metabolic burden. They are more stably maintained and capable of replicating larger pieces of DNA due to their larger native size and the replication mechanism. This makes low-copy number plasmids excellent but undervalued alternatives to medium- or high-copy number plasmids for the vector-based introduction of metabolic pathways.^[6, 71-72]

Furthermore, the Duet expression system does not allow quick modular alterations of single or multiple pathway elements. Changes in the pathway architecture from the pseudo-operon configuration in the Duet vectors to operons, for example, are not feasible (**Figure C-6**). With only strong isopropyl β -D-1-thiogalactopyranoside (IPTG)-inducible T7 promoters (P_{T7}) in place, which is also known to be a leaky expression system, tuning of gene expression is not possible.^[6] Therefore, enzyme stoichiometry is hard to control, which is strictly required for both a balanced flux through synthetic pathways and the reduction of the metabolic burden.

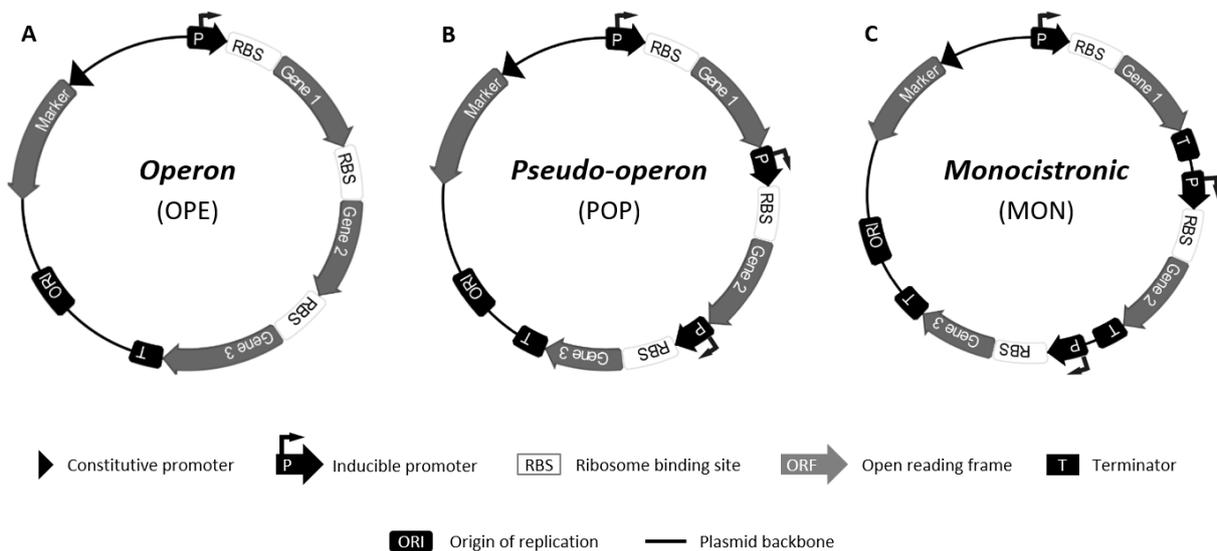


Figure C-6. Genetic configurations of an example plasmid harboring three target genes. (A) In operons, a gene cluster is under the control of an inducible promoter and one terminator. (B) The expression of genes is controlled by individual promoters in pseudo-operons. One common terminator stops transcription. (C) In monocistronic configurations, multiple genes are under the control of an individual promoter and a terminator. The ORI is necessary for plasmid replication and a marker gene (e.g., antibiotic resistance gene) for positive selection and plasmid maintenance under selective pressure. The Duet vectors are designed to harbor two target genes in pseudo-operon configuration. The figure was adapted from T. Bayer *et al.* (2015).^[6]

As noted before, current and future artificial metabolic routes will not only depend on modularity but tools for fine-tuning of expression on different molecular levels (e.g., transcription and translation) to balance the simultaneous production of all pathway enzymes.^[32, 71, 73-74]

Xu *et al.* combined both modularity and the regulation of gene expression in a set of ePathBrick vectors.^[75] The ePathBrick expression system supports the assembly of multi-component pathways in different configurations (**Figure C-6**) and the incorporation of genetic activator and repressor elements enabling transcriptional fine-tuning.^[6, 75] The applicability was demonstrated by the functional assembly of a seven-gene flavonoid pathway (9 kb) on a single ePathBrick vector in monocistronic configuration (**Figure C-6C**), dramatically reducing the plasmid burden.^[6, 75] The construction of ePathBrick vectors was based on restriction enzymes, precisely, the four

isocaudamers *XbaI*, *SpeI*, *AvrII*, and *NheI*.^[76] These restriction enzymes are compatible with the BioBrick™ standard, which was introduced by Knight in 2003 and first contemplated by Rebatchouk, who suggested a general cloning strategy referred to as NOMAD (nucleic acid ordered assembly with directionality) in 1996.^[77] The BioBrick™ standard aims at the dissection of biology into standardized parts (e.g., promoter, terminator, RBS) and their directional assembly into genetic devices and whole (synthetic) pathways by consecutive rounds of molecular cloning, utilizing the isocaudomer pair *XbaI* and *SpeI* (**Figure C-7**).^[76] Isocaudomers are pairs of restriction enzymes that have slightly different recognition sequences but, upon cleavage, generate identical cohesive termini. Ligation results in a scar sequence that cannot be cleaved by either of the initial restriction enzymes.^[6] While the BioBrick™ scar sequence encodes an in-frame stop codon, the isocaudomer scar sequence of *BglII* and *BamHI* encodes a Gly-Ser, which can be utilized to produce fusion proteins.^[78-79]

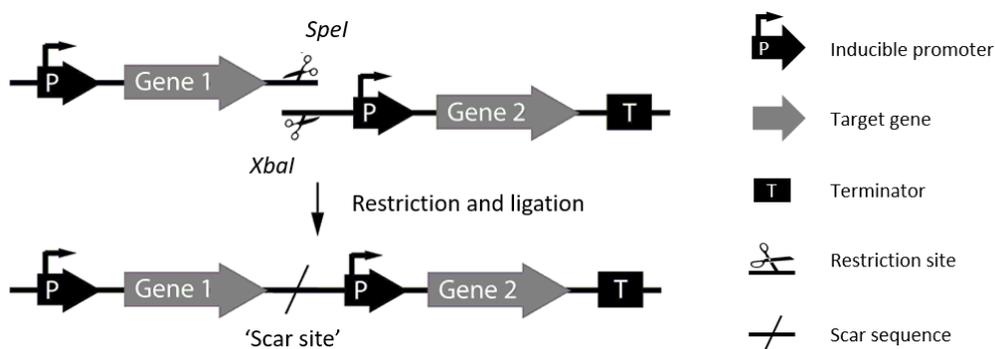


Figure C-7. Schematic BioBrick™ assembly. Restriction by the isocaudomer pair *SpeI* and *XbaI* creates compatible sticky ends that can be joined by ligation. The resulting scar sequence cannot be cut by the original enzymes. Consequently, *SpeI* and *XbaI* can be re-used in a subsequent round of cloning. The figure was adapted from T. Bayer *et al.* (2015).^[6]

Although the simplification of biology into parts and their Lego brick-like assembly are powerful tools for the construction of artificial pathways, the iterative rounds of restriction enzyme digestion and ligation can be time-consuming. Furthermore, the sequence dependency of restriction enzymes can be troublesome if a target gene contains recognition sites for the desired restriction enzymes. Certainly, other restriction enzymes can be used and codon degeneracy allows alterations in a nucleotide sequence to ditch forbidden restriction sites without changing the amino acid sequence. Since such strategies offer not more than a laborious compromise, cloning methods omitting the repeated use of restriction enzymes and ligase were developed to assemble whole metabolic pathways (ideally) in one step.^[6]

C II.2 Advanced cloning techniques

In the last two decades, advances in DNA technologies have dramatically reduced the cost of commercial DNA synthesis and sequencing.^[17, 80-81] The limiting technology for the construction of larger metabolic pathways with a defined architecture featuring the desired regulatory elements has long been its assembly. This obstacle was overcome and several DNA assembly methods allow constructing synthetic pathways and can even assemble whole genomes.^[6, 32, 36, 82-84]

Advanced cloning techniques utilize polymerase chain reactions (PCRs) to produce linear DNA fragments to be joined. Joining is facilitated by terminal, overlapping, and homologous DNA sequences. Hence, multiple DNA fragments can be linked in a seamless and (almost) sequence-independent fashion. This circumvents restriction enzyme digestion, ligation and, consequently, the scar sequences generated during the BioBrick™ and related assembly methods, for example (**Figure C-7**).^[6]

PCR cloning became interesting for pathway construction with the introduction of sequence- and ligation-independent cloning (SLIC), which enabled the assembly of ready-to-transform plasmids harboring target pathway genes.^[6, 85] Originally applied by Li *et al.*,^[86] SLIC has been successfully applied and constantly improved.^[32]

One advanced SLIC protocol repeatedly used in this thesis is Fast Cloning (FC).^[87] The target vector backbone and the desired insert are amplified in independent PCRs. A mixture of both linear DNA fragments is digested with *DpnI* to destroy the methylated PCR templates. Subsequently, competent *E. coli* cells are directly transformed with the mixture. The DNA fragments are joined *in vivo* via the homologous ends (15–30 bases) introduced during PCR. Finally, the bacterial DNA repair machinery seals DNA strand breaks (**Figure C-8**).^[87]

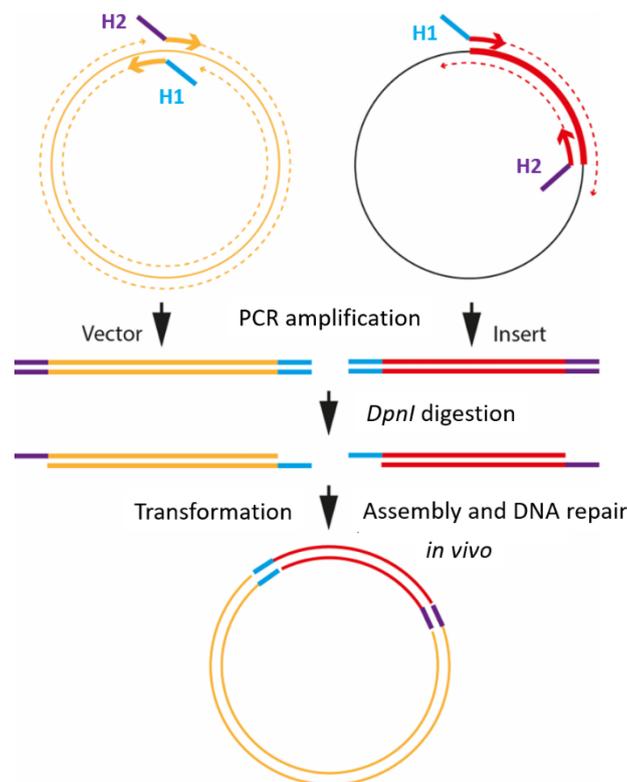


Figure C-8. Vector construction by Fast Cloning (FC). FC is an improved SLIC protocol that involves the separate amplification of vector and insert by PCR with terminal homologous overhangs (H1 and H2). A mixture containing both linear DNA fragments can be transformed after *DpnI* digestion. The overlapping sequences H1 and H2 direct assembly and DNA gaps sealed *in vivo*. The exact mechanism is not fully understood yet.^[87] The figure was adapted from T. Bayer *et al.* (2015).^[6]

Although FC is rather simple and mostly sequence- and ligation-independent, scaling up to assemble entire pathways can be difficult. (Engineered) DNA polymerases exhibiting proof-reading activity, for example, reduced the innate error rate of PCRs and are capable of amplifying nucleotide sequences with an increased GC content. However, plasmids become less efficient at larger sizes and, importantly, complex mixtures only assemble with low efficiency *in vivo* after transformation.^[6, 32]

Higher developed microorganisms such as *S. cerevisiae* outpace the DNA repair mechanisms of bacterial cells. Defects like double strand breaks (DSBs) can be efficiently sealed by processes involving homologous recombination events.^[6, 88] Gibson *et al.* exploited this repair capacity by adapting a long-known cloning protocol in *S. cerevisiae* called transformation-associated recombination (TAR; **Figure C-9**)^[89] to assemble the entire

circular genome of *Mycoplasma genitalium* in a single step in yeast cells.^[84, 90] The mechanism is also based on overlapping sequences that undergo homologous recombination.^[6] TAR was successfully applied to rebuild natural pathways for two polyketides, aureothin^[91] and spectinabilin^[92], which were then heterologously expressed in *E. coli* (Figure C-9).^[6, 93-94]

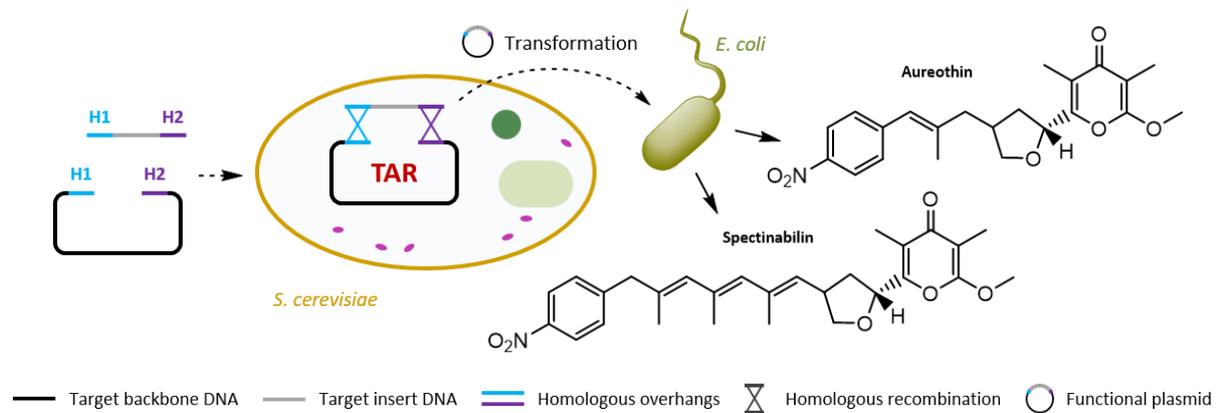


Figure C-9. Transformation-associated recombination (TAR) in yeast. Homologous recombination is facilitated by matching DNA overlaps (H1 and H2) *in vivo*. TAR is capable of assembling more complex mixtures of DNA fragments and was used to re-build natural pathways yielding aureothin and spectinabilin after introduction in *E. coli*.^[93]

Due to the high fidelity and accuracy of enzymes involved in homologous recombination and the fact that *S. cerevisiae* can tolerate large (synthetic) DNA molecules, this and similar yeast-based assembly strategies hold promise to be reliable and powerful tools for *de novo* pathway construction.^[6, 32, 93-95]

A very convenient *in vitro* technique for the ligation of multiple DNA fragments is the ‘Gibson’ isothermal assembly, commonly referred to as ‘Gibson assembly’ (Figure C-10). It omits both potentially forbidden restriction sites of restriction enzyme based cloning strategies and the increasing error-rate of PCRs when amplifying long DNA sequences.^[6, 83] Gibson assembly employs an enzyme cocktail containing a high-fidelity DNA polymerase, a T5 exonuclease, and a *Taq* DNA ligase and was successfully applied to assemble small circular genomes of 16–580 kb in one step.^[83, 96] Short incubation times at a single temperature and the parallel assembly of multiple DNA pieces containing homologous overlaps are obvious advantages. Additionally, the elevated temperature (50°C) can be beneficial to resolve rigid secondary DNA structures (e.g., stem loops) that would otherwise interfere with an efficient assembly.^[6, 32, 83] Since the Gibson assembly depends on three purified enzymes, it is more expensive than *in vivo* SLIC methods. However, the ligase activity may enhance assembly efficiency, which is relatively low for assemblies *in vivo*.^[6]

Zhang and co-workers applied a seamless ligation cloning extract (SLiCE) to assemble linear DNA fragments *in vitro*.^[97] Since different laboratory strains can be used as sources for SLiCE (e.g., *E. coli* TOP10), it is a very cheap alternative to the enzyme cocktail used in Gibson assembly. However, the bacterial extract contains the same enzymes that join DNA fragments *via* homologous overhangs only at low efficiency *in vivo*. Therefore, Zhang *et al.* employed an *E. coli* DH10B strain expressing a λ *Red* recombination system (PPY strain), which greatly enhanced assembly efficiency.^[97]

Assembly tools such as SLIC, Gibson assembly, and the application of bacterial extracts depend on homologous, single stranded DNA (ssDNA) overlaps of certain lengths depending on the method used. If ssDNA adopts stable secondary structures such as hairpins and stem loops (e.g., terminator sequences),^[98-99] they can strongly interfere with directed assembly. Furthermore, identical homologous sequences must not be used repeatedly since this can produce constructs either not containing all desired DNA fragments or in unwanted configurations.

Consequently, the thoughtful design of DNA overlaps is essential, especially for the assembly of complex mixtures containing many different biological parts. This renders such methods not truly sequence-independent and can be a major obstacle.^[6]

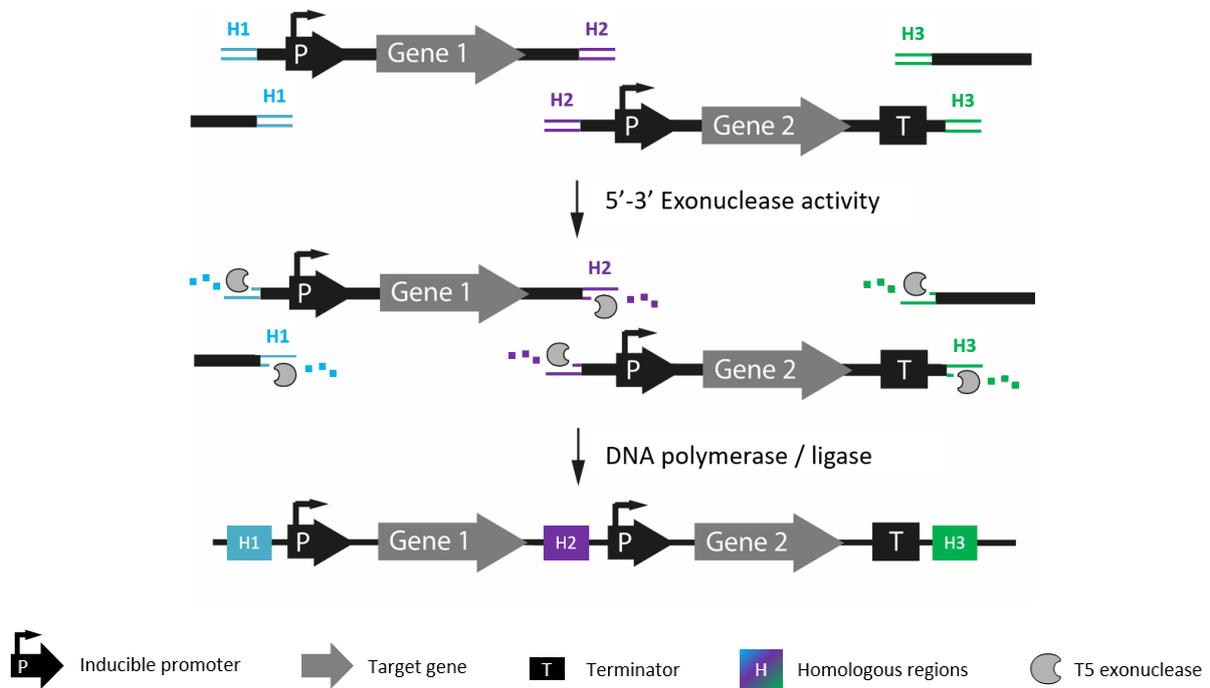


Figure C-10. Schematic Gibson assembly. Gibson assembly utilizes a 5'-3' exonuclease, a DNA polymerase, and a ligase to ligate mixtures of linear DNA fragments *in vitro*. The T5 exonuclease chews back the ends of homologous DNA sequences (H1–H3) creating terminal overlaps. The DNA polymerase fills the gaps and a *Taq* ligase finally seals them (both not shown for clarity). The figure was adapted from T. Bayer *et al.* (2015).^[6]

With advanced DNA assembly techniques available, single biological parts can now be more efficiently assembled into multi-component pathways. Further improvements of established protocols included the exploitation of novel enzymatic activities to enhance assembly efficiencies.^[100-101] Computational tools will highly support DNA assembly strategies in the future and lead to automated DNA assembly processes.^[6, 32, 37, 102]

Since the imposed metabolic burden on the heterologous host due to the drain of resources used to produce enzymes and other pathway components is inevitable, the plasmid burden can be reduced by assembling metabolic pathways on a minimal number of plasmids. Ultimately, the plasmid burden can be abolished by the integration of *de novo* pathways into the host genome.^[6]

C II.3 Genomic modification tools

Although plasmids are easy to insert into a cell and allow strong gene expression, they do not only contribute to the metabolic burden. Plasmids can suffer from genetic instability consequently, decreasing product yield and productivity of the *de novo* pathway due to both the loss and the alteration of plasmids. Major contributors to genetic instability are: segregational instability, in which unequal distribution of plasmids to daughter cells leads to plasmid-free cells; structural instability, in which plasmids acquire altered DNA sequences that may cause incorrect expression of the target pathway enzymes; and allele segregation, in which functional plasmids are displaced by non-functional plasmids, resulting in non-productive cells that are resistant to selection pressure.^[103]

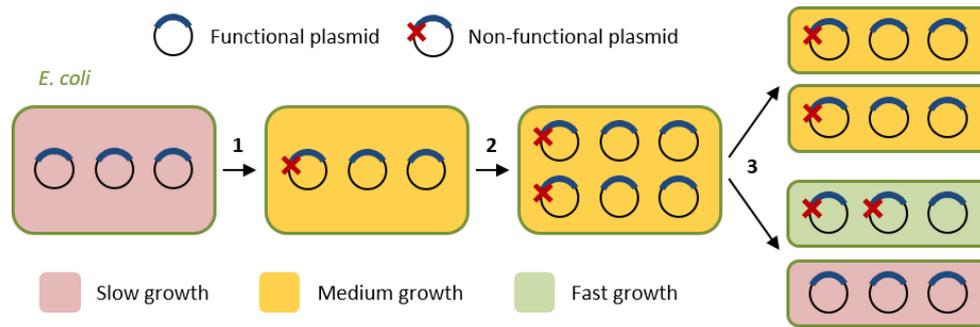


Figure C-11. Allele segregation mechanism. (1) Random mutation (red X) eliminates functional pathway expression but does not affect the selectable marker. (2) Plasmids are copied before cell division. Ordered replication results in one plasmid copy (as shown), whereas random replication leads to varying copy numbers (not shown). (3) Plasmids are segregated randomly during cell division, which leads to different plasmid populations in daughter cells. Daughter cells exclusively harboring functional plasmids (pink) show slow growth and high productivity. Cells predominantly containing functional plasmids (yellow) exhibit medium growth and moderate productivity, while cells accumulating non-functional plasmids (green) grow faster at low productivity. The high growth rate will result in the accumulation of the mutant (non-functional) plasmid. The figure was adapted from K. Tyo *et al.* (2009).^[103]

Under selective pressure (e.g., antibiotic resistance), plasmid-free cells (segregational instability) are not viable. Computational modeling suggests that allele segregation, not random mutations, majorly contributes to the rapid plasmid productivity loss in both ordered and random plasmid inheritance models (**Figure C-11**).^[103] Upon target pathway expression, cells containing non-functional plasmids can grow and replicate faster due to the decreased metabolic burden, accumulate, and lead to an unproductive cell population.

Various strategies have been implemented to reduce segregational and structural instability.^[71] However, allele segregation, which is not soothed by selection markers and the non-viability of plasmid-free cells, has been widely unaddressed in synthetic pathway design and metabolic engineering. Genomic integration ensures the ordered inheritance of recombinant pathways. Random modifications of bacterial genomes by chemical mutation,^[104] UV irradiation,^[105] or transposon mutagenesis^[106] have been used for decades. Recently, the group of Nickel employed transposable elements to insert the entire poly-3-hydroxy butyrate (P3HB) biosynthetic pathway from *Cupriavidus necator* (*C. necator*) into the *E. coli* genome with good integration efficiency.^[107] However, the introduction of undirected (unwanted) alterations in the genome represents a major drawback of these methods. The availability of complete genome sequences facilitated targeted, homologous recombination-based DNA modifications. Such methods often utilize the enzymatic activity of *RecA*^[103, 108] or certain phage-derived enzymes that enable homologous recombination.^[6, 36-37]

Tyo *et al.* circumvented allele segregation by anchoring pathway operons for lycopene and P3HB (**Figure C-12A**) production, respectively, into the *E. coli* chromosome.^[103] Genomic integration was achieved by site-specific recombination of the λ phage (*attP*) and the *E. coli* (*attB*) attachment sites on the donor plasmid and the host genome, respectively (**Figure C-12B**).^[103, 109] The iterative duplication of pathway-coding operons was performed by *recA*-mediated homologous recombination in a process termed chemically inducible chromosomal evolution (CICHe). Finally, the number of genomically inserted pathway copies was stabilized by *recA* deletion (**Figure C-12C**) and the evolved CICHe strain was unaffected by plasmid instabilities in the absence of the antibiotic.^[103]

Although the productivity loss due to allele segregation in plasmid-based pathway engineering was suggested, it has been hardly addressed in industrial biotechnology.^[110] Furthermore, many pathway engineering efforts might have suffered from allele segregation by decreasing product yields and the productivity of the desired compounds in batch and continuous fermentations and, thus, will remain a challenge in future applications.

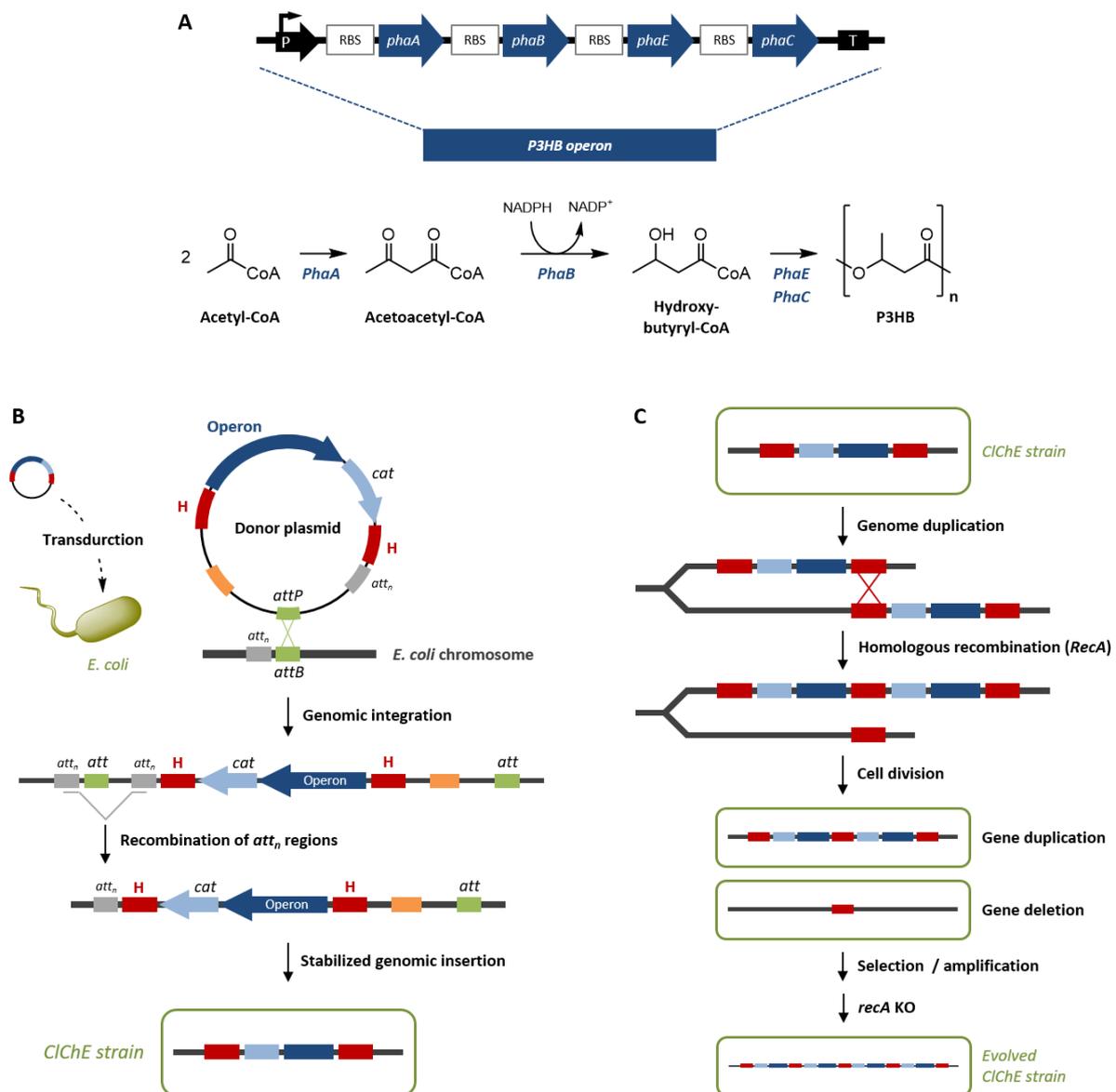


Figure C-12. Chemically inducible chromosomal evolution (CIChE). (A) Operon coding for the P3HB pathway. P3HB is produced from intracellular acetyl-CoA involving four enzymes (*PhaA*, *PhaB*, *PhaE*, and *PhaC*). (B) Stable integration of a single copy of the P3HB operon into the *E. coli* chromosome. The donor plasmid resulted from the recombination of two DNA molecules (not shown): a ColE1-derived plasmid containing the P3HB operon, the chloramphenicol acetyltransferase (*cat*) gene and flanking homologous regions (H) and the λ phage genome. Recombination was facilitated between a portion of the *cat* gene and region homologous to the ColE1 ORI (orange rectangle). *E. coli* is transduced with the recombinant λ phage (top left). The lysogene is formed by site-specific recombination between the *attP* and the *attB* sites (top right). Recombination in two *attn* sites removes remaining λ phage DNA (center),^[109] which results in a stable genomic integration in the CIChE strain (bottom). (C) The construct delivered to the CIChE strain contains the P3HB operon (blue rectangle), the *cat* gene conferring chloramphenicol (Cam) resistance (light blue rectangle), and two homologous regions (red rectangles; top). *RecA* mediates an uneven homologous crossover between the homologous flanking regions, which generates a DNA strand with two insert copies and one strand with a deletion (top center). After cell division, one daughter cell inherits two copies, whereas the other lost the insert (bottom center). Cam is used to select for cells with increased numbers of insertions. Finally, the *recA* gene is knocked-out to prevent further changes in the copy number (bottom).^[103] (A) and (C) were adapted from K. Tyo *et al.* (2009).^[103] (B) was adapted from D. Boyd *et al.* (2000).^[109]

Currently, one routinely employed genomic integration tool for metabolic pathway engineering is the λ *Red* system.^[6, 37, 111] The λ *Red* recombination protocol had been established first by Datsenko and Wanner, involving

a *Flp* recombinase besides λ *Red*,^[112] and was successfully applied to construct the Keio collection of non-essential gene KOs in *E. coli*.^[48] This strategy involved two “helper” plasmids encoding the λ *Red* and the *Flp* recombinase, respectively (**Figure C-13A**). The activity of λ *Red* disrupted the target gene by insertion of an antibiotic resistance marker *via* homologous recombination. In addition to sequences homologous to the gene to be knocked-out flanking the marker gene, also FRT (*Flp* recognition target) sequences are present. Colonies with the disrupted target gene can be selected by their acquired antibiotic resistance. After transformation with the second “helper” plasmid, the antibiotic resistance gene is excised by *Flp* recognizing the FRT sequences (**Figure C-13A**). Both helper plasmids have temperature sensitive replicons and are easily cured.^[6, 112]

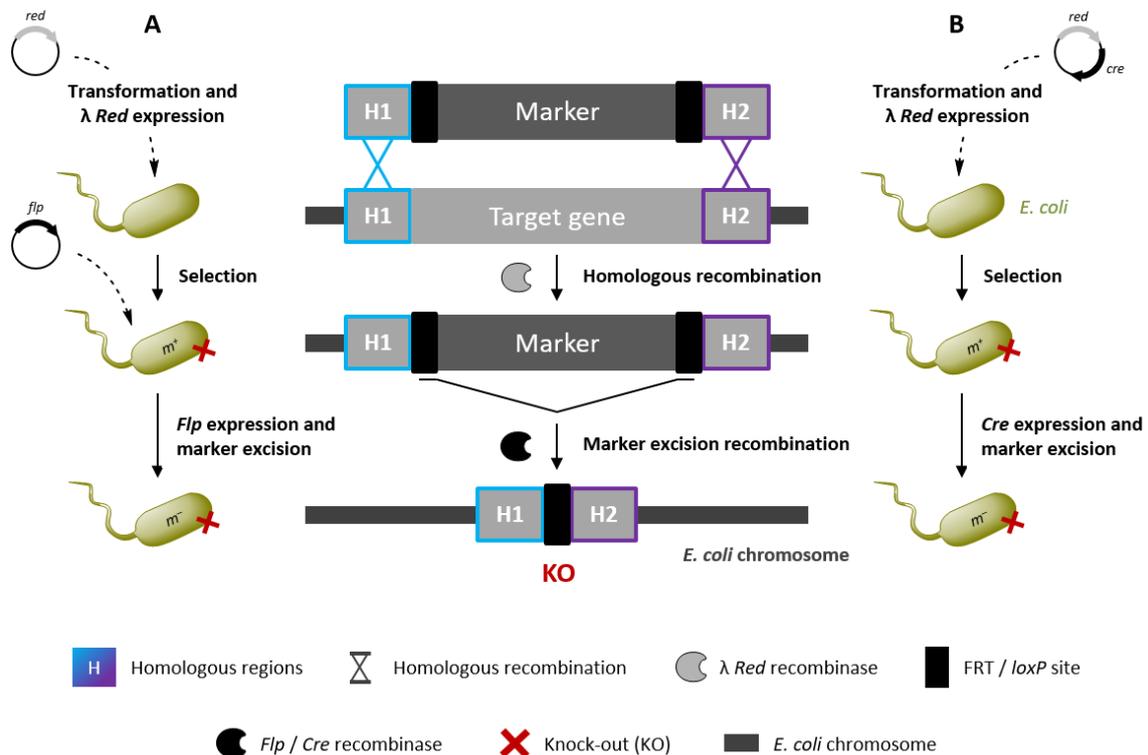


Figure C-13. λ *Red* recombination. (A) Target gene KO in *E. coli* involves the expression of λ *Red* from a first “helper” plasmid post transformation. The insertion of an antibiotic resistance gene *via* λ *Red* mediated recombination between homologous sequences (H1 and H2) disrupts the target gene and enables positive selection. The transformation of a second “helper” plasmid and the expression of *Flp* leads to the excision of the marker gene *via* flanking FRT sites. Plasmids contain temperature sensitive ORIs and can be cured by a temperature shift. Sequential recombinase expression can be induced by the same inducer.^[112] (B) Transformation of only one “helper” plasmid harboring two individually inducible recombinases, λ *Red* and *Cre*, respectively. Upon expression of λ *Red*, marker insertion knocks-out the target gene. After positive selection, *Cre* is expressed with a second inducer, leading to the excision of the marker *via* flanking *loxP* sites. The use of one plasmid omits repeated rounds of “helper” plasmid transformation and curing.^[36]

A modified protocol also involved the λ *Red* system for gene disruption but a *Cre* recombinase for subsequent marker excision. The individually inducible recombinases were encoded on a single “helper” plasmid, conveniently omitting repeated rounds of transformation and curing (**Figure C-13B**).^[6, 113] The system was successfully employed to sequentially delete four endogenous genes, which resulted in a fumaric acid overproducing *E. coli* strain.^[113]

To introduce larger DNA fragments on the *E. coli* chromosome, the λ *Red* system was combined with the yeast mitochondrial homing endonuclease *I-SceI*.^[114-116] The recognition site of *I-SceI* is absent from the *E. coli* genome and rather large (18 bp) for an endonuclease. *I-SceI* introduces a DSB at the target locus, which stimulates *in vivo*

recombination, consequently, increasing recombination efficiency (**Figure C-14**).^[117-118] Notably, to utilize *I-SceI* activity to anchor DNA encoding large synthetic constructs on the *E. coli* chromosome, *I-SceI* restriction sites must be integrated first at the target locus, which can be laborious.^[6, 114-115]

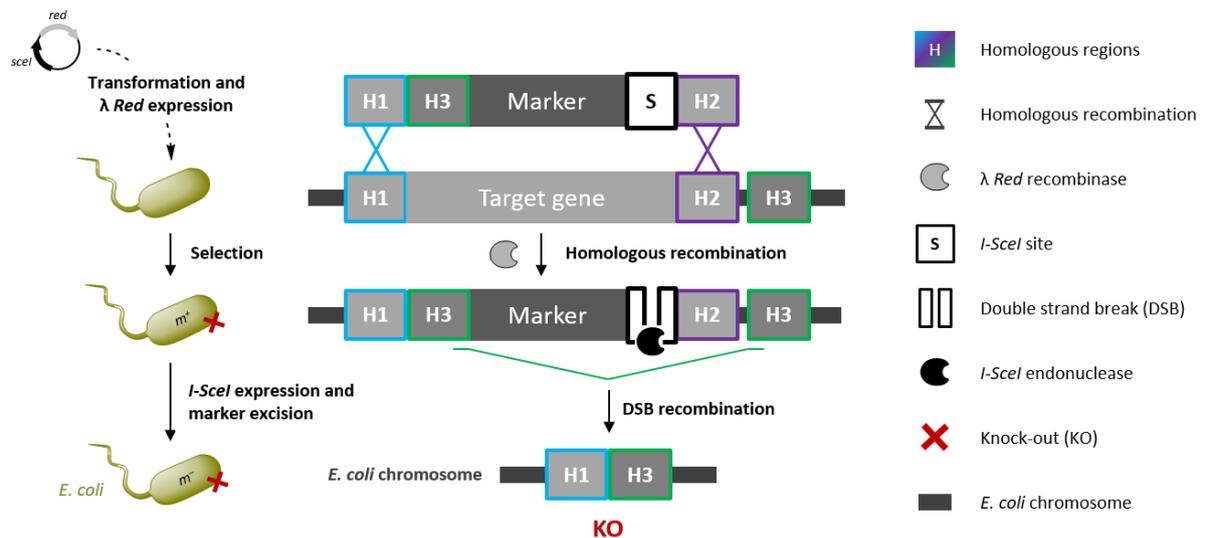


Figure C-14. *I-SceI* mediated double strand break (DSB) recombination. λ Red mediates genomic integration by homologous recombination between H1 and H2. A DSB is introduced by *I-SceI* at its recognition, which stimulates DSB recombination between adjacent homologous regions (H3) and the excision of the marker gene.^[114] However, to utilize *I-SceI* activity to enhance recombination efficiency for target KIs, it must be previously integrated at the desired locus.^[6]

With the λ Red system, target genes can be inserted into the genome of microorganisms at any (non-essential) locus *via* flanking homologous regions. However, the integration efficiency decreases for larger DNA molecules (1.5–2.5 kb),^[112, 114] although insertions of larger fragments (>7 kb) have been reported.^[115] [67a, b] Additionally, insertions (and subsequent excisions) leave scar sequences that might lead to undesired recombinations. Consequently, multiple knock-ins (KIs) utilizing the same mechanisms or similar homologous sequences can be troublesome.^[3, 6, 115] Sabri *et al.* constructed a set of KIKO vectors to integrate target genes in two different *E. coli* strains at three loci (*arsB*, *lacZ*, and *rbsA-rbsR*). KIKO vectors contain a multiple cloning site (MCS) for cloning insertion cassettes. The MCS is flanked by hairpin loops to isolate target genes from genomic DNA sequences at the site of insertion. This minimizes the interference of adjacent DNA elements with target gene expression. Such (undesired) interactions between native and synthetic genetic elements are termed context dependency^[6, 32, 111] and will be discussed later in this introduction. KIKO vectors encode λ Red recombinase. Long homology arms of 0.5 kb increased the insertion efficiency^[86] and achieved the genomic integration of larger cassettes (5.4 kb). *Flp* removed the antibiotic resistance markers (**Figure C-13**).^[115] Although the KIKO vector set provides a standardized tool for the integration of multi-gene pathways in only a few steps, the number of insertion loci is restricted and, importantly, unintended deletions and rearrangements can result *via* FRT scar sequences from previous rounds of genomic integration.^[6]

The discovery of CRISPRs in the *E. coli* genome had not been conclusive when first discovered^[119] but could be attributed to be part of the adaptive immunity of bacteria against bacteriophages.^[120-123] CRISPRs direct sequence-specific restriction of viral DNA to prevent viral replication.^[124] As a genomic engineering tool, Doudna and Charpentier utilized the system to efficiently cleave target DNA sequences with less off-site effects.^[125] Since then, the system has been successfully used to engineer the chromosomes of various organisms, mainly eukaryotes.^[25, 126-128] Besides the *Cas9* endonuclease, which, like *I-SceI*, induces DSBs, CRISPR also requires a

guide RNA (gRNA). The gRNA is a short synthetic RNA composed of a conserved loop necessary for *Cas9* binding and a customized spacer (i.e., targeting) sequence of 20 nucleotides, defining the genomic target to be modified. Conveniently, by altering the target sequence in the gRNA, the genomic target of *Cas9* can be changed. Coexpression of both, *Cas9* and the gRNA, is sufficient to knock-out chromosomal DNA (**Figure C-15A**).^[125]

So far, the CRISPR/*Cas9* system has not been extensively used in the metabolic engineering of prokaryotes since the DNA repair mechanisms of bacterial cells often fail to fix DSBs resulting from *Cas9* endonuclease activity.^[40-41, 82] However, the combination of CRISPR/*Cas9* with the λ *Red* system increased homologous recombination, hence, avoided the otherwise deleterious accumulation of DSBs in *E. coli* (**Figure C-15B**).^[3, 129]

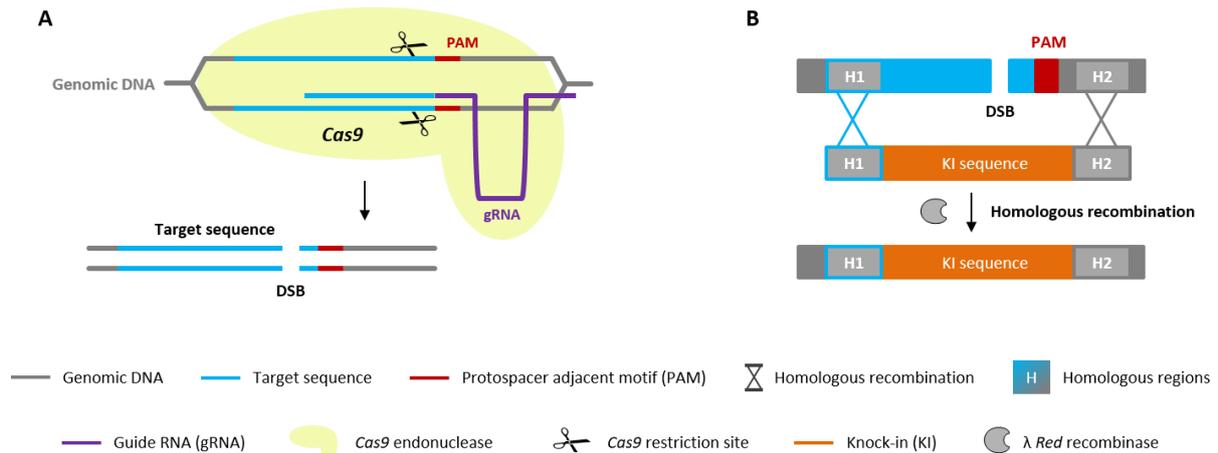


Figure C-15. CRISPR/*Cas9* genome editing. (A) Genome editing depends on a ribonucleoprotein (RNP) composed of the *Cas9* endonuclease and a gRNA, which directs the RNP to the target sequence in the genome. *Cas9* introduces a DSB upstream of the protospacer adjacent motif (PAM). (B) In the presence of a homologous DNA molecule containing the desired sequence to be integrated, λ *Red* facilitates homologous recombination, which is enhanced by the DSB previously introduced by *Cas9*.

By the simultaneous expression of λ *Red*, *Cas9*, and gRNA, whole synthetic pathways encoded on a linear PCR product (10 kb), for example, could be integrated into the *E. coli* genome.^[3, 130-131] Jiang *et al.* established a two-plasmid system to express λ *Red* and *Cas9* from one plasmid (pCas) and the gRNA from another vector (pTarget). This allows to reuse pTarget for multigene editing if the plasmid encodes various gRNAs for different targets^[129] and, in general, reflects the potential of CRISPR/*Cas9* as a versatile tool for specific genomic modifications, deletions and insertions. Hence, a broad applicability lies ahead for the CRISPR/*Cas9* system; not only as a tool for synthetic pathway modification but to cure genetic disorders and disease in humans.^[3]

A completely different genomic integration strategy is the use of mobile group II introns in combination with *Cre* recombinase as suggested by Enyeart and co-workers.^[132] Spliceosomal introns and short terminal repeat retrotransposons comprise roughly 45% of the human genome and are thought to be descendants of group II introns. These autocatalytic and retrotransposable RNA elements are found in all species with a highly conserved secondary structure^[133] and a multifunctional intron-encoded protein (IEP).^[3] Since IEP exhibits reverse transcriptase activity and stabilizes the intron RNA, it enables a remarkable mobility mechanism known as target DNA-primed reverse transcription. The intron RNA, which is spliced during messenger RNA (mRNA) maturation, is integrated into the target DNA, where it is reversely transcribed by the IEP, a process called retrohoming (**Figure C-16**).^[133] If the intron RNA encodes *loxP* sites,^[134-135] *Cre* recombinase can insert a whole donor plasmid with homologous *loxP* sites into the genome. The system was commercialized as a vector set named TargeTron[®] to knock-out target genes in bacteria (**Figure C-16**). One main advantage of this system is that the DNA segment to be inserted is not limited to its size, retaining full control over the locus of integration.^[3, 135]

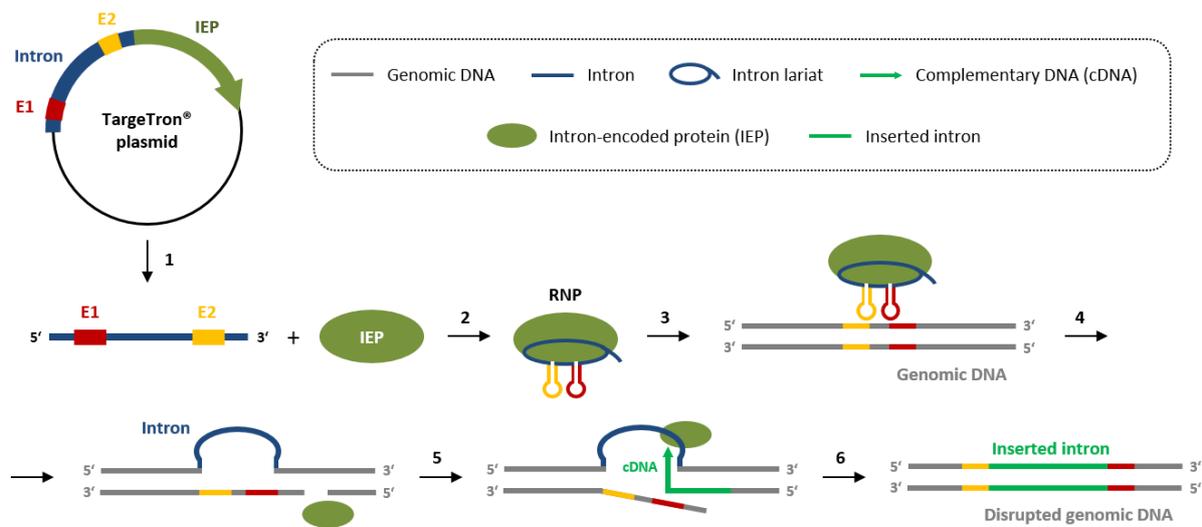


Figure C-16. Group II intron mobility mechanism. The TargeTron[®] system employs group II introns to specifically target genomic DNA and to integrate the intron by IEP activities. The intron containing directing oligonucleotides (E1 and E2) can be easily redesigned to target a different locus, likewise the intron can be re-designed to deliver another insert. The IEP exhibits maturase, endonuclease, and reverse transcriptase activity. (1) Upon expression, intron and IEP are produced from the TargeTron[®] plasmid. The IEP is translated by the host. (2) Following translation, the maturase activity splices the premature intron, which forms a lariat structure and binds to the IEP to form the RNP complex. (3) The RNP scans the genomic DNA. The target site is recognized by base pairing between the target DNA and the complementary E sites (yellow and red) of the intron. (4) Upon recognition, the intron RNA reverse splices into the top DNA strand, while the IEP nicks the bottom strand. (5) The IEP reversely transcribes a cDNA copy of the intron. (6) By assistance of the host DNA replication machinery, a stable insertion (green) is formed that disrupts the target sequence. The figure caption was adapted from Sigma Aldrich (2015).

Oberleitner *et al.* utilized the TargeTron[®] system to subsequently knock-out two genes, *nema*^[136-137] and *fadH*^[138-139], from the *E. coli* genome. The natively expressed enoate reductases (EREDs) N-ethylmaleimide reductase (*NemR*) and 2,4-dienoyl-CoA reductase (*DCR*) potentially interfered with a modular cascade for the transformation of cyclohexenol derivatives *via* intermediate cyclohexenones toward the corresponding lactones^[44] (**Figure C-17**).^[138] The double KO lead to an *E. coli* strain with reduced enoate reducing activity, offering potential to be used as a platform strain for the production of chiral compounds.

The described methods are not only capable of disrupting target genes; if homologous sequences are provided, desired genes, heterologous pathway elements, and whole metabolic pathways can be anchored in the genome. By simultaneous KI and KO, synthetic pathways can be genomically integrated and host genes, interfering with the *de novo* pathway, can be disrupted.^[6] The chromosomal insertion of a 2'-fucosyllactose pathway in *E. coli* by Baumgärtner *et al.* is only one example. To prevent the intracellular degradation of L-fucose, the expression cassette was integrated into the region encoding two degrading enzymes, *fucl* and *fuck*. The carbon flux toward the target compound was enhanced by providing an additional copy of *futC*, which produces 2'-fucosyllactose from GDP-L-fucose and lactose.^[6, 140] Agudo and Reetz inserted the ERED-encoding *yqjM* gene from *B. subtilis* into the *E. coli* genome, simultaneously knocking-out the endogenous *nema* locus (**Figure C-24**). Other than Oberleitner *et al.*, Agudo and Reetz employed λ Red recombination for the deletion of the same target gene.^[141]

The construction of industrially optimized strains requires the integration of whole synthetic pathways,^[103, 107] the deletions of competing endogenous host enzyme activities,^[42] and regulatory modifications.^[6, 129] Such large numbers of genome editing targets depend on efficient and versatile tools to perform time-saving sequential or even multiplex manipulations.^[129]

Metabolic engineering is widely applied to modify *E. coli* to produce biotechnologically relevant chemicals including biofuels and other compounds (e.g., ethanol,^[142] higher alcohols,^[143-144] rare or unusual sugars,^[140] amino acids,^[111, 145] fatty acids,^[146-147] terpenoids,^[148] and polyketides^[93, 149]).

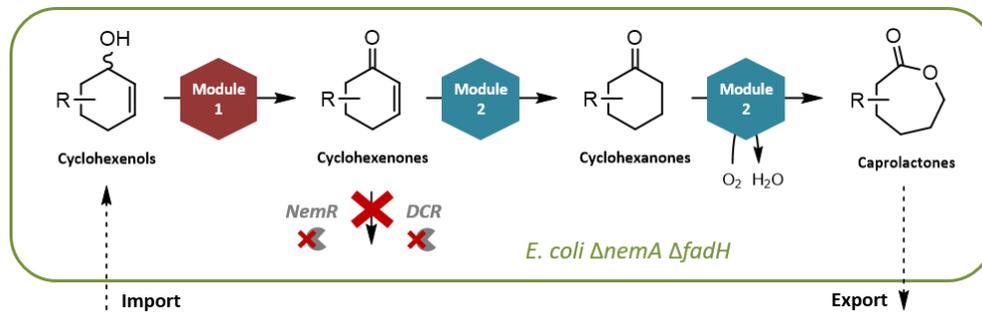


Figure C-17. *E. coli* double knock-out strain with reduced ERED activity. The TargeTron® system was applied to sequentially knock-out (red X) the genes *nemA* and the *fadH*, encoding the endogenous EREDs N-ethylmaleimide reductase (*NemR*) and 2,4-dienoyl-CoA reductase (*DCR*), respectively. Byproduct formation could be reduced for substituted substrates in the *NemR*- and *DCR*-deficient strain, while product yields improved. M1: (*R*)- or (*S*)- specific ADH (*Rhodococcus ruber* or *Lactobacillus kefir*); M2: ERED (*Pseudomonas putida*), cyclohexanone monooxygenase (CHMO; *Acinetobacter sp.*).^[138]

The genomic integration of entire (synthetic) pathways can negate the plasmid burden bestowed upon the host and combined modifications turn simple strains into highly engineered microbial cell factories. With the advent of systems metabolic engineering, combining metabolic and genetic engineering with systems biology and synthetic biology, scientists have already been able to efficiently remove bottlenecks to optimize enzyme cascades *in vivo* to meet industrial performance metrics such as high product titers and productivity.^[6, 36, 57, 150] Genomic modification tools have greatly advanced and will be able to resolve difficult to predict impacts on pathway performance such as the context dependency in the chromosomal environment.^[6]

C III Optimization of synthetic pathway elements

With the research progress made in the last decades, genetic and metabolic engineering tools for the modification of the bacterial hosts such as *E. coli* are available that not only enable the design and application of entire *de novo* pathways. The optimization of (synthetic) pathways is equally important and aims at the identification of bottlenecks, which can decrease host cell viability and product titers.^[6]

The heterologous expression of metabolic pathways, usually consisting of several genes, can lead to cellular stress responses, thus, negatively influencing pathway performance.^[6, 151-153] Advanced cloning techniques enable the assembly of multiple genes on individual plasmids to reduce the plasmid burden and genomic integration tools assist to completely abolish the plasmid burden by anchoring desired pathways in the genome. However, to minimize the metabolic load and to enhance the carbon flux through the target metabolic pathway, enzyme production must be balanced and even coordinated.^[57, 78, 141] An unbalanced enzyme stoichiometry can result in the accumulation of (toxic) cascade intermediates^[6] or a depletion of cofactors (e.g., flavin nucleotides).^[154-155]

Since proteins are produced by the interconnected processes of transcription and translation,^[13] both can be targeted and finely modulated to avoid the overproduction of single cascade enzymes and to balance overall

stoichiometry.^[6] Promoters are the main regulatory elements of transcription and translation is strictly dependent on a functional RBS (**Figure C-18**). Promoters as well as the RBS can be altered to directly change transcription and translation levels, respectively. However, adjacent sequences (and even distant ones) can greatly influence target gene expression, a phenomenon termed context dependency (**Figure C-18**).^[32, 111, 115, 156] The RBS perfectly illustrates context dependency: The only 6 bp long core sequence is located upstream of an open reading frame (ORF), followed by a short spacer sequence of 5 to 9 bp, which is critical for the initiation of translation.^[157-159] Additionally, the 50 bp around the RBS modulate its efficiency and the stability of the transcribed mRNA influences translation.^[73] Whereas the promoter and the RBS are commonly modified to regulate gene expression in bacteria, other regulatory elements have not been used extensively such as riboswitches.^[6, 160-161] More recently, small regulatory RNAs (sRNAs) have been introduced to fine-tune gene expression.^[162-163] Self-regulatory elements can be exploited as metabolic engineering tools (e.g., auto-inducers) but might resemble obstacles in biotechnology processes due to feedback inhibition mechanisms.^[73] Independent of the regulatory element introduced, their influence on the genetic context will modify transcription, translation, or both.^[3, 6, 32, 158, 164]

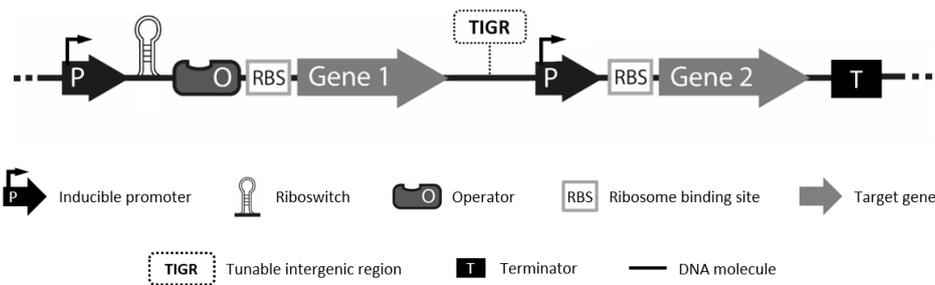


Figure C-18. Genetic regulatory elements. The expression of ORFs can depend on adjacent regulatory elements and include (inducible) promoters, operator sequences, the RBS, and terminators. Riboswitches are structural mRNA elements that relay the binding of a small molecule, for example, *via* conformational changes into a biological function in a concentration dependent manner.^[160-161, 165] The tunable intergenic region (TIGR) can be modified to influence downstream regulatory elements such as the RBS or influence mRNA stability. The figure was adapted from T. Bayer *et al.* (2015).^[6]

C III.1 Optimization on the levels of transcription and translation

Enzymes and their arrangement in cascades build up biosynthetic pathways, which are embedded in highly regulated metabolic networks. Genes encode these enzymes and other cellular proteins. To ensure vital functions including the replication of DNA, protein biosynthesis, bacterial growth, and energy charge homeostasis, environmental stimuli (e.g., nutrients) as well as intracellular signal molecules such as guanosine tetrakisphosphate or pentakisphosphate [(p)ppGpp]^[166] lead to alterations in gene expression.^[13, 167] The alarmone (p)ppGpp, for example, is involved in the stringent response in bacteria, reacting to amino acid^[168] and fatty acid^[169] depletion, heat shock,^[170] and other stress conditions.^[171-172] Stringent control involves extensive changes in gene expression, the transcriptional inhibition of ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), the activation of biosynthetic genes, and the translational repression of non-vital proteins by interfering with 30S and 50S subunit interactions.^[166, 172]

In the context of synthetic pathway engineering, both transcription and translation are suitable targets to balance pathway enzyme production and install a beneficial enzyme stoichiometry that will enhance the flux through the *de novo* pathway. Many regulatory elements (**Figure C-18**) are available to regulate gene expression from the synthetic biology (online) toolbox: <http://parts.igem.org>.

A group of proteins responsible for the initiation of transcription recognizes promoter sequences, which serve as standard regulatory elements to fine-tune enzyme production. For the heterologous production of pathway enzymes, inducible promoters routinely replace native ones. Promoter strength has profound influence on the amount of enzyme produced and, thus, impacts pathway enzyme stoichiometry (**Figure C-19A**).^[71] Frequently used promoter systems such as P_{BAD} or P_{T7lac} are induced by L-arabinose and isopropyl β -D-1-thiogalactopyranoside (IPTG), respectively, and will result in the expression of target genes at high-levels.^[173] IPTG, an α -lactose (α -Lac) analogue, is not readily metabolized by *E. coli* and grants prolonged expression levels. It binds to the *LacI* repressor complex, which is released from its target DNA and, consequently, allows the RNA polymerase to start transcription. Such DNA segments are commonly known as operators, which bind transcription factors (TFs) that prevent downstream gene transcription (**Figure C-19B**).^[174] The *lac* operon of *E. coli* had been the first discovered operon,^[174-176] which built the foundation for today's standard pET system for the heterologous production of proteins.^[177] The pET vectors feature the *LacI*/ P_{T7lac} promoter, which turned out to be highly efficient on the level of transcription. However, transcripts are often translated into vast amounts of inactive (insoluble) proteins.^[178] The combination of P_{T7lac} and the *LacI* repressor exhibits leaky expression, another drawback of the pET system.^[71, 179-180] Promoters for metabolic engineering must show tight control to avoid unintentional metabolic load. Therefore, tightly controlled systems such as *AraC*/ P_{BAD} ^[50] or the more recently introduced cumate gene switch-based expression system may be more suitable to control the production of target pathway enzymes.^[6, 180] Nonetheless, other regulatory elements take precedence over the use of operator sequences and there are only a few literature examples related to metabolic engineering.^[6, 73]

Brautaset *et al.* compared an expanded set of promoters for gene expression in bacteria, which not only included promoters regulated by the addition of sugars such as L-arabinose (L-Ara) and L-rhamnose (L-Rhm); promoters can be also induced by small organic molecules including alkanes, substituted benzenes, and even peptides.^[6, 181] Auto-responsive promoters do not depend on the addition of certain chemicals and represent promising alternatives to established promoter systems. They respond to environmental stimuli (e.g., oxygen^[182] or light^[183-185]) or metabolites produced during microbial growth or stress conditions.^[73, 186] Another process than the stringent response to starvation in bacteria is quorum sensing.^[187-188] It, too, results in the coordinated expression of genes, which leads to the concerted behavior of a whole cell population. It is triggered by specific signal molecules (i.e., auto-inducers).^[6] The auto-inducer concentration can be proportional to the cell density, for example, and was successfully tested to heterologously express different enzymes.^[189-192] Since the production of target genes only occurs at a certain cell density threshold, the optimization of induction initiation can be avoided, eliminating continual monitoring of bacterial growth prior to induction.^[6, 73] Newly discovered promoters will complement the toolbox for transcriptional control for metabolic engineering purposes on a fundamental level. The modification and combination of established promoter systems will enable the precise tuning of gene expression and the coordinated transcription of multiple pathway genes, respectively.^[193-196]

Whereas promoters were thoroughly investigated and have been employed as regulatory elements, their transcriptional counterparts (i.e., terminators) have been widely neglected.^[73] Terminators intrinsically stop transcription.^[99, 197] Terminators adopt stable secondary hairpin structures, formed by palindromic GC rich regions of the newly synthesized mRNA.^[99] A chain of at least six uracil (U) residues immediately follows the hairpin structure. The lower binding energy of the U stretch destabilizes the RNA-DNA duplex, allows it to unwind, and to dissociate from the RNA polymerase.^[98] Depending on the composition of hairpin sequence, its exact secondary structure, and the length of the U chain, both termination strength and fidelity are influenced and, thus, modifies translation. Since adjacent nucleotides influence the terminator functionality, terminators are another example for context dependency.^[3] In addition to stopping transcription, terminators stabilize their own mRNA.^[198] Hence, engineering transcriptional terminators may provide another layer of control to regulate both transcription and translation since the half-life of an mRNA correlates with translation frequency.^[6]

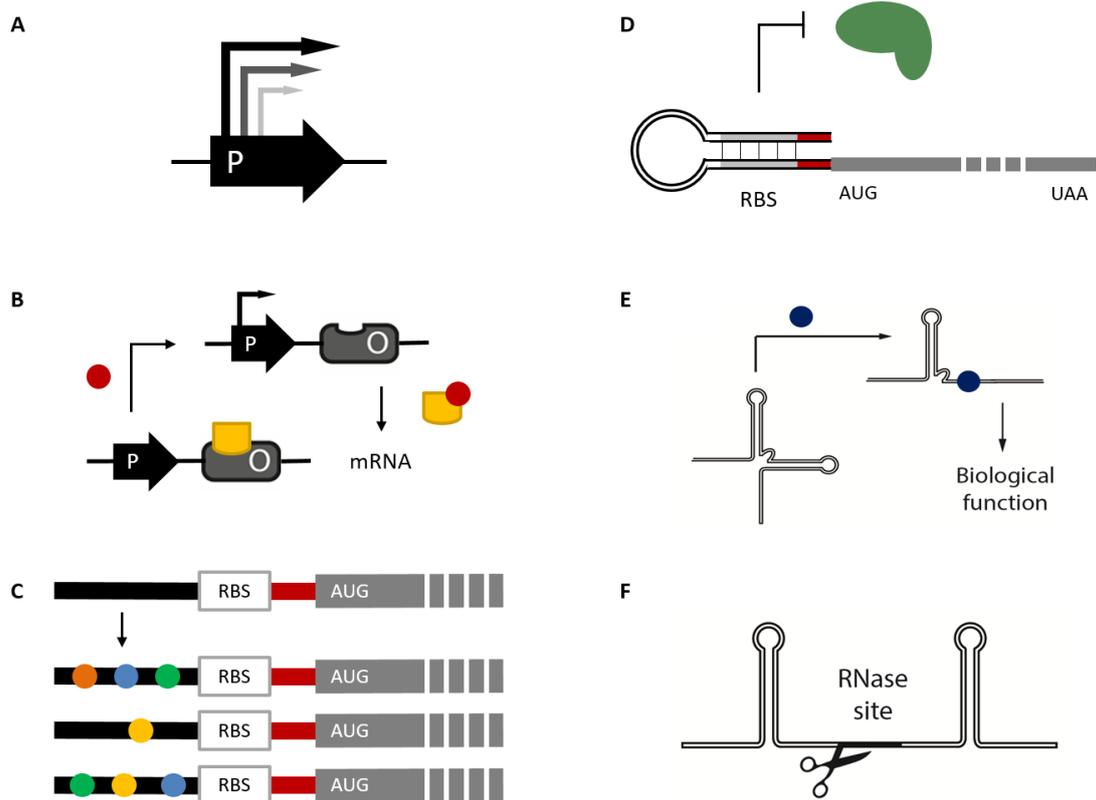


Figure C-19. Transcriptional and translational regulation mechanisms. (A) Promoter strength regulates the rate of transcription and impacts subsequent mRNA translation. (B) Operators recruit repressor proteins (yellow) that regulate the activity of upstream promoter sequences. In the presence of inducer molecules (red), which bind to the inhibitory proteins, the RNA polymerase proceeds and transcription continues. (C) Translation efficiency can be modified by sequence alterations close to the RBS. (D) Sequestration of the RBS due to the formation of stable secondary structures can strongly interfere with translation initiation. (E) Riboswitches are structural mRNA elements. The binding of small molecules (dark blue) leads to conformational changes, which affect the biological activity of the mRNA. (F) The presence of RNase sites can influence transcript stability, thus, impacts subsequent translation. The figure was adapted from T. Bayer *et al.* (2015).^[6]

Besides modifying the persistence of mRNAs *in vivo*, a more routinely applied strategy to control translation targets the sequence of the RBS (**Figure C-19C**). The recognition of the RBS by the ribosome is necessary for protein biosynthesis. The RBS, sometimes referred to as the Shine-Dalgarno (SD) sequence in bacteria, is a short sequence upstream of the ATG start codon, encoding a methionine.^[159, 199-200] As noted above, variations in the core sequence and the nucleotides around the RBS modulate translation efficiency.^[32, 158, 164] Furthermore, stable secondary structures of the RBS or close to it can drastically reduce the rates of translation or even prevent it (**Figure C-19D**).^[201] To avoid a non-functional synthetic RBS, bioinformatic tools are available to assist their design.^[6, 158, 202-203] Wang *et al.* examined a method called multiplex automated genome engineering (MAGE), which was applied to engineer a recombinant *E. coli* strain to accumulate lycopene.^[36, 204] Targeting the 20 genes responsible for lycopene biosynthesis, the RBS sequences were modified through allelic replacements, using oligo-nucleotides containing degenerated RBS sequences: DDRRRRRDDDD; whereas D = adenine (A), guanine (G), or thymine (T) and R = A or G. High similarity between the replaced RBS regions and the proposed SD sequence (AGGAGG) resulted in enhanced translation efficiencies.^[6, 73, 204] The unbalanced translation of several enzymes in a biosynthetic mevalonate to amorpha-4,11-diene (amorphadiene) pathway lead to the accumulation of intermediates. To overcome this bottleneck, Nowroozi *et al.* applied a combinatorial approach to screen suitable RBSs for different cascade enzymes. By taking growth, expression levels, and the accumulation

of intermediates into account, Nowroozi and co-workers could improve the production of amorphadiene, a precursor for the anti-malarial drug artemisinin (**Figure C-1**),^[205] about 5-fold. (**Figure C-20**).^[6, 206]

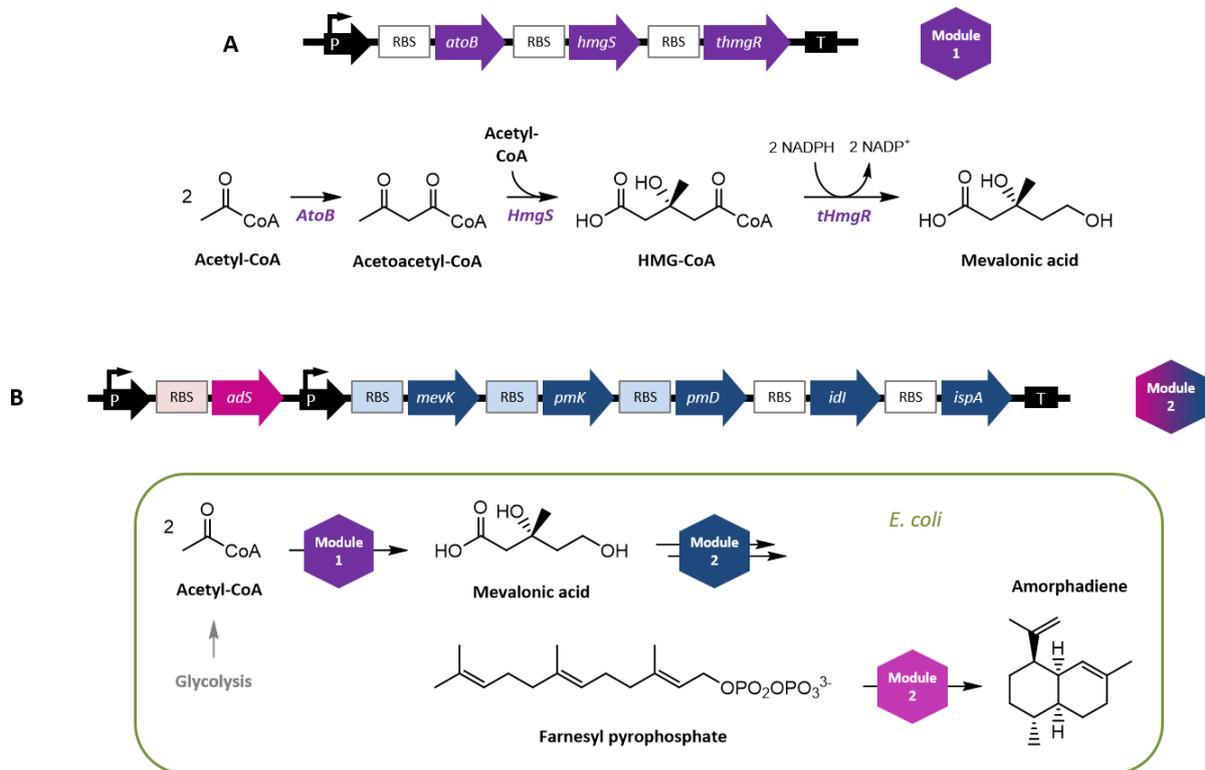


Figure C-20. Amorphadiene synthesis. The upper and the lower mevalonate pathway were arranged in modules. Incorporation of the amorphadiene synthetase (*adS*) gene and balancing enzyme stoichiometry by RBS variants (colored RBS boxes) lead to an increased production of amorphadiene. **(A)** Upper mevalonate pathway (violet) to convert intracellular acetyl-CoA into mevalonic acid. M1: acetoacetyl-CoA thiolase (*AtoB* from *E. coli*), 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase (*HmgS*), truncated HMG-CoA reductase (*tHmgR*; *S. cerevisiae*) **(B)** The lower mevalonate pathway (blue) including the *adS* gene (pink) converts mevalonic acid *via* farnesyl pyrophosphate into the target molecule. Intermediate reactions not shown for clarity. M2: amorphadiene synthetase (*AdS*), mevalonate kinase (*MevK*), phosphomevalonate kinase (*PmK*), mevalonate pyrophosphate decarboxylase (*PmD*; *S. cerevisiae*), isopentenyl pyrophosphate isomerase (*Idl*), farnesyl pyrophosphate synthase (*IspA*; *E. coli*).^[206]

Codon usage also influences the translation efficiency due to the availability of tRNAs corresponding to its codon on the mRNA, which is specific for different species. The use of codon-optimized genes is standard if synthetic pathway enzymes from higher organisms are heterologously expressed in host organism such as *E. coli*.^[6, 207-210]

A post-transcriptional and -translational regulation strategy in bacteria is based on RNA molecules called riboregulators or sRNAs. sRNAs belong to a small subset in the group of non-coding RNAs in prokaryotes and eukaryotes and can exhibit different structural, regulatory, or even enzymatic functions.^[6, 211-213] Noncoding RNAs can also comprise a portion of mRNAs such as riboswitches (**Figure C-19E**). sRNAs can bind chemicals^[214] and respond to environmental changes such as temperature,^[215] which provoke a change in gene expression;^[216] the translation of proteins can be either activated or repressed.^[6, 217-220] Different riboregulators have already been applied successfully.^[221-223] Kang *et al.* artificially overexpressed a sRNA (*rhyB*) to accumulate the metabolic intermediate succinate^[224] and to produce polyhydroxyalkanoates.^[225] More recently, Na *et al.* designed synthetic sRNA arrays to knock-down target genes to increase the production of both tyrosine and cadaverine in engineered *E. coli* strains.^[162] With the sRNA approach, tyrosine titers ($2 \text{ g} \cdot \text{L}^{-1}$) as high as previously reported by

Juminaga *et al.* could be achieved. In contrast, Juminaga and co-workers employed a plasmid-based expression system and had to tune promoter strength and engineer synthetic operons.^[226]

Finally, intergenic regions can be designed to stabilize mRNA secondary structures or modified to encode, for example, RNase sites (**Figure C-19F**).^[227-228] Pflieger *et al.* tuned the expression levels of multiple genes in operon configuration (**Figure C-6**) by customizing TIGRs (**Figure C-18**). Balancing enzyme production of a heterologous pathway producing mevalonate could be increased 7-fold.^[73, 229] The incorporation of RNase sites in the 3'-untranslated region (3'-UTR) could also modulate the expression levels of a BVMO from *Pseudomonas fluorescens* (*P. fluorescens*) in *E. coli*.^[230]

The advent of systems metabolic engineering provides various strategies for the optimization of heterologous enzymes production on the levels of transcription and translation.^[6] The combination of different tools lead to engineered microorganisms, overproducing both natural compounds and non-natural chemicals.^[231] The utilization of synthetic sRNAs offers great potential as alternatives to conventional gene KO strategies. KOs cannot be easily undone, whereas the knock-down of target genes is reversible.^[36] The introduction of sRNAs only conveys a minimal metabolic burden on the host due to their natively small size. Expression is simple and sRNAs can be applied to simultaneously tune the expression levels of multiple target genes, allowing gene to function studies of essential genes that cannot be deleted.^[6, 36, 232] Gene silencing was successfully applied to shift the production of P3HB from glucose toward poly-4-hydroxy butyrate in *E. coli*, for example.^[232] Guidelines for the rational design of novel sRNAs have already been proposed and will assist and accelerate the development of future metabolic engineering applications to boost the performance of artificial biosynthetic pathways.^[233]

C III.2 Engineering enzyme performance

Optimization strategies target different molecular levels and cellular processes including transcription and translation. The modification of regulatory elements, thereby changing the genetic context of associated genes, can enhance (or decrease) expression levels and will produce different amounts of the target proteins. However, unbalanced enzyme stoichiometry is not the only reason potentially impairing the carbon flux through *de novo* pathways. The accumulation of cascade intermediates might result from unfavored enzyme kinetics or poor substrate acceptance. Furthermore, biocatalysts regularly failed to meet industrial process criteria including pH and thermostability, and organic solvent tolerance, for example.^[6, 234] In the 1990s, Arnold and Stemmer pioneered advanced enzyme engineering methods to overcome these shortcomings and efficiently modified the amino acid sequence of biocatalysts by mimicking Darwinian evolution in a test tube at fast pace.^[6, 17, 235] Two strategies have been extensively used: directed evolution^[236-237] and rational design.^[238-240]

Enzyme engineering by directed evolution involves iterative cycles of gene mutagenesis, expression, and selection of mutant enzymes exhibiting the desired property. Commonly used mutagenesis techniques include error-prone PCR, saturation mutagenesis, combinatorial active-site saturation testing (CASTing), and DNA shuffling (**Figure C-22**).^[234] In contrast, rational design largely depends on the availability of structure-function relationships of target enzymes. Initiative approaches were based on the comparison of sequence homology and aimed at enzyme mutants with improved solubility, thermostability, or organic solvent tolerance.^[73, 241] The groups of Arnold, Bornscheuer, Clapés, Fessner, Reetz, Turner, and others have impressively applied directed evolution and rational design to improve the properties of biocatalysts and even equipped them with completely new functions.^[17, 234, 242]

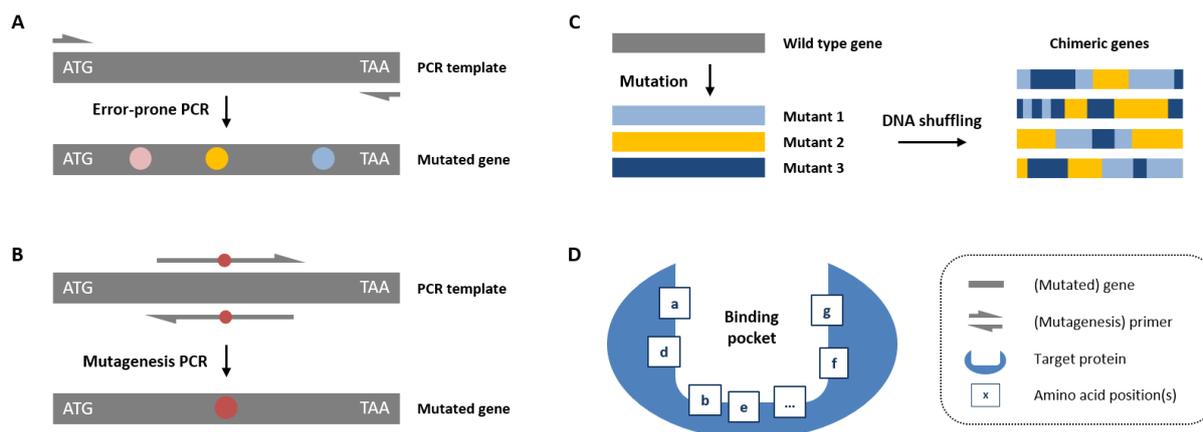


Figure C-21. Commonly applied mutagenesis methods. (A) Error-prone PCR employs conditions to enhance the intrinsic error rate of PCRs and randomly targets the whole amplicon. (B) Directed mutagenesis by designed primers carrying the desired mutation (e.g., QuickChange protocol) targeting specific base pairs. (C) DNA shuffling recombines portions of different mutant genes to generate chimeric genes. (D) CASTing: Amino acids (a, b, c, ...) of the enzyme binding pocket are identified by X-ray structure or homology models and systematically targeted by saturation mutagenesis. (C) and (D) were adapted from M. T. Reetz (2013).^[234]

Some of these achievements were accomplished *en route* to commercially relevant biocatalytic processes. One excellent example was the synthesis of ethyl-(*R*)-4-cyano-3-hydroxy butyrate, a key intermediate in the synthesis of atorvastatin (Figure C-22). Atorvastatin is a cholesterol-lowering blockbuster drug sold under the name Lipitor[®] with a peak sales volume exceeding 10 billion US\$ in 2011. The key enzyme towards ethyl-(*R*)-4-cyano-3-hydroxy butyrate is a halohydrin dehalogenase (HHDH) from *Agrobacterium radiobacter* (Figure C-22).^[6] The combination of the statistical analysis of protein sequence activity relationships (proSAR) and a recombination-based directed evolution approach tremendously improved the cyanation capability of the HHDH. Under process conditions, the productivity of the target intermediate could be increased about 4 000-fold.^[243] Another process utilized a DNA shuffling approach and yielded the same atorvastatin intermediate (Figure C-22).^[244-245] HHDH activity was increased >2 500-fold compared to the unengineered enzyme and included the application of a ketoreductase (KRED) to reduce the precursor ethyl-4-chloroacetoacetate to ethyl-(*S*)-4-chloro-3-hydroxy butyrate, which was subsequently transformed by HHDH into ethyl-(*R*)-4-cyano-3-hydroxy butyrate. Additionally, a glucose dehydrogenase (GDH) was applied to recycle the cofactor NADPH (nicotinamide adenine dinucleotide, reduced; Figure C-22).^[6] Several rounds of DNA shuffling improved GDH and KRED activity 13- and 7-fold, respectively.^[16]

Chiral amines are important motifs in drug molecules and the development of efficient synthetic routes to optically pure derivatives remains a major goal for the pharmaceutical industry. ω -transaminases (ω -TAs) and monoamine oxidase variants from *Aspergillus niger*, for example, represent two important families of biocatalysts to access optically pure chiral amines. TAs mediate the pyridoxal phosphate (PLP) dependent, reductive amination of prochiral ketones and allows the synthesis of the corresponding (*S*)- or (*R*)-amines.^[246-247]

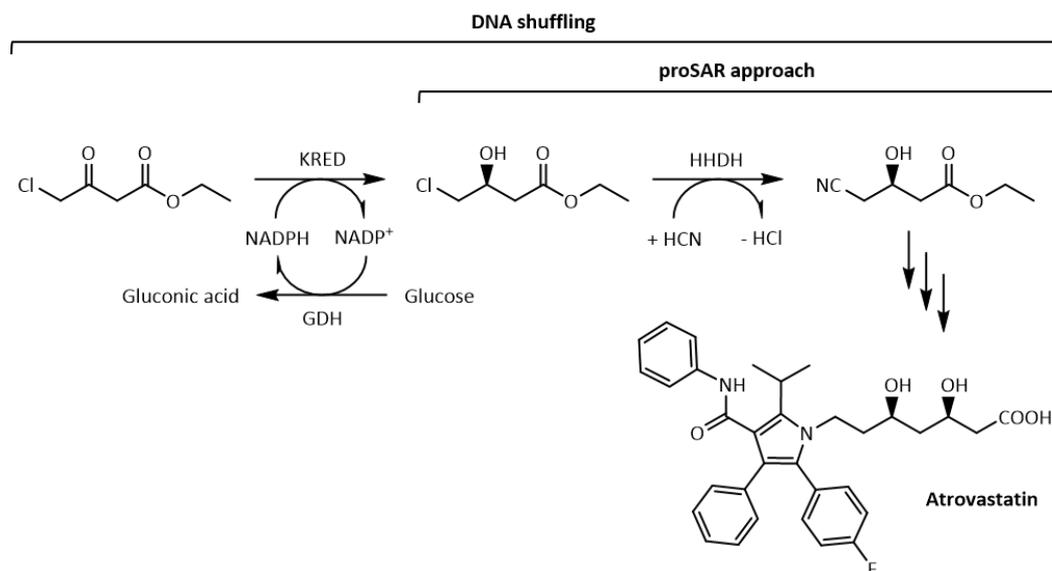


Figure C-22. Improved atorvastatin precursor synthesis by directed evolution of enzymes. A hybrid approach of directed evolution and proSAR greatly increased the HHDH activity leading to the target intermediate ethyl-(*R*)-4-cyano-3-hydroxy butyrate (top right) in the synthesis of atorvastatin (bottom).^[243] Another approach employed a DNA shuffling methodology not only to improve the HHDH activity; a KRED and a GDH were engineered for the reduction of ethyl-4-chloroacetoacetate (top left) to ethyl-(*S*)-4-chloro-3-hydroxy butyrate (top center) and cofactor recycling, respectively.^[16]

By combining *in silico* design and a potpourri of protein engineering technologies, Savile *et al.* equipped a transaminase lacking activity for prositagliptin, a ketone precursor for the type II diabetes drug sitagliptin (**Figure C-23**).^[248] One initial objective driving this development was the omission of problematic Pd/C catalyst separation after chemical reductive amination,^[249] which could partly be solved by second generation processes employing a Rh catalyst.^[250] Furthermore, the transaminase was engineered to meet process parameters including the tolerance toward dimethyl sulfoxide (DMSO), acetone, and isopropyl amine (*i*-PrNH₂) at elevated temperatures. By employing the mutant transaminase, productivity could be increased significantly.^[251]

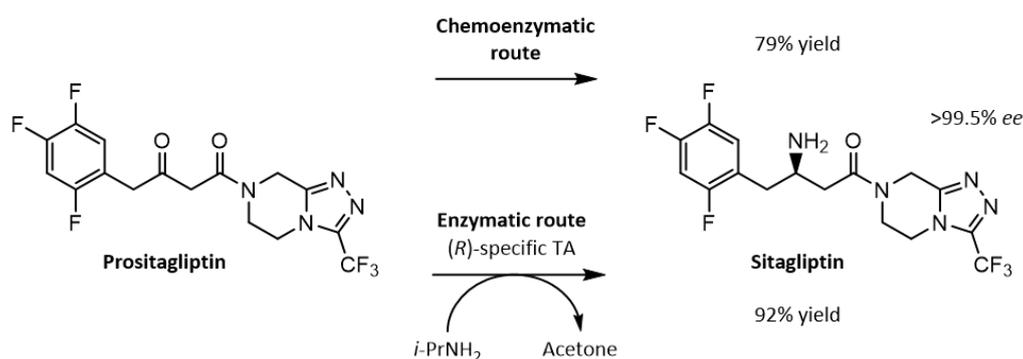


Figure C-23. Superior enzymatic route toward sitagliptin. Directed evolution of an (*R*)-specific TA led to the acceptance of the sterically demanding prositagliptin and meeting process conditions including organic solvent tolerance and higher reaction temperatures; *i*-Pr = isopropyl. The figure was adapted from T. Bornscheuer *et al.* (2012).^[17]

Agudo and Reetz assembled an *in vivo* redox cascade encompassing two successive regioselective oxidations catalyzed by an engineered cytochrome P450 enzyme from *Bacillus megaterium* (P450-BM3). The two regioselective oxidation steps were followed by stereoselective olefin reduction by (*R*)- or (*S*)-selective mutants

of the ERED *YqjM*.^[141] Both enzymes, *P450-BM3* and *YqjM*, have been subjected to directed evolution. *P450-BM3* mutants showed increased activity toward the substrate 1-cyclohexene carboxylic acid methyl ester. *YqjM* mutants installed the desired stereoselectivity in the target compounds. Apart from enzyme engineering strategies, the *nemA* gene coding for the endogenous ERED *NemR*, was knocked-out from the *E. coli* genome by λ *Red* recombination. As noted earlier, the natively expressed *NemR* interfered with the activity of the heterologously expressed *YqjM* mutants (**Figure C-24**).^[6]

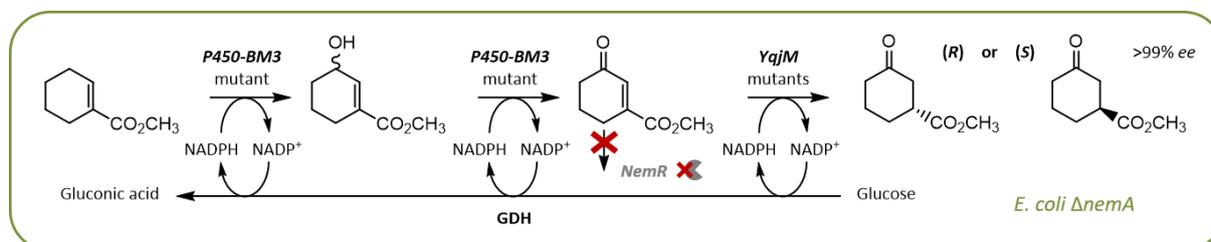


Figure C-24. Synthetic enzyme cascade optimization by directed evolution. Optimization of a redox cascade reaction consisting of consecutive regioselective oxidations by a *P450-BM3* mutant and the stereoselective olefin reduction by either a (*R*)- or (*S*)-selective *YqjM* mutant. Cofactor recycling was performed by GDH. The KO of endogenous *NemR* (red X) abolished unselective byproduct formation. The figure was adapted from R. Agudo and M. T. Reetz (2013).^[141]

Another recent example of directed evolution reversed the enantioselectivity of a phenylacetone monooxygenase (PAMO) to perform the asymmetric sulfoxidation of prochiral thioethers. Saturation mutagenesis resulted in an engineered PAMO containing four single point mutations. The mutations synergistically turned the wild type enzyme with (*S*)-preference for sulfoxide formation (90% *ee*) into the mutant enzyme with (*R*)-preference (95% *ee*).^[6, 252]

Protein engineering methods were not only employed to optimize enzymes catalyzing functional group transformations. Carbon-carbon bond formations are key reactions in synthetic organic chemistry to construct the carbon framework of complex molecules.^[253] Aldol reactions have been proven to be a powerful methodology for C–C bond formation reactions amongst others. Particularly, aldolases have been subjected to protein engineering since they offer a green, uniquely regio-, and stereoselective tool for carbonyl condensations (**Figure C-25**). New aldolases derived from the transaldolase scaffold (e.g., fructose-6-phosphate aldolase) have been shown to be unusually flexible regarding their substrate scope. In general, aldolases display a relaxed scope toward acceptors but are very specific for the donor molecule.^[254]

An intensively studied aldolase is the D-fructose-6-phosphate aldolase from *E. coli* (*Fsa1*). *Fsa1* exhibits promiscuous nucleophile selectivity, accepting glycolaldehyde as well as several hydroxylated ketones.^[254] The wild type enzyme catalyzes homo- and cross-aldol additions of glycolaldehyde, accessing small sugars including D-threose, L-glyceraldehyde, and functionally related derivatives.^[255] The substrate specificity of *Fsa1* variants has been tailored by structure-guided engineering of the active site to expand the application scope. Szekrenyi *et al.* assembled hexoses from simple precursors in a stereodefined way (**Figure C-25A**). Mutations at positions L107 and, particularly, A129 were found to be crucial to both donor activity and selectivity, which resulted in *Fsa1* variants with relaxed donor preference (**Figure C-25B**).^[256-259]

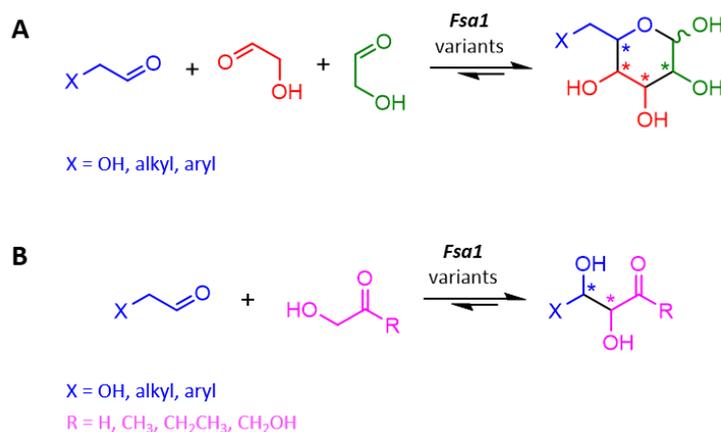


Figure C-25. Synthesis of sugar derivatives from achiral starting materials by *Fsa1* mutants. (A) Stereoselective trimerization of glycolaldehyde (red and green) and derivatives (blue) by engineered *Fsa1* aldolase from *E. coli*. (B) Rational design of wild type *Fsa1* lead to variants with relaxed donor specificity (pink). (A) was adapted from A. Szekrenyi (2015).^[260]

Other than changing the stereoselectivity or enhancing the activity toward known and novel (sterically) demanding substrates, protein engineering was applied to equip enzymes with completely new functions. Scientists in the Arnold group engineered the previously introduced *P450-BM3* by site-saturation mutagenesis and screened variants for enhanced cyclopropanation activity.^[261-262] Cyclopropanation activity was a new function desperately needed for the enantioselective synthesis of levomilnacipran,^[263] an antidepressant sold as Fetizma® (**Figure C-26**). Zhang *et al.* also demonstrated the possible alteration of the *P450-BM3* scaffold to fine-tune the acceptance of other demanding substrates and established non-natural biocatalytic cyclopropanations as a complementing tool in organic synthesis.^[264-265]

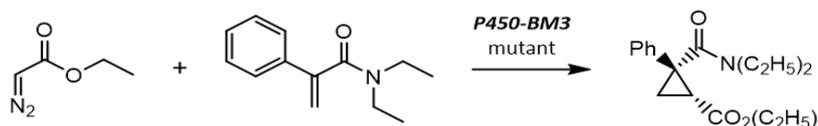


Figure C-26. Levomilnacipran synthesis by enzymatic cyclopropanation. The wild type *P450-BM3* from *Bacillus megaterium* was adapted by directed evolution for non-natural cyclopropanation reactions. Mutant *P450-BM3* was used to synthesize the antidepressant levomilnacipran on gram-scale; Ph = phenyl.^[264]

The selected examples above undeniably demonstrate the power of protein engineering technologies. Directed evolution and rational design have dramatically accelerated the test tube evolution of biocatalysts. The catalytic properties of enzymes could be greatly improved^[266] and engineered enzymes operate under harsher conditions (e.g., high temperatures, organic solvents), accept non-natural substrates,^[251] or even exhibit new functions including Diels-Alder reactions^[267] or cyclopropanations.^[6, 264]

The engineering of high performance biocatalysts depends on steadily increasing numbers of amino acid substitutions. Whereas one to five mutations were typical in the early 2000s, 35 amino acid exchanges were made on average during protein engineering in 2010.^[17] The directed evolution of the HHDH for the precursor synthesis of atorvastatin (**Figure C-22**) changed ≥ 35 of the 254 amino acids^[243] and 27 of 330 residues were substituted in the transaminase variant accepting prositagliptin (**Figure C-23**).^[17, 251] Consequently, the sequential introduction of mutations is not feasible. Methods to add multiple mutations simultaneously relied on statistical or bioinformatic methods like proSAR.^[243] Other approaches limited the location of amino acid substitutions in

the protein sequence to the active site and the types of changes to those known from sequence comparisons to occur often at these sites.^[268]

Despite these advances, major challenges remain since changing 30–40 amino acids and screening tens of thousands of candidate mutants certainly is laborious and time-consuming.^[17] An immediate solution to this problem has been the development of high-throughput screenings (e.g., fluorescence-activated cell sorting; **Table C-2**).^[152, 269-273]

Table C-2. Screening and selection strategies

Method	Genotype/phenotype linkage	Detection	Library size (throughput)	Advantages	Disadvantages
Colonies on solid media	Spatial separation of variants	Manual inspection of visible signals ^[a, b]	10 ² – 10 ⁴ (low)	Straightforward implementation	Laborious; qualitative and semi-quantitative
Isolated liquid cultures	Spatial separation of variants	Biochemical assays ^[b, c]	10 ² – 10 ⁴ (low)	Straightforward implementation; flexibility in reporter/detection method; quantitative	Laborious
Cell surface display	Compartmentalization (cell membrane)	ELISA, FACS ^[b, d]	10 ⁸ (high)	Yeast display: eukaryotic gene expression, post-translational modification	Limited application to certain biocatalysts ^[e]
IVC ^[f]	Emulsions ^[g]	FACS	10 ⁷ – 10 ⁸ (high)	Quantitative	IVC and microfluidic techniques require expertise and optimization

^[a] For example, fluorescence or colorimetric assays of surrogate substrates. ^[b] Selection might also be coupled to organismal fitness/survival or auxotrophy. ^[c] GC, HPLC, nuclear magnetic resonance (NMR), mass spectrometry (MS), fluorescent or colorimetric (microplate) assays. ^[d] Enzyme-linked immunosorbent assay (ELISA), fluorescence-activated cell sorting (FACS). ^[e] For example, esterases or proteases.^[273] ^[f] *In vitro* compartmentalization. ^[g] For example, water-oil emulsions or (self-assembling) polyelectrolyte shells.^[273] The table was adapted from M. S. Packer and D. R. Liu (2015).^[273]

The engineering of enzymes solved limitations such as low stability *in vitro* and low activity toward unusual substrates and led to biocatalytic processes on industrial scales. Future protein engineering will have to face challenges emerging through the interfacing of individual biocatalysts with other enzymes and proteins in synthetic enzyme cascades;^[17] a problem that has already been addressed by synthetic biology tools to enhance the performance of *de novo* pathways in living cells.

C IV Improving synthetic pathway performance in the cellular context

As early as in 2001, Schmid *et al.* predicted an increasing portfolio of biocatalysts in application and the continuous regeneration of expensive cofactors (e.g., NADP⁺/NADPH) by 2010.^[30] However, by then, the use of non-metabolizing cells for biocatalysis had proven to be more challenging than estimated and the use of engineered enzymes shifted to crude and semipurified forms. Whereas historically whole cells had offered a simple and effective option for cofactor recycling and enhanced enzyme stability *in vivo*, protein engineering and the use of single enzymes was considered more economic and convenient. Engineered enzymes tolerated harsher reaction conditions and the absence of cell membranes eliminated potential limitations by diffusion.^[17]

Today, two- and three-step biotransformations are routinely accomplished by the combination and engineering of enzymes from different sources *in vivo*. Recombinant microbial whole cell biocatalysts (i.e., ‘designer cells’)

have been successfully constructed, optimized, and applied to produce value-added chemicals.^[6-8, 43, 274] Under this approach, drawbacks of *in vitro* preparations are dispensable. Required enzymes are simultaneously produced by host cells. Hence, cell disruption, clarification, and concentration of target enzymes is not necessary (**Figure C-2**). The separation of the biocatalyst after biotransformation can easily be done by flocculation and filtration of the biomass.^[275] Substrate and product transfer limitations through cell membranes have hardly been observed using lyophilized biocatalysts, for example.^[276]

At present, designer cells contain synthetic pathways consisting of up to nine heterologously expressed enzymes.^[43, 277] In living cells (with intact cell membranes), bottlenecks can originate from limited permeability of substrates and/or products.^[278] The leaking of pathway intermediates out of the cell, their accumulation due to different kinetics or an unbalanced heterologous production is also possible.^[279] Hazardous intermediates can dramatically reduce cell viability unless host cell responses compensate cellular toxicity.^[280-281] In this context, the metabolic background of the host may lead to the formation of byproducts from cascade intermediates by endogenous enzyme activities.^[42, 44, 141] Furthermore, bottlenecks can also be caused by a depletion of redox cofactors.^[11, 154-155]

Several optimization strategies have already been suggested earlier in this introduction and target different molecular levels. The use of balanced promoter systems (transcription)^[178] and codon-optimized genes (translation)^[282] is well-established.^[71] With the advent of systems metabolic engineering, combining metabolic and genetic engineering with systems and synthetic biology, limitations of *in vivo* applications have been addressed in more elaborate ways including, for example, the spatial organization of pathway enzymes along scaffold proteins.^[283] Recent advances aim at improved cofactor availability, the identification as well as the removal of competing background reactions.^[6, 36-37, 42]

C IV.1 Substrate uptake and product release

Whereas substrate and product transfer limitations are neglectable in dried or frozen whole cell biocatalysts, the permeability of the intact cell membranes might interfere with the diffusion of substrates and/or products. The outer membrane of bacteria is a lipid bilayer and only lipophilic substances can pass the membrane by diffusion. Consequently, the uptake of more polar substrate molecules might be limited. For unpolar compounds, the velocity of uptake can be limiting since diffusion is a passive and rather slow process.^[6]

Treatment of bacterial cells with solvents and/or detergents is a simple yet unspecific approach to improve the permeability of cell walls and commonly used in biocatalysis.^[284-285] Many protocols involve treatment of cells with ethylenediaminetetraacetic acid (EDTA) and toluene. One example is the bioconversion of ethyl-4-chlorooxobutanoate (ECOB) to ethyl-(*R*)-4-chloro-3-hydroxybutanoate (ECHB) by *E. coli* whole cells expressing the yeast reductase *YOL151W* and a GDH for cofactor recycling. ECOB could be fully converted to the target ECHB within 3 h reaction time. Importantly, no conversion was observed without pretreatment.^[6, 286] Although such protocols are straightforward, the adaption of incubation time and the addition of additives might be required. Unpredictable host cell responses in the presence of solvents and detergents or due the increased permeability of the cellular membrane represent drawbacks to be considered.^[287]

The introduction of transporter proteins provides an elegant solution to facilitate the uptake of specific substrates or a certain group of compounds. The coexpression of the alkane transporter *AlkL* from *Pseudomonas putida* (*P. putida*) overcame the rate limiting step of an artificial *in vivo* cascade leading to amino-functionalized fatty acid methyl esters (or the corresponding carboxylic acids; **Figure C-27**).^[288] The activity of the *AlkL* transporter improved yields up to 100-fold for the oxidation of C₁₂ and higher alkanes to the corresponding fatty alcohols and acids.^[6, 288] Generally, *AlkL* shows transport activity for hydrophobic (long chain) substrates.^[289] By implementation of an ω-TA, the amino-functionalized product was generated, whereas amounts of the

carboxylic acid byproduct were reduced. Importantly, the production of *AlkL* had to be tightly regulated since the overexpression of this transporter was cytotoxic.^[288]

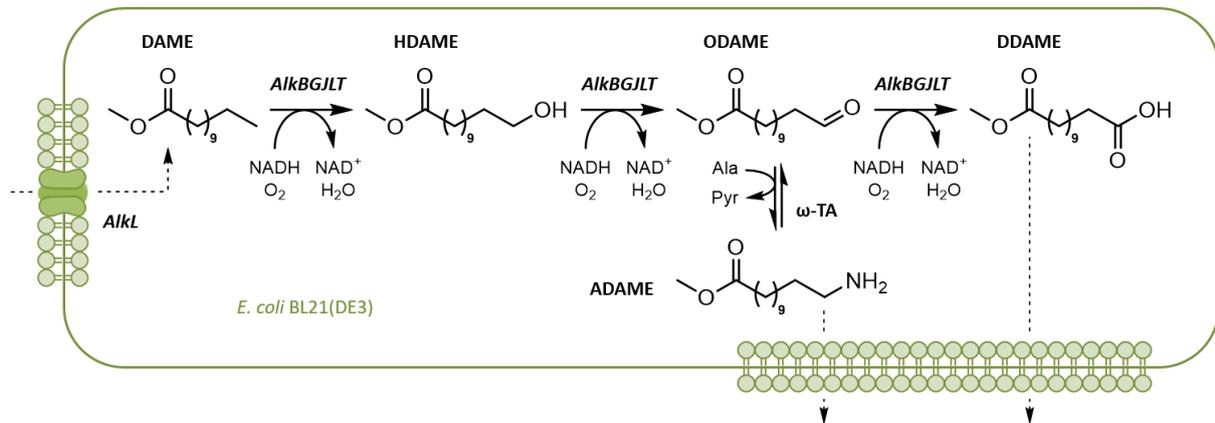


Figure C-27. The alkane transporter protein *AlkL* enhances target substrate uptake. Since the uptake of dodecanoic acid methyl ester (DAME) represented the rate-limiting step of the outlined *in vivo* cascade, the plug-in of *AlkL* overcame this limitation. DAME was oxidized via 12-hydroxy dodecanoic acid methyl ester (HDAME) to 12-oxododecanoic acid methyl ester (ODAME) by the *alkBGLT* operon from *P. putida*. The same operon-encoded enzymes lead to the formation of dodecanedioic acid methyl ester (DDAME). The implementation of an ω -TA produced 12-amino dodecanoic acid methyl ester (ADAME) with L-alanine (Ala) as the amine donor; Pyr = pyruvate. The figure was adapted from T. Bayer *et al.* (2015)^[6] and S. P. France *et al.* (2017).^[277]

Gram-negative bacteria natively express outer membrane proteins (OMPs), so-called porins, which enable the passage of hydrophilic solutes (e.g., nutrients) across the cell membrane and the secretion of waste products. Consequently, porins can be utilized to increase the influx of polarer substrates. Three known porins in *E. coli* are *OmpC*, *OmpF*, and *NmpC*, all trafficking various (non-natural) compounds including antibiotics.^[290-291] Environmental stimuli such as osmolarity, pH, temperature, or the concentration of nutrients regulate both the expression and the permeability of porins.^[292] *OmpF*, for example, is expressed at low temperatures. *PhoE*, another transporter protein, is exclusively expressed under phosphate starvation, which makes the expression of *PhoE* inducible^[291] and may be utilized as a regulatory tool for synthetic pathway design.^[6]

At present, the utilization of endogenous transporter proteins to regulate substrate uptake and product release is underrepresented in whole cell biocatalysis. Moreover, heterologous transport proteins, especially from higher organisms, can be difficult to express in microbial cells such as *E. coli*. The different composition of cell membranes potentially interferes with the native folding required for integral membrane function. Highly polar unusual target molecules (e.g., phosphorylated compounds) might have to be harvested by cell disruption. Alternatively, they can be converted by an additional enzymatic activity (e.g., dephosphorylation) into less polar products that can freely diffuse through the cellular membrane.^[293]

C IV.2 The intracellular environment and metabolic background

Historically, whole cells were exploited to accumulate natural metabolites. The fermentative production of ethanol by *S. cerevisiae* is only one prominent example.^[294] Organic acids of the tricarboxylic acid (TCA) cycle including citric and succinic acid are produced by different microorganisms on industrial scales. For example, *Aspergillus niger*, *Lactobacillus rhamnosus*, and *E. coli* are commonly employed to produce citrate, lactic acid, and succinate, respectively.^[295]

With increasing knowledge about the biochemistry and the architecture of metabolic networks of different organisms, productivities of designer cells have been improved by rational KO/KI strategies to remove unwanted

side reactions and reroute metabolic fluxes and to permanently insert target genes into the genome, respectively. Additionally, the adjustment of target protein production has been intensively used to enhance the flux through (synthetic) pathways (e.g., shikimate production in engineered *E. coli* cells).^[6, 296] Regarding genomic integration, Koma *et al.* observed differences in the activity of single copy insertions of reporter genes such as *lacZ*, which encodes β -galactosidase, depending on the insertion loci. Although the deleted genes were not essential, they might directly or indirectly influence the expression of other genes involved in protein production.^[111] As discussed earlier, the altered genetic context post genomic integration might also influence expression levels.^[6] Koma *et al.* combined λ Red recombination, *Flp*/FRT recombination, and P1 transduction^[297] to insert multiple genes into target loci on the *E. coli* chromosome. They successfully incorporated essential genes of the shikimate pathway to accumulate aromatic amino acids. By integrating two non-native decarboxylases from *Lactobacillus brevis* and *P. putida*, Koma *et al.* efficiently produced the aromatic compounds tyramine and phenethylamine.^[111]

A contrary strategy is the design of 'minimal genomes' and the generation of synthetic cells and strains. Such microbial strains will feature a strongly reduced metabolic background, consequently, will be less likely to interfere with *de novo* enzyme cascades. One way to build a minimal genome that includes only the genes essential for life was pursued by the team around Craig Venter in 2010.^[298] Parts of the genome of *Mycoplasma mycoides* (1 079 kb) were chemically synthesized and transplanted into cytoplasm, giving rise to the JCV-syn1.0 strain. Hutchison III *et al.* used a design, build, and test cycle to further reduce this genome to 531 kb encoding 473 genes. The resulting JCV-syn3.0 retained genes involved in key processes such as transcription and translation but also contained 149 genes of unknown function.^[299]

However, the biocatalytic application of 'minimal cells' as platform strains will probably not be realized soon and many *in vivo* cascades have not been optimized beyond laboratory scale. At present, low productivity still represents a major obstacle, which must be overcome to enable industrially relevant and profitable processes. Generally, there are two different strategies: a classical *in vivo* approach based on the increased metabolic network information of microorganisms and a more comprehensive *in silico* approach.

The *in vivo* methods specifically target side reactions to rationally knock-out genes. KOs aim at rerouting the flux through the (synthetic) pathway to accumulate target products and improve overall productivity.^[6] Whereas examples so far involved the KO of only one or a few genes, some strains have been immensely engineered to accumulate (natural) target compounds.^[300-302] Recently, Kunjapur *et al.* constructed an *E. coli* strain exhibiting reduced aromatic aldehyde reduction (RARE) activity by the deletion of up to six genes (*dkgA*, *dkgB*, *yeaE*, *yqhD*, *yahK*, and *yjgB*) with reported activity on the model substrate benzaldehyde.^[42] Other deletions targeted a transcriptional activator (*yqhC*), an endonuclease (*endA*), and a recombinase (*recA*). The latter two had been knocked-out to increase plasmide stability. The RARE strain was used to heterologously produce vanillin from vanillic acid by the expression of a CAR from *Norcadia iowensis* (*N. iowensis*; **Figure C-28A**). The vanillyl alcohol byproduct generated from the target vanillin by endogenous ADHs and aldo-keto reductases (AKRs) was not detected. To demonstrate that the RARE strain may serve as an aldehyde production platform strain, Kunjapur *et al.* synthesized the chiral pharmaceutical intermediate L-phenylacetyl carbinol (L-PAC) by C–C bond formation between benzaldehyde and glycolytic pyruvate (**Figure C-28B**) by recombinantly coexpressing *CAR_{Ni}* and a mutant PDC. Benzoic acid was reduced by *CAR_{Ni}* to benzaldehyde, which was not further reduced to the corresponding benzyl alcohol in the RARE strain. Compared to the unengineered *E. coli* strain, L-PAC production could be increased 10-fold.^[6, 42]

Like KO strategies, computational approaches depend on a comprehensive knowledge of biocatalytic reactions taking place in the target host microorganism. Whole cell biocatalysts are modelled as systems of a set of chemical reactions sets, resembling the metabolic pathways within the cell.^[6] Two complementary approaches

to predict and measure both the operation and regulation of metabolic networks are the genome-based flux balance analysis (FBA)^[303] and the steady-state isotopic labeling-based (¹³C) metabolic flux analysis (MFA).^[6, 304]

Genome-scale reconstructions of metabolic networks are available from online platforms such as BIGG^[305] or SEED.^[306] These models represent stoichiometric reaction equations of the metabolic reactions of the target microorganism and can be either analyzed with the freely available MATLAB COBRA toolbox or online tools (e.g., <http://www.theseed.org>).^[6, 307-308]

Systems biology progressed toward a full understanding and (quantitative) description of the metabolome^[309] of various microorganisms by computational models. Knowledge about the distinct metabolic flux distribution in production hosts has become essential for biotechnological applications with respect to productivity. Additionally, FBA is a valuable tool for the identification of lethal KOs prior to engineering whole cell biocatalysts. Three KOs in the central carbon metabolism of *E. coli* resulted in a >7-fold increase in the production of 3-hydroxy propionic acid, a precursor of the biodegradable poly-3-hydroxy propionic acid. The strain was constructed after gene KO simulations *in silico*.^[310]

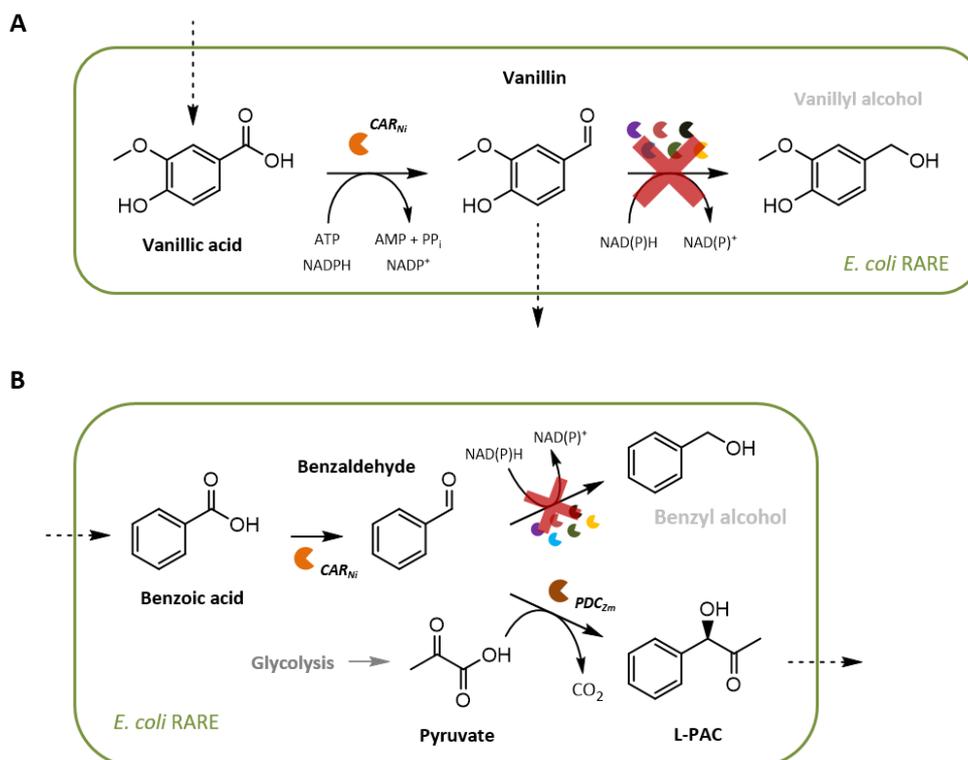


Figure C-28. The engineered metabolic background in the RARE strain enables the production of aromatic aldehydes *in vivo*. (A) Vanillin was produced by the reduction of vanillic acid by heterologously expressed CAR_{Ni} . The vanillyl alcohol byproduct formation was eliminated by the KO of six endogenous enzyme activities (red X). (B) The pharmaceutical intermediate L-PAC was produced in a two-step transformation from benzoic acid by CAR_{Ni} and subsequent C–C bond formation with glycolytic pyruvate by a pyruvate decarboxylase from *Zymomonas mobilis* (PDC_{Zm}). Reduction of the intermediate benzaldehyde to the corresponding benzyl alcohol was not observed in the RARE strain.^[42]

In a recent study, ¹³C-MFA was successfully employed to optimize the accumulation of lysine in *Corynebacterium glutamicum*. MFA attributed the enhanced production of lysine to the *de novo* NADPH generation pathway resulting from the alteration of the cofactor specificity of the native glyceraldehyde-3-phosphate dehydrogenase from NAD^+ (nicotinamide adenine dinucleotide) to NADP^+ and not to a previously hypothesized change in metabolic fluxes toward the pentose phosphate pathway (PPP).^[6, 311] The absolute quantification of metabolites

in host cells is an important tool as it provides information about maximal product yields and can identify potential metabolic bottlenecks (e.g., cofactor availability and recycling).^[312] In the context of synthetic pathway design, metabolomics was applied as a control strategy to follow the changes through intracellular fluxes in *E. coli* after the introduction of a synthetic pathway to produce isopropyl alcohol.^[313]

Different computational approaches enable the construction and optimization of whole cell biocatalysts *in silico* prior to their construction *in vivo*. KO candidates can be easily determined and verified in fewer experiments, consequently, reducing time, costs, and the overall workload, which is essential toward the industrial application of designer cells. Importantly, FBA and MFA can assist the identification of bottlenecks if artificial pathways connect to the metabolism of host cells by the utilization of (redox) cofactors (**Figure C-17**, **Figure C-27**, and **Figure C-28**) or hijacking metabolites such as acetyl-CoA (**Figure C-20**), pyruvate (**Figure C-5**), or dihydroxyacetone phosphate (DHAP; **Figure C-39**) to further improve product yields.^[6, 314]

C IV.3 Cofactor balancing and recycling

Since two decades, industrial biocatalytic processes have become increasingly common.^[30] Advances were driven by the development in the *omics* disciplines, genetic and metabolic engineering tools for pathway assembly and optimization, and protein engineering techniques to tune enzyme performance to meet industrial requirements. However, the large scale use of biocatalysts has long been limited to cofactor independent enzymes such as hydrolases.^[17, 30, 315] Cofactor dependent enzymes like oxidoreductases can perform demanding, synthetically useful reactions and have been used for the asymmetric reduction of carbonyl groups to alcohols and amines, for example.^[316]

Cofactors such as NAD(P)⁺/NAD(P)H and acetyl-CoA are low molecular weight compounds that are essential for many enzymatic reactions in host cells and usually recycled by the reactions of the central metabolism. Some cofactors including PLP, biotin, and flavins are tightly bound to their enzymes and self-regenerating in most cases. However, NAD(P)⁺/NAD(P)H cofactors and nucleoside triphosphates (NTPs) act more like cosubstrates. They are rather loosely bound and facilitate functional group transfer reactions and, hence, are consumed stoichiometrically.^[316-317] The introduction of cofactor dependent enzyme cascades can result in their depletion as reported by Milker *et al.* recently.^[155] The additional consumption of cofactors by *de novo* pathways inevitably affects cellular growth and probably the production of target compounds.^[6] Consequently, different strategies to balance or recycle cofactors have been employed to improve both cell viability and the productivity of (synthetic) pathways.^[318]

A simple approach to balance cofactor consumption and regeneration is the establishment of self-sustaining redox cascades (**Figure C-29**).^[319-320] In multistep transformations, a cofactor that is consumed in a previous step can be regenerated in a subsequent step dependent on the consumed form of the cofactor (**Figure C-29A**). Alternatively, an additional enzyme activity for cofactor recycling can be employed. The transformation of a cosubstrate into the corresponding byproduct depends on the exhausted cofactor, which is regenerated (**Figure C-29B**). Oberleitner *et al.* utilized both approaches to assemble a self-sustaining redox cascade *in vitro* containing an ADH from *L. kefir* (*ADH_{Lk}*), EREDs from either *S. cerevisiae* (*OYE1*) or *Pseudomonas sp.* (*XenB*), a CHMO from *Acinetobacter sp.* (*CHMO_{Acineto}*), and a D-glucose-6-phosphate dehydrogenase (G6P-DH) for NADPH regeneration.^[322] The applicability of the enzymatic cascade was exploited to synthesize structurally different ϵ -caprolactones as precursors for biorenewable polyesters (**Figure C-30**). Nonetheless, G6P is an expensive cosubstrate and the produced D-glucono-1,5-lactone-6-phosphate (G-1,5-L6P) cannot be reused. This is a drawback of many recycling approaches transforming cosubstrates into (unusable) byproducts, eventually causing complications in their separation during down-stream processing.

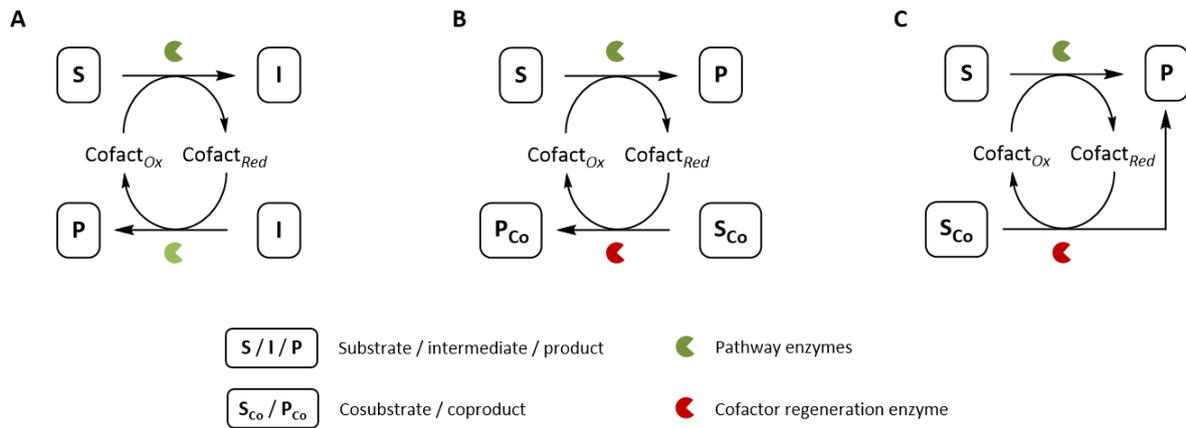


Figure C-29. Exemplary self-sustaining redox cascades. (A) Redox neutral two-step cascade regenerating cofactors through the intermediate step. This strategy is also known as ‘closed loop’ cascade. (B) Cofactor regeneration by an additional cofactor recycling system at the expense of a cosubstrate. (C) Redox neutral ‘convergent’ cascade.^[319-321]

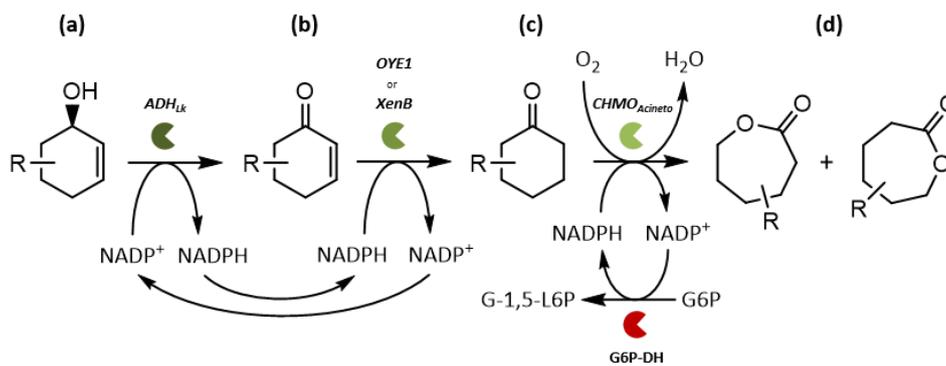


Figure C-30. Self-sustaining redox cascade to produce ϵ -caprolactones *in vitro*. Transformation of (*S*)-2-cyclohexen-1-ols (a) via 2-cyclohexenones (b) to cyclohexanones (c), in which NADP^+ is regenerated via the intermediate step. The subsequent oxidation of c to ϵ -caprolactones (d) by a BVMO depends on NADPH , which is regenerated by the oxidation of D-glucose-6-phosphate (G6P) to D-glucono-1,5-lactone-6-phosphate (G-1,5-L6P) catalyzed by G6P-DH.^[322]

To synthesize ϵ -caprolactone, Bornadel *et al.* designed a convergent two-enzyme cascade consisting of *CHMO*_{Acineto} for the oxidation of cyclohexanone and an ADH from *Thermoanaerobacter ethanolicus* (*ADH*_{Te}) for the oxidation of the ‘double-smart’ cosubstrate 1,6-hexanediol and the regeneration of NADPH (Figure C-29C and Figure C-31).^[321] Recently, a two-step optimization of the convergent cascade through design of experiments and a biphasic system was achieved.^[323]

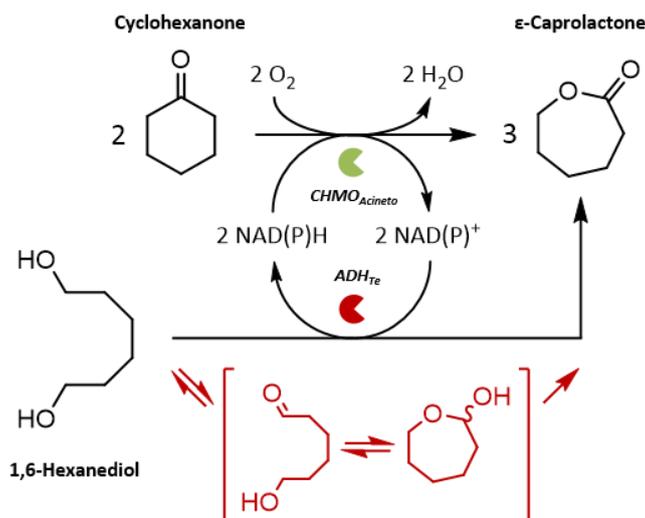


Figure C-31. Example of a 'convergent' cascade. The convergent cascade involved the oxidation of cyclohexanone by $CHMO_{Acineto}$ and a 'double-smart' cosubstrate (1,6-hexanediol), which is used for cofactor regeneration and transformed into the target product (ϵ -caprolactone) by ADH_{Te} without the formation of a byproduct. The figure was adapted from A. Bornadel *et al.* (2015).^[321]

Smart cosubstrates provide an elegant solution for the optimization of redox cascades with various advantages since the high excess of conventional cosubstrates (e.g., ethanol, isopropanol), which might negatively affect enzyme activities, can be omitted and the amounts of byproducts and waste are reduced (**Figure C-29C**).

In general, DH-catalyzed reactions are involved in biodegradation reactions and depend on $NAD^+/NADH$, whereas $NADP^+/NADPH$ facilitate biosynthetic reactions. Both nicotinamide cofactors play essential roles in biochemical redox reactions and enable the application of synthetically useful biocatalysts.^[13]

Several enzymatic strategies have been developed to regenerate $NADH$ such as the conversion of formate, glucose, or G6P by formate dehydrogenase (FDH), GDH, or G6P-DH, respectively; ADHs are routinely applied as well and utilize simple molecules (e.g., isopropanol) as cosubstrates.^[316] $NADPH$ is preferentially regenerated by GDH,^[316] G6P/G6P-DH,^[316, 322] engineered FDHs accepting $NADP^+$ instead of the natural NAD^+ ,^[324] and phosphite/phosphite dehydrogenase (PTDH).^[325-326] PTDHs catalyze the (almost) irreversible oxidation of phosphite to phosphate and were engineered toward a relaxed specificity for both $NADP^+$ and NAD^+ .^[316, 327] A thermostable PTDH mutant was also linked to several BVMOs to spatially converge the enzymatic transformation and the cofactor regeneration activities.^[328] However, phosphite cannot diffuse through cellular membranes and is, therefore, not applicable for cofactor regeneration *in vivo*.^[6]

Changing the cofactor specificity of enzymes by protein engineering techniques has been employed *in vitro* to develop industrial feasible processes. Recently, Beier *et al.* switched the cofactor specificity of $CHMO_{Acineto}$ from $NADPH$ to $NADH$. Many BVMOs display a strong preference for $NADPH$, which is less stable and ten times more expensive than $NADH$.^[329] The same strategy has been employed to optimize enzyme cascades *in vivo* and to increase the flux through target pathways.^[311]

In the context whole cell biocatalysis, it is essential to manipulate the intracellular redox state as well as cofactor levels.^[11, 330] Thus, various strategies have been applied to control the cofactor regeneration system or balance the enzyme activities of redox reactions. For example, $NADH$ oxidase (NOX) has been applied for cofactor regeneration by an engineered whole cell biocatalyst for chiral compound production.^[318] Intracellular cofactor recycling capacities and availability can be achieved by the KI of cofactor regenerating enzymes and/or the KO of competing metabolic reactions consuming the target cofactor.^[330]

Recently, the redox neutral amination of alcohols in *E. coli* was reported. It involved an ADH from *Geobacillus stearothermophilus* (*ADH-ht*) to oxidize primary aliphatic and aromatic alcohols to the corresponding aldehydes and an ω -TA from *Vibrio fluvialis* (*VflH6*), which aminated the intermediate aldehydes. An AlaDH from *B. subtilis* was coexpressed to recycle NADH (**Figure C-32**).^[331]

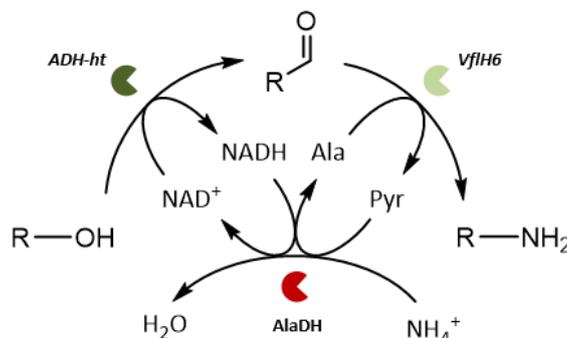


Figure C-32. Redox self-sufficient whole cell amination of primary alcohols. NAD⁺ dependent *ADH-ht* oxidized alcohol substrates to the corresponding aldehydes. Subsequent amination was catalyzed by *VflH6*. L-alanine (Ala) served as the amine donor. An AlaDH from *B. subtilis* simultaneously regenerated both Ala from Pyr and the NAD⁺ cofactor. The figure was adapted from S. Klatt and V. F. Wendisch (2014).^[331]

NAD⁺/NADH availability *in vivo* can be enhanced by the expression of well-known regeneration enzymes such as GDH and FDH.^[316] The GDH from *B. subtilis* has been broadly applied in biocatalysis^[141, 286] and engineered to adapt to industrial process conditions;^[16, 332] on the other hand, the FDH from *Candida boidinii* has been intensively used under aerobic and anaerobic growth conditions.^[316, 333] Both enzymes were simultaneously expressed to engineer the NADH regeneration capacity of *E. coli* to produce (2S,3S)-butanediol, a valuable building block in asymmetric synthesis.^[334] Zhou *et al.* employed NOX and further increased intracellular NAD⁺ levels by co-expressing a NAD⁺ transporter for the uptake of extracellularly supplied cofactor.^[335]

An example involving the KO/KI of genes is depicted by the engineering of *E. coli* to produce succinic acid by the fixation of CO₂ by a phosphoenolpyruvate (PEP) carboxylase. The KO of competing pathways led to accumulation of pyruvate and limited the regeneration of NAD⁺ from glycolytic NADH. To increase the regeneration of NAD⁺, a nicotinic acid phosphoribosyltransferase (NAPRTase) and a pyruvic acid carboxylase (PYC) were coexpressed. NAPRTase is the rate-limiting enzyme of NAD⁺/NADH synthesis in *E. coli* and PYC produces oxaloacetate from pyruvate.^[6, 336] The KIs led to a significant increase in both cell mass and succinic acid production, which was not possible without the NAPRTase and PYC activities since the glycolytic flux was blocked by the insufficient regeneration of NAD⁺.^[336-337]

For NADP⁺/NADPH recycling, alternative KO/KI strategies must be pursued as for NAD⁺/NADH regeneration, which are recycled differently *in vivo*.^[338] Many engineering approaches focus on enhancing the flux through the PPP.^[6] FBA assisted to the measure intracellular redox cofactor levels and suggested the cyclization of the PPP,^[339] which was confirmed by ¹³C-MFA.^[340] In case of regenerating NADPH under anaerobic conditions, the overexpression of PPP enzymes is inefficient. Hence, the activities of NADP⁺ transhydrogenase and ATP-dependent NAD⁺ kinase were modulated to produce NADPH.^[6] The combination of both strategies was applicable for the aerobic as well as the anaerobic regeneration of NADPH.^[341] Alternatively, the regeneration of cofactors can be accomplished through photochemically or electrochemically ways.^[342-344]

Apart from NAD(P)⁺/NAD(P)H cofactors, the regeneration of NTPs is of increasing importance, particularly adenosine-5'-triphosphate (ATP; **Figure C-33A**).^[316] ATP not only serves as a phosphorylating agent; it is a common cosubstrate for enzymes forming (high energy) bonds (e.g., acetyl-CoA; **Figure C-34**) in nature.

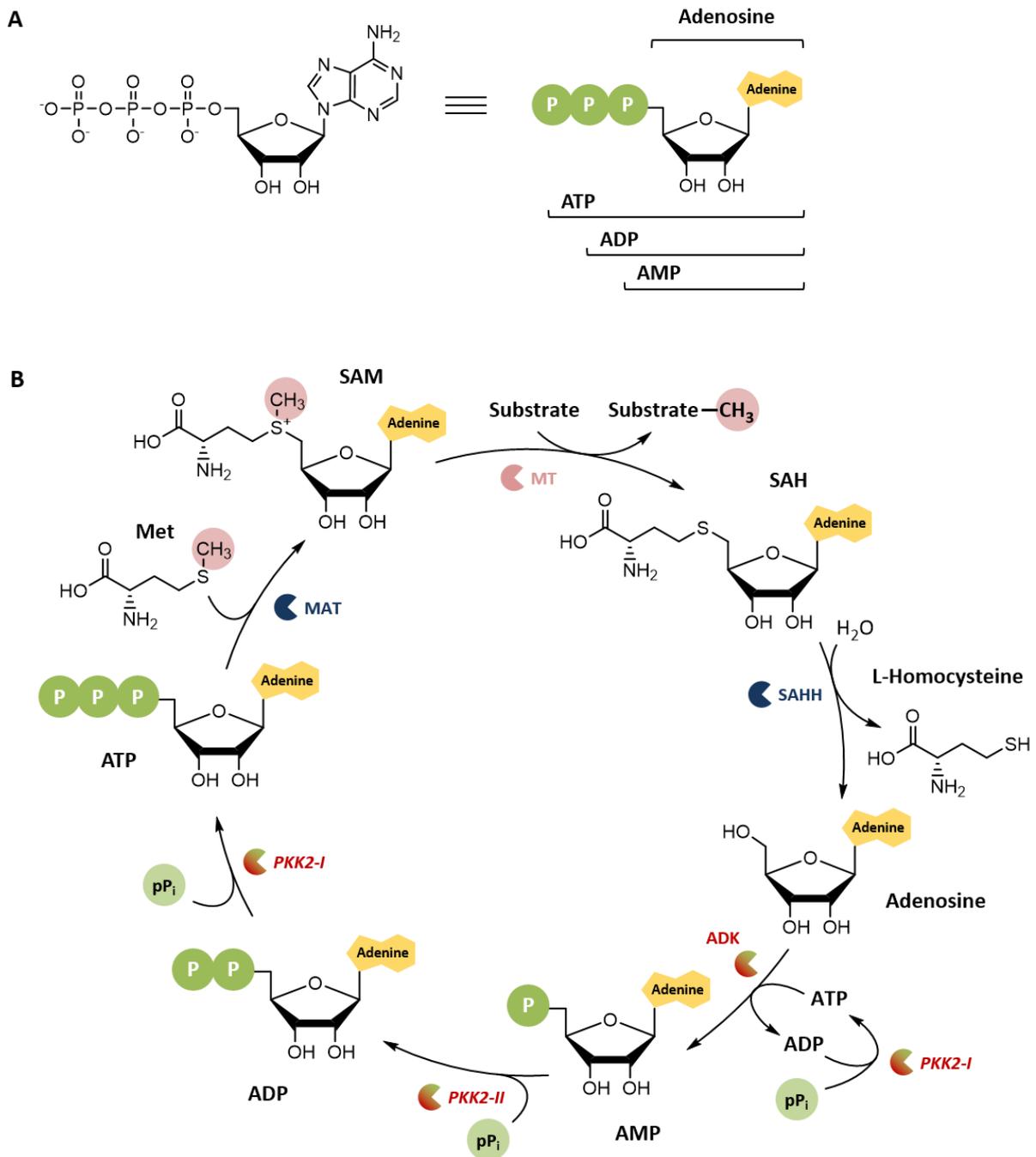


Figure C-33. SAM regeneration cycle. (A) Detailed and simplified structure of ATP. (B) The biocatalytic alkylation with cofactor regeneration is powered by polyphosphate (pP_i) and uses L-methionine (Met) or a derivative (not shown) as alkyl donor. A Met adenosyl transferase (MAT; *E. coli* or *Thermococcus kodakarensis*) forms SAM from Met and ATP. The target substrate is methylated by specific methyl transferases (MTs).^[346] S-adenosylhomocysteine (SAH) is hydrolyzed by SAH hydrolase (SAHH; *Mus musculus*). ATP is regenerated from adenosine and pP_i by adenosine kinase (ADK; *S. cerevisiae*), pP_i kinase (PKK2-II from *Acinetobacter johnsonii*; PKK2-I from *Sinorhizobium meliloti*). Met and the target MT substrate were added stoichiometrically, with an excess of pP_i and a catalytic amount of AMP as cofactor precursor. (B) was adapted from S. Mordhorst *et al.* (2017).^[346]

Consequently, the enzymes involved have an intrinsic application potential in the syntheses of value-added products. Several enzymatic strategies have been developed but do not match the ATP regeneration capabilities of whole cells. Although regeneration systems for ATP starting from adenosine-5'-diphosphate (ADP) are available, certain limitations exist for both *in vitro* and *in vivo* applications requiring ATP regeneration from

adenosine-5'-monophosphate (AMP) or adenosine (**Figure C-33**).^[345] Up to now, well-studied regeneration systems involved PEP/pyruvate kinase, acetylphosphate/acetate kinase, or polyphosphate (pP_i)/polyphosphate kinase (PPK), for example.^[316, 345] Most recently, Mordhorst *et al.* coupled different ATP regeneration strategies and a biomimetic cyclic S-adenosylmethionine (SAM) regeneration system to perform catalytic alkylation reactions (**Figure C-33B**).^[346]

SAM-dependent methyltransferases (MTs) are versatile tools for the specific alkylation of various compounds including pharmaceuticals. The cyclic SAM regeneration system assembled by Mordhorst *et al.* promises the robust and economical application of MTs soon. In addition to the substrate to be methylated, only methionine and pP_i must be added in stoichiometric amounts and the system acts catalytically with respect to the cofactor precursor adenosine in alkylation reactions.^[346]

Cofactor recycling and balancing has intensively applied to optimize redox cascades and other functional group transfer reactions. As mentioned earlier, C–C bond formations are pivotal reactions in nature and syntheses to expand the carbon framework to build up larger molecules.^[253] Whereas sugar derivatives can be assembled, for example, by DHAP-dependent aldolases (**Figure C-39**), fatty acids, polyketides, and steroids can be synthesized by enzymes requiring acetyl-CoA. Among the various enzymatic acetyl-CoA regeneration systems available,^[316] acetyl-CoA synthetase was successfully applied to synthesize P3HB *via* the biosynthetic pathway from *Ralstonia eutropha* from acetate *in vitro*.^[347] The condensation of acetate and CoA at the expense of ATP regenerated acetyl-CoA (**Figure C-34**). AMP could also be recycled from pP_i by PPK2-II and -I (**Figure C-33B**), for example, or recycled by whole cell biocatalysts heterologously expressing the P3HB pathway (see also **Figure C-12A**).

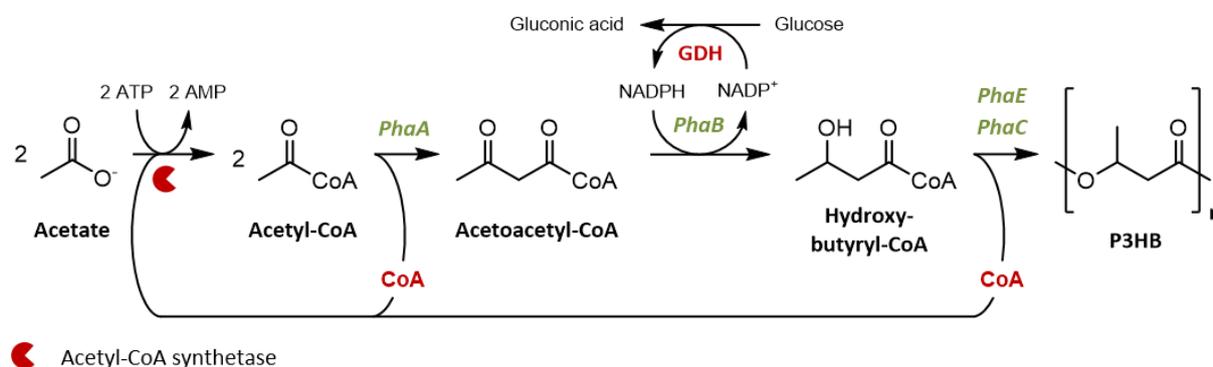


Figure C-34. Acetyl-CoA regeneration. The acetyl-CoA dependent production of P3HB involves two separate β -ketothiolase reactions (*PhaA* and *PhaC*). Acetyl-CoA is regenerated from acetate and CoA by acetyl-CoA synthetase. NADPH used in the reduction of acetoacetyl-CoA (*PhaB*) is regenerated by GDH.^[347] The figure was adapted from R. D. Woodyer *et al.* (2008).^[316]

In the past decade, substantial progress has been made to further improve cofactor/cosubstrate regeneration systems and to implement them *in vivo*. Recycling systems cover nicotinamide cofactors such as NAD(P)⁺/NAD(P)H, which are essential for many redox enzymes and their applications in syntheses, ATP/NTPs, and small molecules such as acetyl-CoA, involved in the expansion of carbon chains and frameworks *en route* to assemble complex chemicals. Advances in metabolic and protein engineering enabled the application of cofactor recycling systems on larger scales, including the production of pharmaceutical intermediates^[16] and precursors with industrial relevance.^[17, 323] Most recently, a regeneration system for the cofactor SAM was introduced.^[346] This will rapidly facilitate the development of more applicable, economical, and environmentally friendly uses of enzymes like MTs to expand the biocatalytic toolbox in the future.

C IV.4 Synthetic biology tools for carbon flux enhancement

The apparent advantages of *in vivo* biocatalysis include the efficient recycling of cofactors and the simultaneous production of target enzymes by host cells and are routinely applied. Control over the carbon flux through target cascades can be challenging. Strategies to enhance the flux such as the optimization of enzyme stoichiometry (**Figure C-19**) and the KO of competing endogenous enzyme activities (**Figure C-28**) have proven their applicability beyond doubt.^[6, 348]

These strategies evolved through millions of years of evolution to produce metabolic pathways and networks in cellular systems capable of one-pot multi-step catalysis and have been successfully transferred into test tubes and bioreactors. In nature, enzymes catalyzing sequential reactions have frequently evolved structures with stoichiometric arrangements of active sites and chemical and/or physical mechanisms, which control the mass transport through cellular systems and intermediates along metabolic pathways.^[349] This controlled flux and the direct transfer of chemical entities from one active site to another (without first diffusing to the bulk environment) is termed 'substrate channelling'.^[350-351] Substrate channelling promotes the processing of metabolites along a designated pathway; intermediates are directed from one active site to the next, thereby, reducing exposure to competing side reactions and protecting cells from toxic or labile intermediates. Additionally, high local concentrations of metabolites may overcome otherwise unfavorable thermodynamics in the bulk or cellular environment.^[349]

Intramolecular tunnels, electrostatic guidance, spatial organization (i.e., clustering), and covalent attachment of intermediates in multienzyme structures represent naturally evolved channelling mechanisms and are complemented by the colocalization of enzymes in microcompartments.^[349, 352-353] Consequently, substrate channelling and the underlying mechanisms inspired synthetic chemists and biologists to optimize enzymatic cascade reactions, where intermediates are not in equilibrium with the bulk solution, to direct diffusion processes, enhance the flux through target cascades, and increase yields (**Figure C-35**).

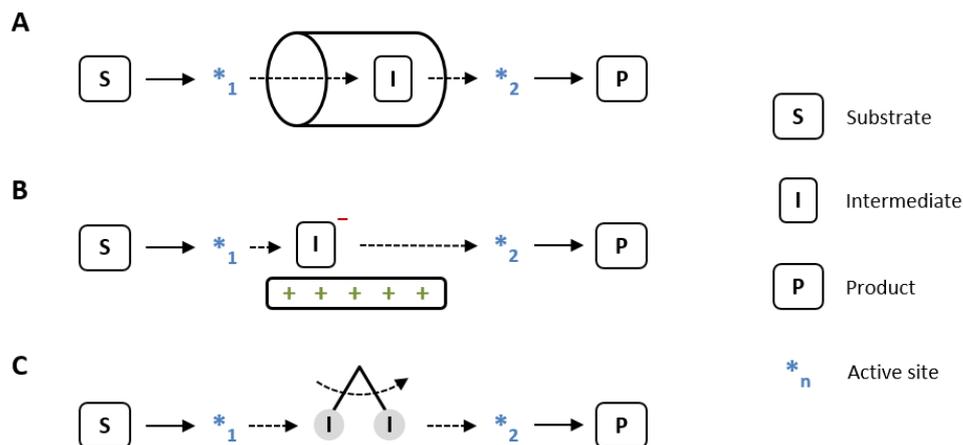


Figure C-35. Substrate channelling. (A) Intramolecular tunnels channel intermediate from one active site (blue) to the next (e.g., tryptophan synthetase). (B) Electrostatic guidance directs intermediates between active sites by interactions between positive (green) and negative (red) charges (e.g., TCA cycle metabolon). (C) Intermediates are covalently bound to chemical swing arms and transferred from one active site to another (e.g., PDC). The figure was adapted from I. Wheeldon *et al.* (2016).^[349]

One well-studied example of substrate channelling by an intramolecular tunnel is tryptophan synthase, a bifunctional enzyme catalyzing the conversion of indole-3-glycerolphosphate to indole and glyceraldehyde-3-phosphate at one active site and condenses indole with serine to tryptophan at the second active site (**Figure**

C-35A). A hydrophobic tunnel connects both active sites, facilitating the diffusion of indole along the tunnel.^[354-355] Intramolecular tunnels are also found in bifunctional aldolase/dehydrogenases and carbamoylphosphate synthase, channelling short chain aldehydes and carbamates, respectively.^[13, 349]

An alternative mechanism of substrate channelling is electrostatic guidance that is utilized, for example, by a complex of malate dehydrogenase and citrate synthase, two TCA cycle enzymes.^[13] Diffusion of the negatively charged cascade intermediate oxaloacetate is bounded by a positively charged patch on the protein surface between the active sites of the upstream malate dehydrogenase and the downstream citrate synthase (**Figure C-35B**).^[349, 356]

Regarding the covalent attachment of intermediates on chemical swing arms (**Figure C-35C**), there are various examples in nature. The PDC catalyzes the oxidative decarboxylation of pyruvate to acetyl-CoA, metabolically linking glycolysis to the TCA cycle.^[13] The complex is comprised of three enzymes: E1, E2, and E3. In *E. coli*, 24 subunits of E2 form a core with multiple E1 and E3 subunits around that comprise the surface of the multienzyme complex.^[357] E1 decarboxylates pyruvate and transfers the resulting acetyl group to the lipoamide swing arm on E2. Subsequently, E2 transfers the acetyl group from the chemical swing arm to free CoA. Finally, the swing arm is reactivated by NADH-dependent E3.^[13] Swing arms are also found in fatty acid synthases, polyketide synthases, and fatty acid β -oxidation complexes.^[358-360] In fatty acid and polyketide synthases, intermediates are attached to acyl carrier proteins *via* a thiol ester to a phosphopantetheine moiety. Acyl carrier proteins are channelled between enzymatic modules, where the intermediates are successively extended and modular assembly results in structural highly diverse products. Such assembly lines channel intermediates by protein/protein interactions, which are less dependent on the chemical properties of the target intermediates.^[349]

The phosphopantetheine swing arm in CARs shuttles carboxylic acid substrates from the adenylation domain to the reduction domain, where activated substrates are finally reduced to the corresponding aldehydes at the expense of NADPH (**Figure C-36**). PPTases posttranslationally attach the phosphopantetheinyl residue to the apo-CAR to form the functional holo-CAR enzymes.^[62-64]

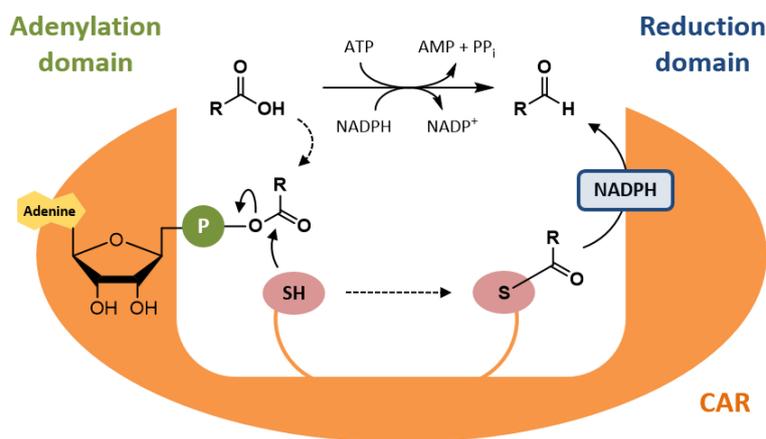


Figure C-36. Proposed mechanism of CAR enzymes. Carboxylate substrates and ATP enter the adenylation domain, where a phosphoester intermediate is formed and PP_i released. The thiol of the phosphopantetheine arm nucleophilically attacks the carbonyl carbon to form a thioester intermediate and AMP is released. The swing arm transfers the intermediate to the reduction domain, in which NADPH reduces the thioester bond. The phosphopantetheine thiol group is regenerated, the aldehyde products and $NADP^+$ released; reaction equation given at the top center. The figure was adapted from W. Finnigan *et al.* (2017).^[64]

The recent interest in the assembly of nanostructures with multiple organic and inorganic active sites that catalyze desired cascade reactions has been successfully adapted for the synthetic biology toolbox, which, at

present, features protein, nucleic acid, and polymer scaffolds to spatially organize non-natural cascades *in vitro* and *in vivo*.^[6, 361] Scaffolds enable the directed assembly and colocalization of multiple enzymes without the genetic and irreversible fusion of proteins.^[362]

Synthetic fusion proteins are designed by incorporating multiple proteins and domains into one complex to achieve improved properties and new functionalities, respectively, with a wide range of biotechnological and pharmaceutical applications.^[362] A prominent example was presented by the groups of Fraaije and Mihovilovic, who covalently fused a BVMO to a PTDH.^[328] One example is the construction of a synthetic bifunctional aldolase/kinase by gene fusion through overlap extension. The fusion enzyme consists of fructose-1,6-bisphosphate aldolase (*FruA*) from *Staphylococcus carnosus* (*S. carnosus*) and dihydroxyacetone kinase (*DhaK*) from *Citrobacter freundii* (*C. freundii*) with an intervening linker of five amino acid residues. The fusion protein retained both enzymatic activities and the proximity of the active centers promoted a kinetic advantage, indicated by a 20-fold increase in the initial velocity of the overall aldol reaction.^[363]

On the other hand, protein scaffolds were engineered and expressed in *E. coli* to colocalize a three-enzyme cascade producing mevalonate.^[283] Dueber *et al.* assembled the heterologous pathway consisting of the endogenous acetoacetyl-CoA transferase (*AtoB*), 3-hydroxyl-3-methylglutaryl-CoA synthase (*HmgS*), and 3-hydroxyl-3-methylglutaryl-CoA reductase (*HmgR*) from *S. cerevisiae* (see also **Figure C-20A**). The *de novo* pathway performance suffered from an unbalanced enzyme stoichiometry and the cytotoxicity of the intermediate 3-hydroxyl-3-methylglutaryl-CoA (HMG-CoA).^[364] To overcome these bottlenecks, a synthetic protein scaffold was constructed from metazoan binding domains: a GTPase binding domain (GBD) from rat, the Src homology 3 (SH3) domain, and PSD-95/DlgA/ZO-1 (PDZ) from mouse.^[283] *AtoB*, *HmgS*, and *HmgR* were tagged with the corresponding peptide ligands, respectively (**Figure C-37**).

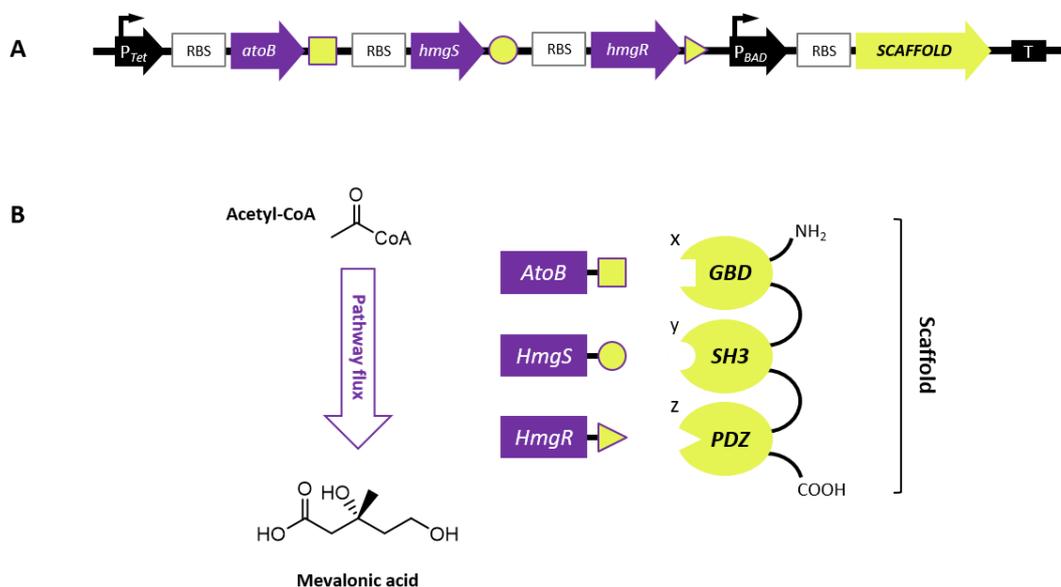


Figure C-37. Modular control over metabolic fluxes by synthetic protein scaffolds. (A) Pseudo-operon encoding the independently inducible mevalonate pathway genes (*atoB*, *hmgS*, and *hmgR*) and the synthetic scaffold. (B) The biosynthetic enzymes for mevalonate production *AtoB*, *HmgS*, and *HmgR* were tagged with the peptide ligands to bind to the protein scaffold *via* their target binding domains GBD, SH3, and PDZ, respectively. The number of domain repeats *x*, *y*, and *z* controls enzyme stoichiometry to maximize mevalonate titers. The figure was adapted from J. E. Dueber *et al.* (2009).^[283]

In comparison to the unoptimized cascade, the spatially organized mevalonate pathway achieved a 77-fold improvement in product titer. The natural modularity of GBD, SH3, and PDZ domains enabled the recruitment of

the three mevalonate biosynthetic enzymes to a synthetic complex and balancing of enzyme activities. Consequently, the low target enzyme expression reduced the metabolic load, channelling of the toxic HMG-CoA improved cell viability, and the overall flux through the artificial pathway could be greatly increased.^[6, 283]

Furthermore, nucleic acid scaffolds have been used to assemble two-step cascades leading to *trans*-resveratrol, propane-1,2-diol, and hydrogen;^[365-366] DNA aptamers have been used to colocalize enzyme cascades *in vitro*;^[367-369] and polymersomes and microcapsules have been used to assemble two- and three-step enzyme cascades.^[370-371]

So far, many colocalization and compartmentalization systems have been developed for *in vivo* applications with the goal of increasing the efficiency of metabolic pathways. Each of the two strategies, colocalization and compartmentalization, has its challenges: scaffolds ideally should allow control over the recruitment and proximity of the target cascade enzymes, while encapsulation of enzymes in compartments should likewise be programmable. Currently, colocalization onto protein or DNA-scaffolds exclusively relies on reversible interactions, which could lead to disassembly *in vivo* and diminish robustness of applications *in vitro*. Additionally, the low number of binding domains (e.g., GBD, SH3, or PDZ) available limits the number of enzymes that can be colocalized in a controlled fashion.^[372-374]

Bacterial protein microcompartments and smaller capsids are currently under investigation as alternatives to (synthetic) scaffolds.^[372] So far, virus capsids have been engineered to encapsulate enzymatic cascade reactions.^[375-376] In contrast to viral capsids, bacterial microcompartments are larger and have evolved to compartmentalize metabolic pathways including the ethanolamine degradation pathway in *Salmonella* sp. or ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) in cyanobacteria. Encapsulation peptides target metabolic enzymes into the self-assembling compartments;^[377-379] to connect the metabolically active lumen of the compartment and the bacterial cytoplasm, pores are formed by shell proteins.^[372] Such compartments can be heterologously produced and non-native cargo proteins targeted into the lumen.^[377, 380-382] In the context of synthetic cascade optimization and biotechnological application, compartments may be engineered as robust nano-scale 'bioreactors' *in vitro* with functionalized shell surfaces for immobilization and pore selectivities tailored for specific cascade reactions. Another area of vivid research is the encapsulation of enzymes into synthetic vesicles and, more recently, polymersomes have been used to entrap complex multi-enzyme systems, for example, the transcription/translation machineries of cells and other metabolic cascades.^[372]

Advances in systems biology and biological engineering have provided elegant solutions and tools for the industrialization of (synthetic) biology.^[6] The colocalization and assembly of enzymes into controlled aggregates and organized clusters is a recently developed strategy to enhance the overall performance of enzyme cascades, leading to enhanced selectivities and yields. The goal of these efforts is to construct artificial nano-bioreactors *in vitro* or cellular compartments *in vivo*. Both can function as chemical microenvironments that are equipped with the ability to produce their own enzymes, facilitate transport across their membrane, and, potentially, communicate with other vesicles performing the same or different tasks.^[370, 372, 383-385]

C V State-of-the-art microbial cell factories

With the genetic strategies to assemble artificial metabolic pathways and to introduce them in bacterial hosts like *E. coli*, biocatalytic systems have matured to produce chemicals that could not be easily made by known (chemical) processes. With the advent of systems biology from 2000 onwards, the description of complex biological systems led to genetic engineering tools that resolved bottlenecks impairing synthetic enzyme cascades in the cellular context due to the metabolic background of host cells.^[6, 17, 386-388]

These combined strategies resulted in the construction of ‘microbial cell factories’ and optimization approaches aimed at balancing enzyme stoichiometry by the implementation of different regulatory elements (**Figure C-18** and **Figure C-19**) and reduced endogenous host enzyme activities through targeted gene KO, for example, preventing the formation of byproducts from cascade intermediates (**Figure C-17** and **Figure C-28**).^[6, 389-390] Enhanced fluxes through synthetic cascades were achieved by substrate channeling (e.g., scaffolding; **Figure C-37**) to ultimately increase the overall pathway performance.^[6, 283, 372] Besides the adaption of cellular properties including its genetics and cell physiology (e.g., membrane permeability^[286] and substrate uptake,^[288, 391] **Figure C-27**) and catalyst design by protein engineering techniques, reaction engineering was intensively examined to improve both selectivities and yields.^[17, 392-394] The adaption of reaction conditions can include the reaction medium (e.g., buffer solutions), temperature, pH, or the use of biphasic systems.^[6, 393-394] The previously introduced concept of biocatalytic retrosynthesis supports the design of *de novo* enzyme cascades and the identification of appropriate biocatalysts from a rapidly expanding portfolio.^[68] The use of whole cell catalysts has been shown as more resource efficient with a reduced environmental impact for the efficient production of natural compounds, fine chemicals, and complex molecules including pharmaceuticals.^[6, 34, 395]

C V.1 The synthesis of aldehydes as products and valuable intermediates

Aldehydes are a class of chemicals with various industrial applications such as flavors and fragrances.^[396] Importantly, the high reactivity of the carbonyl group makes aldehydes valuable intermediates and precursors, not only for pharmaceuticals.^[280, 397-398] Regarding the high value applications and large markets for several aldehydes (e.g., benzaldehyde, cinnamaldehyde, and vanillin), focus has been put on microbial aldehyde synthesis in recent years.^[399] However, aldehydes are not known to accumulate in most natural microorganisms since their high reactivity contributes to cellular toxicity.^[280-281, 400] Initially, target aldehyde products had to be isolated from bacterial hosts by evaporation or extraction *in situ*. Alternatively, efforts have been made to engineer *E. coli* and other microbes with emphasis on *de novo* aldehyde biosynthetic routes and the accumulation of target aldehyde compounds, contemplating cellular toxicity.^[280]

Because most industrial host strains do not accumulate aldehydes, microbial production of these molecules from simple carbon sources required two parallel approaches: pathway construction for product synthesis and strain engineering for product accumulation. Pathway constructions considered enzymatic reactions that can produce desired aldehydes including the oxidation of primary alcohols and the reduction of carboxylic acids, for example. Carboxylic acids are ubiquitous in cellular metabolism and many can be converted to aldehydes in whole cells, usually involving only a single enzymatic activity.^[401-403] Roughly a decade ago, significant advance occurred by the identification and characterization of a CAR from *N. iowensis* (*CAR_{Ni}*). After cloning into *E. coli*, soluble expression, and purification, *CAR_{Ni}* was active on several aromatic carboxylic acids *in vitro*.^[404] The unfavored conversion from carboxylates to aldehydes requires the activation of substrates by the hydrolysis of ATP to AMP and pyrophosphate (PP_i). The activated intermediates are shifted intraenzymatically by a phosphopantetheine residue to be reduced in the presence of NADPH (**Figure C-36**).^[405] Since then, various CARs from organisms such as *Mycobacterium* sp. or *Streptomyces* sp. have been added to the portfolio of available biocatalysts.^[62-64]

One major challenge of aldehyde production *in vivo* has long been the rapid endogenous conversion of nearly all aldehydes to their corresponding alcohols.^[42, 280, 404] In 2012, Rodriguez and Atsumi reported the accumulation of isobutyraldehyde in *E. coli* by sequentially deleting eight genes (*yqhD*, *adhP*, *yiaY*, *ahr*, *betA*, *fucO*, *eutE*, and *eutG*).^[390] In 2014, they constructed a platform *E. coli* strain (with 13 gene KOs) that minimally converted exogenously supplied aliphatic aldehydes to the corresponding alcohols.^[406] Kunjapur *et al.* constructed an *E. coli* strain displaying reduced aromatic aldehyde reduction (RARE) activity by deleting up to six genes (*dkgA*, *dkgB*, *yeaE*, *yahK*, *ahr*, and *yqhD*) with reported activity on their model substrate benzaldehyde. They applied the RARE

strain to heterologously produce vanillin and L-PAC upon expression of *CAR_{Ni}* (**Figure C-28**). Both engineered strains provide versatile microbial cell factories to produce structurally different aliphatic and aromatic aldehydes, respectively.^[42, 406]

Whereas the reduction of carboxylates is now readily established on laboratory scale and non-oxidative biosynthetic routes exist that decarboxylate 2-keto acid substrates (e.g., PDC, *KivD*), the oxidation of primary (aromatic) alcohols to the corresponding aldehydes remains a major challenge in biocatalysis with only a few examples in the literature.^[280, 407-409] One issue toward the production of aldehydes is the overoxidation to the corresponding carboxylic acids. In this regard, whole cells (e.g., *Glucanobacter* sp. or *Brevibacterium* sp.) are commonly used to carry out the two-step oxidation from primary alcohols to carboxylates. Such systems exploit endogenous activities of ADHs and aldehyde dehydrogenases (ALDHs); the use of isolated enzymes is less common for these oxidations.^[407, 410] However, aldehydes as endproducts only could be accessed by *in situ* extraction, employing two-liquid-phase systems to avoid overoxidation,^[289, 411] or subsequent reaction of the aldehyde with a 'scavenger' molecule.^[412] The group of Bühler presented an efficient whole cell biocatalyst for the oxidation of hydrophobic long-chain substrates (e.g., DAME).^[289] The heterologous pathway contains the alkane monooxygenase *AlkBGT* and the ADH *AlkJ* from *P. putida*. As mentioned earlier, the OMP *AlkL* was introduced to overcome the limited substrate uptake and increased the carbon flux through the cascade. To isolate target aldehydes, for example, an organic solvent served as a substrate reservoir and extracted target products *in situ*. Best results were achieved with an organic phase composition of 75% (v/v) bis-(2-ethylhexyl) phthalate and 25% (v/v) DAME as the substrate. The heterologous expression of *AlkJ* shifted the reaction equilibrium from the primary alcohol HDAME to the thermodynamically unfavored aldehyde moiety, yielding ODAME, which was extracted *in situ*^[391] or subsequently transformed into DDAME (**Figure C-27**).^[6, 389]

From the perspective of *de novo* aldehyde synthesis, biocatalytic oxidations may provide new opportunities for primary alcohols as valuable starting materials.^[280] In general, ADHs are more often applied than oxidases since the biocatalyst portfolio lacks native enzymes from the latter class that accept a broad variety of substrates.^[407] Horse liver dehydrogenase, on the other hand, has been widely applied *in vitro* and shows a broad substrate specificity for primary alcohols.^[413-414] Recently, the group of Li introduced synthetic pathways for the functionalization of styrenes in *E. coli* (**Figure C-4**). Pathway assembly and product flexibility was achieved by the combination of enzyme modules encoded by four compatible pET vectors in one cell. Target products included α -hydroxy carboxylic acids, 1,2-amino alcohols, and α -amino acids and were all synthesized *via* aldehydes from their primary alcohol intermediates.^[43]

The minimization of the endogenous reduction of aldehydes in model organisms such as *E. coli* has provided a foundation for microbial aldehyde synthesis and broader utilization of aldehydes as intermediates for other biosynthetically challenging compound classes including primary amines,^[418-419] esters,^[420] and chiral building blocks such as cyanohydrins^[421] or polyfunctional carbonylation products (**Figure C-38**).^[42, 280, 397-398, 422-424] Especially aldolases have been used to enzymatically elongate carbon chains from aldehyde substrates. As pointed out earlier, protein engineering techniques have been applied to enhance the applicability of aldolase mutants as powerful biocatalysts in organic synthesis (**Figure C-25**).^[256, 260]

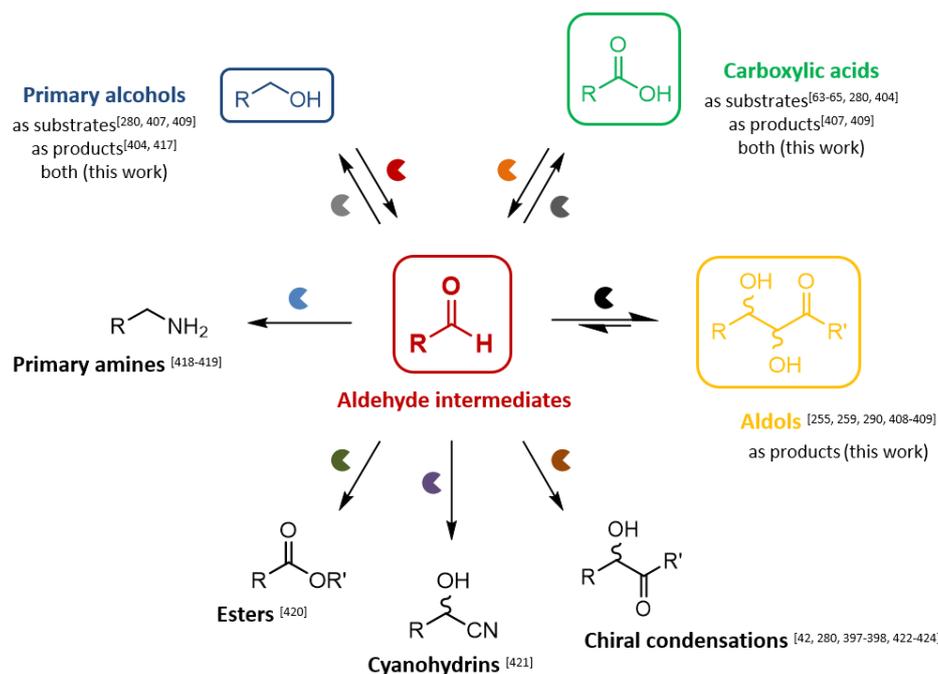


Figure C-38. Enzymatic aldehyde-derived products and building blocks. Aldehydes (red) can be produced from primary alcohols (blue) and carboxylic acids (green) and are valuable intermediates toward the synthesis of primary amines, esters, chiral cyanohydrins, and carbonylation products. Like carbonylation products, aldols (yellow) are synthesized by the formation of a new C–C bond, resulting in two new chiral centers.^[256, 260, 293, 415-416] The metabolic host background of *E. coli*, for example, can be utilized to produce alcohols^[417] or carboxylates.^[348] Structures in frames are the focus of this work. The figure was adapted from A. M. Kunjapur and K. L. J. Prather (2015).^[280]

C V.2 Carbon framework expansion by (engineered) aldolases

The stereoselective formation of C–C bonds is essential for organic synthetic chemistry to build up and expand carbon frameworks and increase complexity of target compounds.^[425-427] In this regard, biocatalysts with their high chemo-, regio-, and stereoselectivity offer major advantages to chemical catalysts and have accelerated research on enzymes for C–C coupling reactions.^[428] In living systems, C–C bonds are usually formed from activated precursors that are joined stereoselectively by highly specific enzymes, which play vital roles in the biosynthesis of carbohydrates. Enzymes involved in these biosynthetic pathways are ketolases/transketolases and aldolases/transaldolases. Naturally, ketolases and aldolases cleave C–C bonds rather than forming them.^[415, 428] From a synthetic point of view, aldolases have been widely applied as biocatalysts to expand carbon frameworks. Not only up to two stereogenic centers can be formed upon C–C coupling from simple, unprotected substrates; aldolase tandem reactions produce (cyclized) carbohydrates, simultaneously introducing as many as six chiral centers (**Figure C-25**).^[260] The use of aldolases has greatly improved the synthetic opportunities for the atom-economic asymmetric synthesis of chiral molecules with potential pharmaceutical relevance.^[260] New aldolases derived from the transaldolase scaffold (based on fructose-6-phosphate aldolase from *E. coli*) have been shown to be unusually flexible in their aldol acceptor scope; this makes them particularly valuable for addressing an expanded molecular range of complex polyfunctional targets.^[254]

Aldolases catalyze C–C coupling between an aldehyde acceptor and another carbonyl compound (e.g., ketone or aldehyde) that can form an enol and acts as the aldol donor. Aldolase-mediated reactions yield polyhydroxylated chiral compounds that are of special interest for the pharmaceutical industry since the polar hydroxyl functionalities aid water solubility and hydrogen bond interactions may convey biological specificity.^[260, 415] Depending on the nature of the reaction, aldolases belong to either the lyase enzyme class [EC 4] or the

transferases [EC 2]. They usually show relaxed specificities regarding the electrophilic group (acceptor), but strictly depend on a specific nucleophile (donor). Aldolases can be grouped by their reaction mechanisms into type I and II aldolases. Type I aldolases form a protonated Schiff base intermediate, linking a highly conserved lysine in the active site with the donor carbonyl carbon (e.g., DHAP). Additionally, tyrosine residues are crucial since they act as stabilizing hydrogen acceptors. On the other hand, class II aldolases polarize the carbonyl group with a divalent cation such as Zn^{2+} for subsequent C–C bond formation.^[428-429]

Amongst the carbohydrate aldolases, the commercially available fructose 1,6-bisphosphate aldolase (*FruA*) from rabbit muscle (RAMA) has been most widely employed in preparative synthesis.^[430] This class I aldolase reversibly catalyzes the formation of D-fructose 1,6-bisphosphate (F-1,6-bisP) from DHAP and D-glyceraldehyde 3-phosphate (GAP). RAMA and *FruA* aldolases from other organisms display a broad tolerance for many other aldehydes and sugar analogs and yield (3*S*,4*R*) products (**Figure C-39**).^[415] Other known DHAP-dependent aldolases fucose 1-phosphate aldolase (*FucA*), rhamnose 1-phosphate aldolase (*RhuA*), and tagatose 1,6-bisphosphate aldolase (*TagA*) provide the complementary product configurations (3*R*,4*R*), (3*R*,4*S*), and (3*S*,4*S*), respectively (**Figure C-39**).^[293, 415]

Another DHAP-dependent type I aldolase was first described by Schürmann and Spenger.^[431] The D-fructose 6-phosphate aldolase *Fsa1* from *E. coli* natively links dihydroxyacetone (DHA) and DHAP. Within recent studies by the groups of Clapés and Fessner, *Fsa1* was evolved by protein engineering tools and the donor preference was impressively expanded. Both groups modified the specificity toward hydroxyacetone (HA) and DHA,^[258, 432] hydroxyethanal,^[259] and various, more sterically demanding nucleophiles.^[257, 432-433] Like other aldolases, *Fsa1* accepts a vast number of different aldehyde acceptors,^[434-435] which makes *Fsa1* a versatile biocatalyst to produce iminocyclitols, a compound class of attractive drug candidates,^[436] and other biologically relevant substances.^[416]

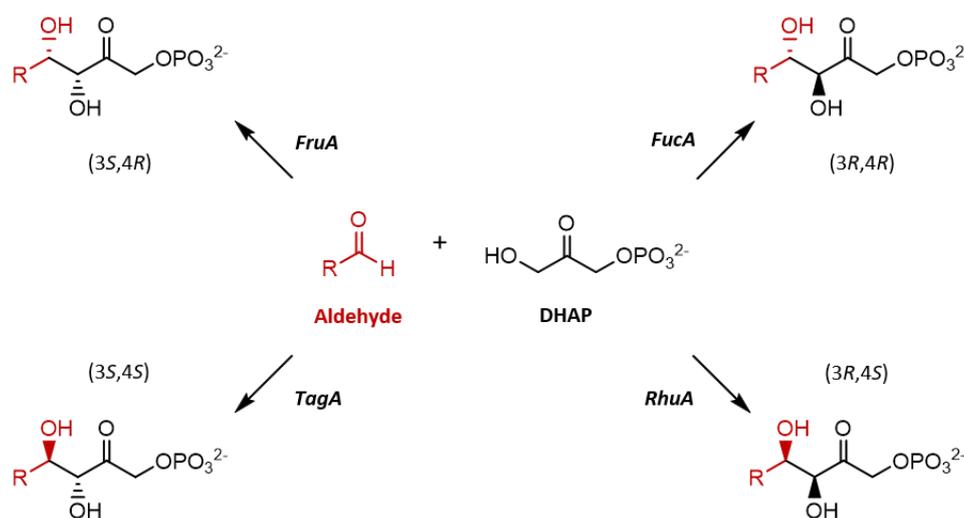


Figure C-39. The four complementary DHAP-dependent aldolases. DHAP-dependent aldolases mediate the aldol condensation of an aldehyde acceptor (red) with the natural donor DHAP (black). *FruA*, *FucA*, *RhuA*, and *TagA* provide the full set of four aldol products. The figure was adapted from M. Wei *et al.* (2015).^[293]

As outlined before, protein engineering by directed evolution and rational design has altered properties such as enzyme stability,^[437] substrate specificity,^[438-439] and stereoselectivity to produce tailor-made aldolase biocatalysts.^[416, 440-441] Computational design of *de novo* aldolases from inactive protein scaffold backbones and designed active sites has also been successfully demonstrated. Initially low activities were increased through site-directed mutagenesis and laboratory evolution and resulted in the RA95 aldolase mutant.^[442] Recently, the group

of Hilvert applied an ultrahigh-throughput, droplet-based microfluid screening platform to further improve the activity of the artificial *RA95* aldolase 30-fold.^[443]

Although manipulation of aldolases led to new biocatalysts and their application in cascade reactions, they are predominantly employed in transformations *in vitro*.^[260, 415-416, 444] H elaine *et al.* set up a multienzymatic procedure consisting of two kinases and *Fsa1* in one-pot to obtain terminally phosphorylated, rare L-sugars (**Figure C-40**).^[445]

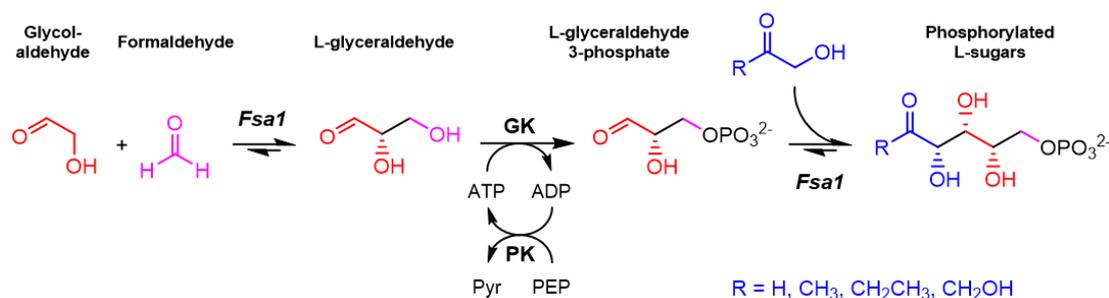


Figure C-40. *In vitro* cascade leading to phosphorylated L-sugars. The three-step cascade involves *Fsa1* to link the achiral compounds glycolaldehyde (red) and formaldehyde (pink) in the first step and another donor molecule (blue) in the third step. The intermediate step employs a glycerol kinase (GK) for phosphorylation of L-glyceraldehyde to L-glyceraldehyde 3-phosphate and a pyruvate kinase (PK) to regenerate ATP from PEP; Pyr = pyruvate.^[316] The figure was adapted from V. H elaine *et al.* (2015).^[445]

By utilizing an enzymatic route to produce DHAP as the acceptor for subsequent aldolase reaction, H elaine and co-workers circumvented the addition of notoriously labile DHAP. However, their cascade was still dependent on an excess PEP, which, like many other phosphorylated compounds, is labile and expensive. This drastically impairs the synthetic applicability of these chemicals. Besides often low yielding chemical syntheses,^[446] alternative enzymatic routes to produce DHAP *in situ* include (i) the phosphorylation of glycerol by a glycerol kinase (GK) and the subsequent oxidation of GAP by a GAP oxidase and (ii) the ATP-dependent phosphorylation of DHA by the *DhaK* from *C. freundii*, for example.^[293, 363, 447]

The oxidation of GAP by an oxidase generates hydrogen peroxide (H₂O₂), which must be mitigated by a catalase,^[448] especially *in vivo*.^[6, 400] Even though (i) and (ii) start from the cheap substrates glycerol and DHA, respectively, they employ multiple, costly isolated enzymes to produce DHAP. This adds complexity to the enzymatic procedure and renders such *in vitro* approaches inefficient.^[6] A desirable solution to eliminate these bottlenecks was realized by the group of Wang. They transformed DHAP-dependent aldolase-mediated *in vitro* reactions into an engineered *E. coli* and solved the issue of DHAP availability and stability by hijacking the glycolytic intermediate from the central carbon metabolism of the host cell.^[293] In living cells, glucose is metabolized into (F-1,6-bisP) via three consecutive enzymatic steps. *FruA* splits F-1,6-bisP into DHAP and GAP, which are interconverted by triosephosphate isomerase (TMI) with a highly favored formation of DHAP.^[13, 449-450] The recombinant expression of *FruA*, *FucA*, and *RhuA* omits tedious enzyme isolation and purification. The coexpression of the *E. coli* phosphatase *YqaB* irreversibly dephosphorylates aldol adducts *in vivo*, thus, achieving phosphate recycling by the host cell and shifting the reaction equilibrium toward the nonphosphorylated target aldols (**Figure C-41**).^[293]

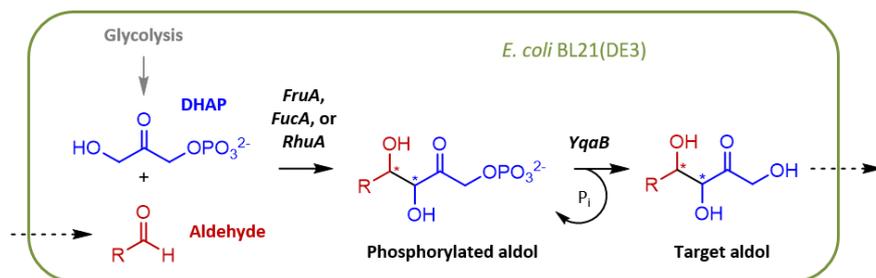


Figure C-41. Aldolase cascade hijacking metabolic DHAP to produce polyhydroxylated compounds *in vivo*. The engineered *E. coli* cells harbored one of the DHAP-dependent aldolases *FruA*, *FucA*, or *RhuA*, which linked substrate aldehyde acceptors (red) to the intracellular donor DHAP (blue), which was produced from glucose *via* glycolysis. Subsequent dephosphorylation by *YqaB* led to the formation of target aldols. The figure was adapted from M. Wei *et al.* (2015).^[293]

Such carbohydrate derivatives play essential roles in biological functions, including cell–cell interactions, signaling processes during bacterial and viral infection, inflammation, and cancer metastasis. In this regard, polyhydroxylated chiral compounds are precursor molecules for potential glycosyltransferase inhibitors.^[451-452] Consequently, they are valuable targets for the pharmaceutical industry and medical research.^[260, 293, 453]

C VI Scope and aim of this thesis

Synthetic enzyme cascades are constructed by sequentially combining biocatalytic reactions that are metabolically nonrelated in nature. This work blends in with this rapidly developing field, which has led to the production of complex valuable chemicals from simple precursors. Aim of this thesis is the design, assembly and implementation, characterization, and optimization of artificial enzymatic pathways in the model organism *E. coli* (Figure C-42).

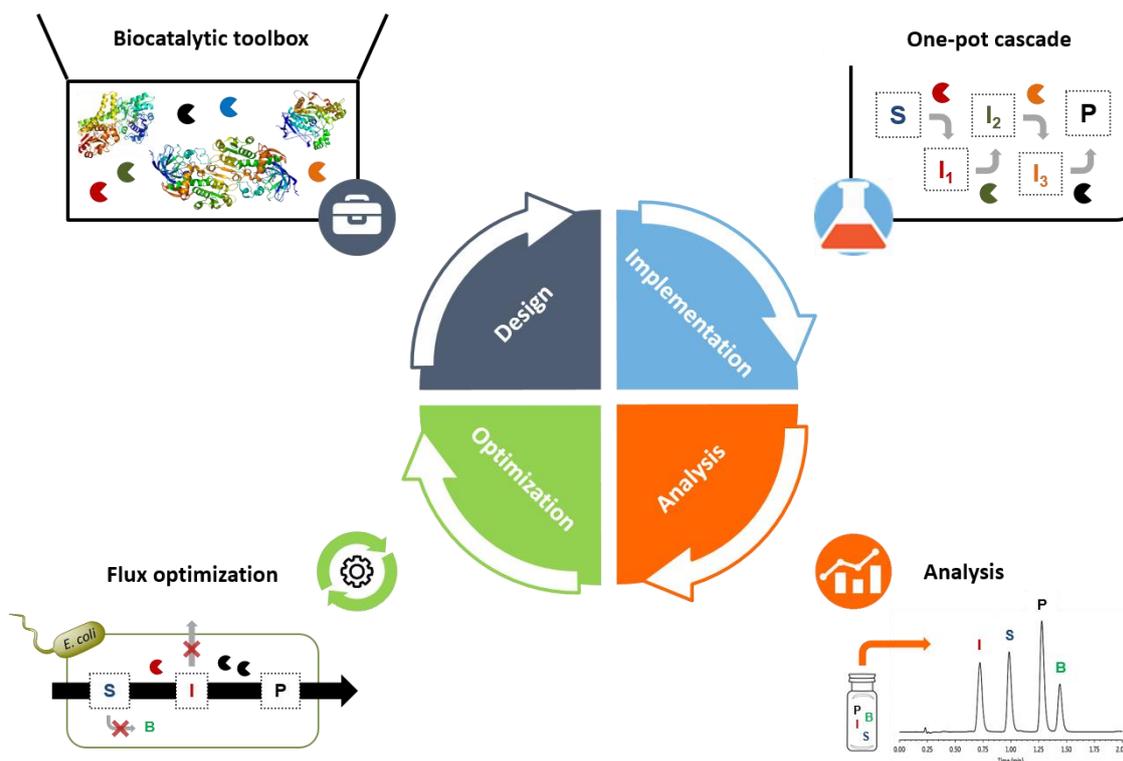


Figure C-42. Synthetic cascade design cycle. The construction of *de novo* pathways goes through rounds of design (i.e., catalyst selection), cascade implementation (e.g., *in vitro*, *in vivo*, or hybrid systems), analysis to identify bottlenecks, and optimization to enhance the pathway flux, for example, by strain engineering and synthetic biology tools.

In two main chapters (chapter D and E), the evolution of biocatalytic routes producing polyhydroxylated compounds – from the design on paper to the optimized implementation in microbial cell factories – will be described and include:

- The initial biocatalytic retrosynthetic approach to identify suitable biocatalysts to produce target sugar derivatives (D I),
- the characterization and selection of enzymes from different classes including esterases, ADHs, CARs, (D)HA- and DHAP-dependent aldolases, and phosphatases (D II),
- the assembly into ‘cascade core modules’ (D III) featuring either the engineered (D)HA-dependent aldolase *Fsa1-A129S* (D III.1) or a DHAP-dependent aldolase (*FruA*, *FucA*, and *RhuA*; D III.2), and
- the identification of bottlenecks and subsequent optimization of the flux through the *de novo* cascades (chapter E).

Chapter D focuses on both biocatalyst selection and subsequent pathway assembly. Although the constantly expanding toolbox of available biocatalysts (e.g., natural, engineered, and evolved enzymes) enables the biocatalytic retrosynthesis (D I) of target molecules, certain transformations including the oxidation of primary alcohols to the corresponding aldehydes resembled a challenge in biocatalysis not only *in vivo*. The first subchapter (D II) discusses the identification of a novel biocatalyst (*AlkJ* from *P. putida*) for the oxidation of primary aromatic alcohols, a key transformation toward polyhydroxylated compounds. Furthermore, known biocatalysts for *in vitro* transformations such as esterases like *BS2* from *B. subtilis* and different aldolases from *E. coli* were characterized toward the synthesis of sugar derivatives. The latter enzyme class catalyzes the C–C coupling between the *in situ* produced aldehyde acceptors and a donor molecule. Whereas one-step transformations involving aldolases are commonly established *in vitro*, their implementation in living cells has long been a vain endeavor. Therefore, the first subchapter also describes the transfer of two aldolases, the

engineered *Fsa-A129S* and the wild type *FucA*, from *in vitro* to *in vivo* applications. The latter is dependent on the labile (and expensive donor) molecule DHAP. Hence, the realization of a biosynthetic route hijacking DHAP from the metabolite pool of *E. coli* in this work is highly desired but depends on the addition of a subsequent enzymatic dephosphorylation step (e.g., *PhoN-Sf* from *S. flexneri* or *YqaB* from *E. coli*), which is also discussed. Finally, pathway assembly is achieved not only by standard molecular cloning techniques involving restriction enzymes and ligases; SLIC methods can assemble whole metabolic pathways in distinct genetic arrangements from linear DNA fragments *via* homologous overhangs. Only introduced recently, such cloning methods are far from being routinely used in biocatalysis and the construction of multi-enzyme pathways. Therefore, this work aims at promoting the applicability and advantages of such assembly methods.

The impact of the key transformations, *in situ* aldehyde preparation and aldol formation, characterized in previous chapter D and their influence on both *E. coli* host cell viability and target aldol adduct titers are addressed in chapter E. Consequently, subchapters are dedicated to different optimization strategies from genetic and metabolic engineering, synthetic and systems biology to enhance the flux through *de novo* pathways and, more precisely, include:

- enzyme balancing from improved genetic architectures by the introduction of genetic regulatory elements such as synthetic terminators and analysis of context effects (E 1.1),
- the identification of endogenous enzyme activities toward cascade intermediates such as aldehydes and the utilization of host strains (e.g., *E. coli* K-12 RARE) featuring engineered metabolic backgrounds to reduce byproduct formation (E 1.3),
- a complementing flux enhancement strategy based on the reversion of undesired byproduct formation (e.g., carboxylic acids) by the introduction of a nonnative enzyme with opposing functional group transformation activity (E 1.4), and
- engineering of reaction parameters and improving reaction work-up and aldol product isolation by SPE (E 1.5).

Aldehydes are important chemicals with large markets in food and pharmaceutical industries. However, their production as intermediates and products in microbes is an unmet challenge due to the high cytotoxicity of aldehydes. This special issue of aldehyde toxicity is addressed in subchapter E 1.4 by the formation of a 'hidden reservoir' for reactive aldehyde species *in vivo*. Aldehydes can be contained by *AlkJ* and a CAR from *N. iowensis* below nonviable concentrations, yet freely available for subsequent aldol formation. Additionally, the rerouting of carbon fluxes by the introduction of a reversing enzymatic activity has been largely unperceived and offers potential to maximize product titers by transforming byproducts in desired cascade intermediates. The aromatic aldol adducts produced in this thesis are precursors for sugar derivatives with potential glycosyltransferase inhibitor activity (e.g., D-fagomine). Hence, the optimized enzymatic routes established in this work provide alternatives to access polyhydroxylated compounds with promising applications in the pharmaceutical industry and medical research. Regarding aldol product isolation, *in vitro* (and *in vivo*) systems suffer from low yields. Not only are polyhydroxylated chemicals highly soluble in water, acyclic aldol adducts are temperature sensitive. Consequently, this work tackled the issue of aldol isolation and purification by refining the downstream processing by an easy to apply solid phase extraction (SPE) procedure.

The last subchapter (E 1.6) discusses potential bottlenecks of cascades featuring DHAP-dependent aldolase like *FucA* and offers different solutions to optimize this synthetic metabolic route that interconnects with the central carbon metabolism of host cells.

D Biocatalyst selection and *de novo* pathway assembly

This chapter comprises the initial retro-biosynthetic approach applied to dissect the target polyhydroxylated molecules into simple, achiral substrates. The identification of suitable chemical transformations in opposite direction of the bioretrosynthesis enabled the selection of biocatalysts from enzyme classes. Their characterization will be discussed in the first part of this chapter. In the second part, the assembly of *de novo* pathways by (advanced) cloning techniques, their transfer into suitable *E. coli* host cells, and pathway validation *in vivo* will be described.

D I Biocatalytic retrosynthesis: From sugar derivatives to simple, achiral substrates

Over the last decade, an increasing number of biocatalysts has been added to the portfolio and cherished by organic chemists to solve, by classical chemical means, challenging transformations. With advanced protein engineering techniques, biocatalysts have been tuned to function in the usually harsher environments of chemical flasks and reactor vessels and have been even equipped with new transformation activities.^[6, 17] This recent developments have inspired the design of (complex) synthetic routes involving multiple biocatalysts (i.e., enzymes) and their arrangement in cascade-type reactions.^[234, 372, 454] In 2013, Turner and O'Reilly transferred the principles of chemical retrosynthesis, a concept introduced by Corey that revolutionized synthetic planning, into guidelines for 'biocatalytic retrosynthesis'.^[68, 455]{Hönig, 2017 #1013}

In the beginning of this thesis, target compounds, precisely, polyhydroxylated chiral molecules, were disassembled *via* biocatalytic retrosynthesis (**Figure D-1**). First, the polyhydroxylated product was identified as an aldol adduct and disconnected between the two stereocenters, resulting in an aldehyde (acceptor) and an aldol donor molecule. By functional group interconversion (FGI), aldehydes were transformed into primary alcohols. A second FGI converted primary alcohols to esters, which are thought to freely pass the cellular membrane of *E. coli* and resemble suitable substrates as entry point for the *de novo* cascade.

Following the direction of synthesis, the proposed enzymatic route starts from simple and readily available carboxylic esters. Regarding the implementation of the suggested enzyme cascade in *E. coli*, esters are suitable substrates that likely pass the cellular membrane of host cell by diffusion owing to their high lipophilicity. In the first (bio)transformation, an esterase converts carboxylic esters to the corresponding primary alcohols,^[456-457] which can be oxidized to the corresponding aldehydes, for example, by an ADH.^[408] Subsequently, aldehydes and suitable donor molecules form the desired aldol adducts in an aldolase mediated reaction.^[256, 415] The resulting polyhydroxylated compounds (i.e., carbohydrate derivatives) execute biological functions including cell-cell interactions and are involved in signaling processes during bacterial and viral infections.^[453] Thus, they are interesting targets for medical research and valuable precursors in the pharmaceutical industry.^[260, 293, 348, 435, 458]

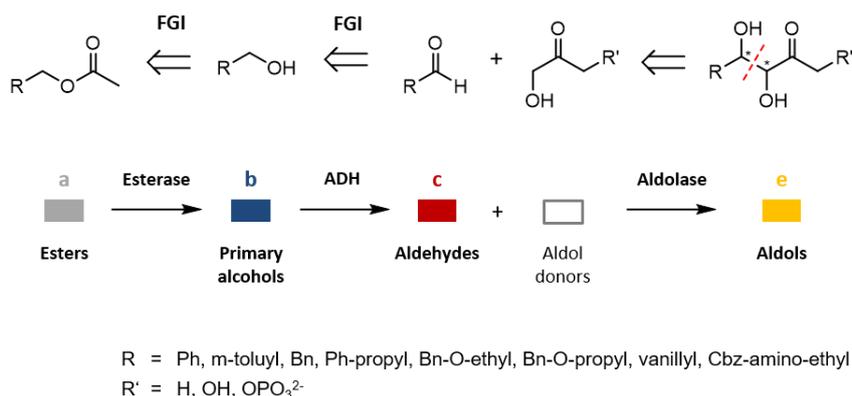


Figure D-1. Biocatalytic retrosynthesis. In the disconnection approach (top), target aldols (**e**) can be transferred into an aldehyde (**c**) and a suitable aldol donor. Functional group interconversions (FGIs) lead to primary alcohols (**b**) and, subsequently, to the corresponding carboxylic esters (**a**). In the synthesis (bottom), esters can be transformed into primary alcohols by an esterase, which are oxidized to the corresponding aldehydes, for example, by an ADH. Mediated by an aldolase, aldehydes and an aldol donor form the desired aldol adducts. The donor specificity (R') is dictated by the selected aldolase. All enzymes should accept a variety of aromatic substrates and intermediates (R ; Ph = phenyl, Bn = benzyl).

Candidate biocatalysts for subsequent pathway assembly should exhibit broad substrate acceptance toward structurally different aromatic residues R including phenyl, benzyloxy, and carbobenzyloxy (Cbz)-amino moieties (**Table D-1**) since such molecules are precursors for D-fagomine (**Figure C-1**) or 1-deoxynojirimycin.^[260, 293, 458] The utilization of different aldol donors is dictated by the aldolases, which usually exhibit relaxed substrate profiles toward the aldehyde acceptors but are very specific for the donor molecules (R') (**Figure D-1**). Wild type and engineered aldolases have been added to the biocatalytic portfolio that utilize donors such as HA ($R' = \text{H}$) and DHA ($R' = \text{OH}$). Alternatively, DHAP-dependent aldolases strictly depend on DHAP ($R' = \text{OPO}_3^{2-}$), yielding a phosphorylated aldol adduct. To facilitate the transport of aldol products out of the cell, the synthetic route needs to be expanded by an additional dephosphorylation step catalyzed by a phosphatase (**Figure D-36**).

Table D-1. Substrate pool

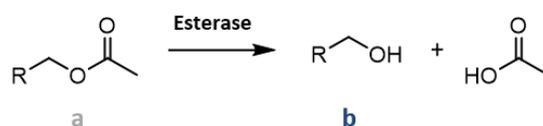
Substrate		Substrate	
1		6	
2		7	
3		8	
4		9	
5			

D II Identification and characterization of suitable biocatalysts

Based on biocatalytic retrosynthesis, biocatalysts were characterized belonging to the following enzyme classes: carboxylic ester hydrolases [EC 3.1.1] (section D II.1), ADHs [EC 1.1.1.1 and 1.1.1.2] (section D II.2), C–C lyases [EC 4.1.2] (section D II.4), and acid phosphatases [EC 3.1.3.2] (section D II.5).^[459]

During the identification of suitable biocatalysts for the oxidation of primary alcohols to the corresponding aldehydes *in vivo* (D II.2.2), carboxylic acids were identified as undesired byproducts. To potentially reduce the carboxylate side products to the target aldehyde intermediates, the characterization of CARS [EC 1.2.99.6] was also pursued.^[348, 459]

D II.1 Getting started: Expression and characterization of esterases



(Non-lypolic) esterases represent a diverse group of carboxylester hydrolases catalyzing the cleavage and formation of ester bonds in animals, plants, and microorganisms. Besides lipases with preference for triacylglycerides, esterases exhibit preference for the hydrolysis of esters bearing short-chain acyl residues.^[457]

The potential of esterases as enantioselective biocatalysts has increased in the last two decades due to the progresses achieved in different scientific areas including the recombinant overproduction of enzymes, structural information useful for understanding the rational behind properties such as enantioselectivity, efficient methods to engineer these properties, and the establishment of rapid screening methodologies.^[456-457] Since then, esterases had significant impact on the development of new robust biocatalysts with high stereo- and enantioselectivities and have been successfully applied as biocatalysts in organic synthesis, not only for the resolution of racemates.^[456-457, 460-462]

As outlined above, esters can be cleaved by carboxylester hydrolases into a primary alcohol and the acyl compound. Compared to the corresponding primary alcohols, esters are rather apolar compounds that are thought to freely pass cellular membranes. Regarding substrate uptake and the biocatalytic retrosynthesis approach (**Figure D-1**), esterases from different microorganisms were characterized in this work and will be described in the following.

Three microbial esterases from *B. subtilis* (BS2),^[463] *P. fluorescens* (PfeI),^[464-465] and *Pyrobaculum calidifontis* VA1 (PEST)^[466] were cloned and/or contributed by the group of Bornscheuer. All three esterases were successfully expressed in *E. coli* BL21(DE3) transformants harboring pET28a_{bs2}, pGASTON_{pfeI}, and pET21a_{pest}, respectively (**Figure D-2A–B**). CFEs were prepared as described in G I.5 and their activities toward p-nitrophenyl acetate (pNPA) were tested (**Figure D-2C**).

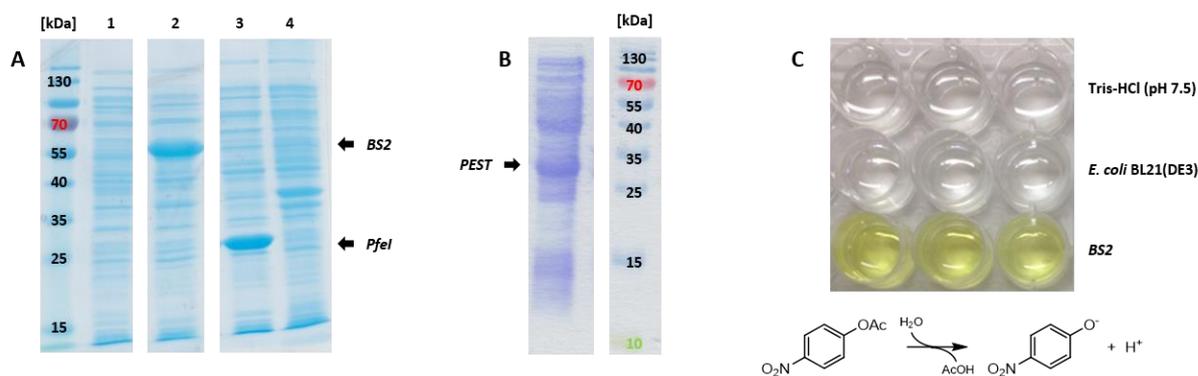


Figure D-2. Expression of different esterases. (A) SDS-PAGE analysis of CFEs containing *BS2* (2) and *Pfel* (3). For comparison, untransformed *E. coli* BL21(DE3) cells were cultivated as *BS2* and *Pfel* transformants (1 and 4, respectively). (B) CFE from *E. coli* BL21(DE3) expressing *PEST*. Protein production was performed as in G IV. Sample loading normalized to 10 µg total amount of protein per lane. (C) Photometric assay monitoring the hydrolysis of pNPA in buffer (top), CFE from untransformed cells (center), and in the presence of *BS2* (bottom); Ac = -COCH₃. Assay performed in triplicates according to G VI.1.1.2.

CFEs containing *BS2* and *Pfel* rapidly hydrolyzed pNPA in less than 10 min incubation time, whereas *PEST* reached the maximal absorption of $A_{405} = 1.0$ after 30 min reaction time (data not shown). Subsequently, biotransformations were performed in duplicates with resting cells (RCs) expressing *BS2* or *Pfel*, following the standard screening procedure outlined in G III.2. *BS2* quantitatively converted the selected esters 2-phenylethyl acetate (2a), 2-(benzyloxy)ethyl acetate (3a), and 3-(benzyloxy)propyl acetate (4a) to the corresponding primary aromatic alcohols 2-4b in only 2 h reaction time. *Pfel* accepted the same model substrates producing the target alcohols at slower rates than *BS2* (Table D-2).

Table D-2. Substrate scope of *BS2* and *Pfel* for the production of primary aromatic alcohols

Substrate	Product formation [%]			
	<i>BS2</i>		<i>Pfel</i>	
	2 h	24 h	2 h	24 h
2a	≥99	≥99	20	92
3a	≥99	≥99	75	≥99
4a	≥99	≥99	15	91

Results presented as mean values of duplicates measured by GC/FID with recoveries >75%. Screenings were performed in *E. coli* BL21(DE3) RCs (OD₅₉₀ = 10.0) expressing esterases, starting with 5 mM substrate and 5% (v/v) ACN as cosolvent at 25°C (250 rpm).

In summary, *BS2*, *Pfel*, and *PEST* showed esterase activity in preliminary photometric assays hydrolyzing pNPA. The tested esters 2-4a could pass the cell membrane of *E. coli* BL21(DE3) by passive diffusion and were efficiently converted to the corresponding aromatic alcohols 2-4b. Hence, RC experiments demonstrated the permeability of selected esters from compound library B I through the cellular barrier and *BS2* from *B. subtilis* as well as *Pfel* from *P. fluorescens* were identified as suitable biocatalysts for the preparation of primary aromatic alcohols from esters *in vivo*. Ultimately, esterases were not implemented in the construction of the synthetic enzyme cascade since it was shown that all alcohol substrates (B I) can diffuse through the cellular membrane (Figure D-6; see D II.2.2), importantly, without limiting the rate of conversion (Figure D-29).

D II.2 Enzymes for the oxidation of primary alcohols to aldehydes *in vivo*



The next cascade step involves the oxidation of primary alcohols to the corresponding aldehydes. Aldehydes do not accumulate in most native microorganisms and are rapidly metabolized *in vivo*.^[42, 280, 407] This issue was tackled by the selective inactivation of ADHs, AIDHs, and AKRs by conventional gene KO strategies to accumulate aldehydes.^[42, 467] However, the high reactivity of aldehydes contributes to cellular toxicity, an obstacle of *in vivo* applications that has not been solved yet.^[280-281, 400]

Oxidative and electrophilic stress in bacteria can damage proteins, lipids, RNA, and DNA, drastically reducing cell viability. Unlike free radicals like hydroxyl radicals generated from H₂O₂, a product of the respiratory chain in *E. coli*,^[468] aldehydes are rather long lived. Therefore, they can diffuse from the site of their origin and can form adducts with proteins and DNA, inactivating enzymes, damage DNA, and, ultimately, causing cell death.^[281, 400, 469] Bacterial genetic responses to oxidative stress are controlled by two major regulators: *OxyR* and *SoxRS*. Regarding oxygen-derived radicals, defense mechanisms include the reduction of such radicals to viable levels or the repair of oxidative damage. The transcriptional activator *OxyR* mediates defense against peroxides and *SoxRS* against superoxide.^[470-471] Constitutively present molecules including NAD(P)H, ascorbic acid, iron (Fe²⁺), or glutathione help to maintain an intracellular reducing environment or to scavenge reactive oxygen species.^[400] To specifically scavenge aldehydes, the expression of AIDHs is upregulated in bacteria. Besides various ADHs, the AIDH superfamily represents an essential enzyme class to reduce oxidative/electrophilic stress in living cells.^[281] Endogenous AIDHs depend on NAD(P)⁺ to oxidize aldehydes to the corresponding carboxylates and can metabolize a variety of chemically and structurally diverse aldehydes.^[281, 472]

Besides the metabolic background and host cell responses disfavoring the production of aldehydes, there are only a few examples of biocatalysts for the oxidation of primary (aromatic) alcohols to the corresponding aldehydes in the literature.^[280, 407-409] *Pichia pastoris* whole cells overexpressing an endogenous alcohol oxidase (AOX) were employed to oxidize benzyl alcohol (**1b**) and 3-phenyl propanol (**3b**) to benzaldehyde (**1c**) 3-phenyl propanal (**3c**) using a two-liquid system for *in situ* product removal.^[411] Similarly, 2-Phenyl acetaldehyde (**2c**) was produced from 2-phenyl ethanol (**2b**) using acetic acid bacteria.^[407] The salicyl alcohol oxidase (SAO) from *Chrysomela* sp. was shown to oxidize **2b** with 21% activity in comparison to the natural substrate salicyl alcohol in the native host, whereas activity was reduced to <4% after recombinant expression of SAO in *E. coli*.^[473]

However, biocatalysts with a broad and overlapping substrate profile are the basis to transform simple *E. coli* strains into microbial cell factories, capable of producing a variety of target compounds *via* the synthetic enzyme cascade. According to the BRENDA database (www.brenda-enzymes.org), there has been no entry of a single biocatalyst to show a broad substrate profile for the primary aromatic alcohol substrates in B I. Consequently, screening of different ADHs for the desired transformation was conducted and will be described in the following. AOXs were excluded to circumvent the coexpression of a catalase and to keep the initial metabolic burden on the host low.

D II.2.1 Unsuccessful substrate acceptance screening of *ADH_{Lk}*, *ADH_{Rr}*, *ADH-A*, and *ADH-ht*

The first two ADHs tested were *ADH_{Lk}* from *L. kefir*^[474] and *ADH_{Rr}* from *R. ruber*.^[44] Therefore, both enzymes were expressed in *E. coli* BL21(DE3) and CFEs were produced as described in G VI.2.1.1 (**Figure D-3A** and **Figure G-4**).

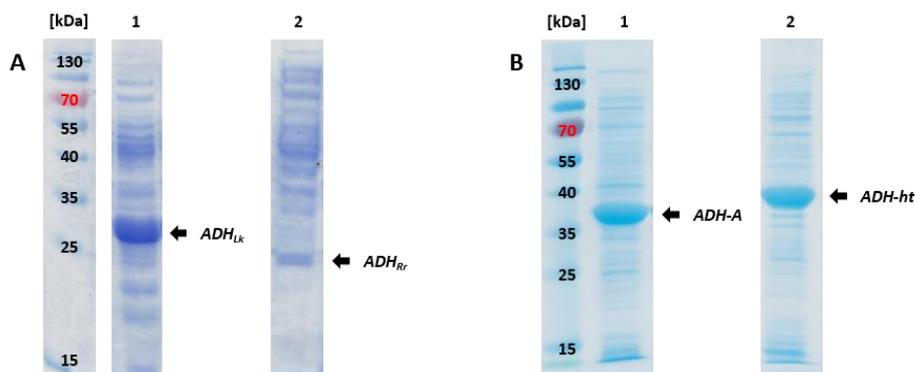


Figure D-3. Successful expression of different ADHs. (A) SDS-PAGE analysis of CFEs containing *ADH_{Lk}* (1) and *ADH_{Rr}* (2). Sample loading normalized to 10 μ g total amount of protein per lane. (B) Whole cell samples containing *ADH-A* (1) and *ADH-ht* (2). Sample loading normalized to $OD_{590} = 7.0$. Single enzyme productions were performed in *E. coli* BL21(DE3) under optimized conditions according to G IV.

Biotransformations were carried out following the procedure in G VI.2.1.2 with a slight excess (4.25 mM) of $NADP^+$ and NAD^+ for the substrate acceptance screenings with *ADH_{Lk}* and *ADH_{Rr}*, respectively. Alcohols **1–3b**, 2-benzyloxy ethanol (**4b**), or 3-benzyloxy propanol (**5b**) were added last (4 mM). The reaction was monitored by GC/FID at 2 h and 24 h reaction time.

Traces of the target aldehyde **3c** and the 3-phenyl propionic acid (**3d**) byproduct were detected in the presence of *ADH_{Lk}* after long reaction times. The target aldehydes **1–2c** and **4–5c** could not be detected (Figure D-4A). *ADH_{Rr}* did not display oxidation activity toward the tested primary alcohols **2b** and **4–5b**. Traces of aldehyde **1c** and carboxylate byproduct **3d** were detected after 24 h reaction time (Figure D-4B). Since *ADH_{Lk}* and *ADH_{Rr}* did not accept the desired substrates and predominantly oxidized secondary alcohols, they did not represent suitable biocatalysts for *de novo* pathway assembly.^[435]

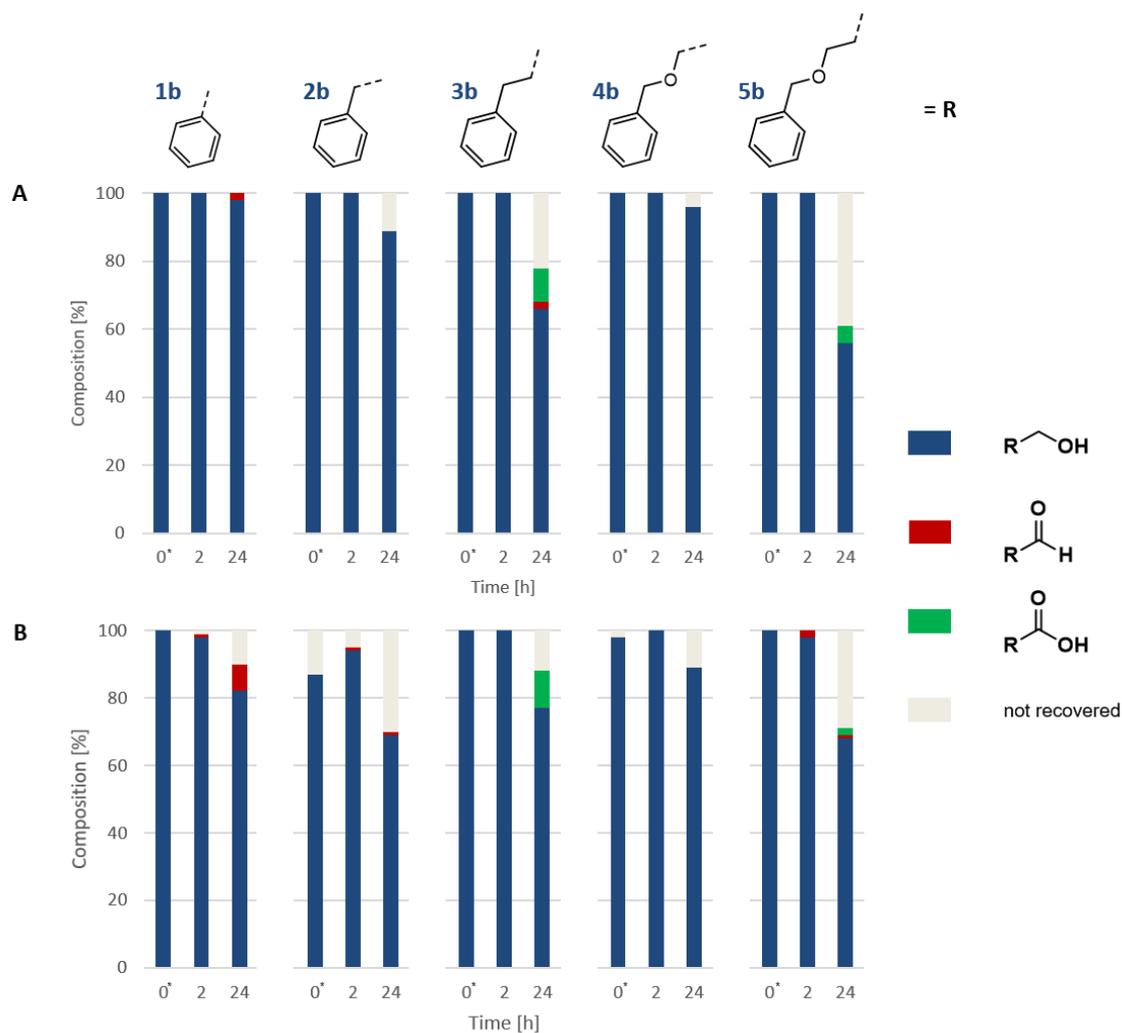


Figure D-4. Substrate scope of ADH_{Lk} and ADH_{Rr} toward primary aromatic alcohols. (A) ADH_{Lk} displayed only neglectable activity toward the alcohol substrates **1–5b**. (B) ADH_{Rr} insufficiently oxidized the same substrates to the corresponding aldehydes, if at all. Screenings were performed according to G I.5. Results presented as mean values of two independent experiments monitored by calibrated GC/FID; t_0^* sample taken immediately after addition of substrate and mixing. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey)

$ADH-A$, another secondary ADH from *R. ruber*, was reported to oxidize **2b** to **2c**.^[475] After optimization of $ADH-A$ production in *E. coli* BL21(DE3) (Figure D-3B),^[476] biotransformations were performed with resting cells (RCs) in duplicates, following the standard screening procedure outlined in G III.2. However, $ADH-A$ showed no activity for the reported **2b**, substrates **3–5b**, and 3-(benzyloxycarbonylamino) propanol (**6b**) under experimental conditions (Table D-3).^[435] Furthermore, the thermostable $ADH-ht$ from *B. stearothermophilus* was shown to oxidize primary alcohols.^[331, 477] The $adh-ht$ gene was successfully subcloned into pET26b(+) applying the refined ‘Florida’ cloning procedure according to G VI.2.2.1 and sequence integrity was confirmed by Sanger sequencing (H I.1.3.2).^[435] $ADH-ht$ was expressed under optimized conditions in *E. coli* BL21(DE3) as outlined in G VI.2.2.1 (Figure D-3B and Figure G-10). However, under standard screening conditions, only the Cbz-protected amino alcohol **6b** was accepted, yielding 27% of the desired 3-(benzyloxycarbonylamino) propanal (**6c**) after 2 h reaction time according to calibrated GC. Alcohols **1–5b** were not accepted (Table D-3).^[435]

Table D-3. Substrate scope of *ADH-A* and *ADH-ht* toward primary aromatic alcohols

Substrate	Product formation [%]			
	<i>ADH-A</i>		<i>ADH-ht</i>	
	2 h	24 h	2 h	24 h
2b	n.c.	[a]	n.c.	n.d.
3b	n.c.	n.c.	n.c.	n.d.
4b	n.c.	[a]	n.c.	n.d.
5b	n.c.	n.c.	n.c.	n.d.
6b	n.c.	[a]	27	n.d.

Results presented as mean values of duplicates measured by GC/FID with recoveries >75%; n.c. = no conversion, n.d. = not determined. The results for *ADH-ht* were adapted from T. Wiesinger *et al.* (2017).^[435]

[a] Traces of target aldehyde and/or carboxylate byproduct

In summary, all four tested ADHs were successfully expressed in *E. coli* BL21(DE3) (**Figure D-3**). *ADH_{Lk}*, *ADH_{Rr}*, and *ADH-A* showed no or only minimal activity toward a few aromatic alcohol model substrates **1–5b**; *ADH-ht* exclusively converted **6b** to **6c** (B I). Consequently, these ADHs had to be discarded as candidates for *de novo* pathway assembly to produce target aldehyde intermediates for subsequent aldol formation. However, *ADH_{Lk}* and *ADH_{Rr}* were successfully implemented in a redox cascade for the asymmetric synthesis of lactones *in vivo*.^[44] *ADH_{Lk}* was also N-terminally tagged with a self-cleaving intein tag for purification.^[478-480] Preliminary expression was performed as described in G VI.2.1.3 (**Figure G-7**) and purification endeavored by S. Milker.^[154] Furthermore, the optimized production of *ADH-A* in *E. coli* BL21(DE3) and the preparation of lyophilized cells resulted in a robust and powerful biocatalyst to reduce ketones to the corresponding secondary alcohols.^[476]

D II.2.2 Successful cloning and expression of *AlkJ* and substrate scope expansion toward non-native compound classes

The *alkJ* gene is encoded by the *alkBGHJKL* cluster in *P. putida* (H I.1.11.1) and is translated into a membrane-associated ADH that was previously reported as an efficient biocatalyst for the oxidation of primary aliphatic alcohols.^[481] Sequence analysis of the *alkJ* gene reveals that many recognition sites of commonly used restriction enzymes (e.g., *NcoI*, *NdeI*, and *XbaI*) are present within the ORF. Hence, conventional molecular cloning into, for example, standard pET vectors was not feasible due to the forbidden restriction sites. The FC procedure by Li *et al.*^[87] was adapted to circumvent the use of any restriction enzymes and DNA ligases to seamlessly assemble the *alkJ* gene and the target pKA1 vector. Primers for the *alkJ* gene and the pKA1 backbone were designed with matching homologous overhangs (33–35 bp homology) for the directed assembly of the two linear fragments post transformation of competent *E. coli* TOP10 cells. The detailed FC procedure is given in G VI.2.4.1. Colony PCR confirmed the presence of the *alkJ* gene in transformants (**Figure D-5A**). The sequence was confirmed by Sanger sequencing (H I.1.3.6).^[348] Since *AlkJ* is a membrane-associated enzyme, it was exclusively found in insoluble protein fractions after IPTG-induced production in *E. coli* BL21(DE3) (**Figure D-5B**).^[481]

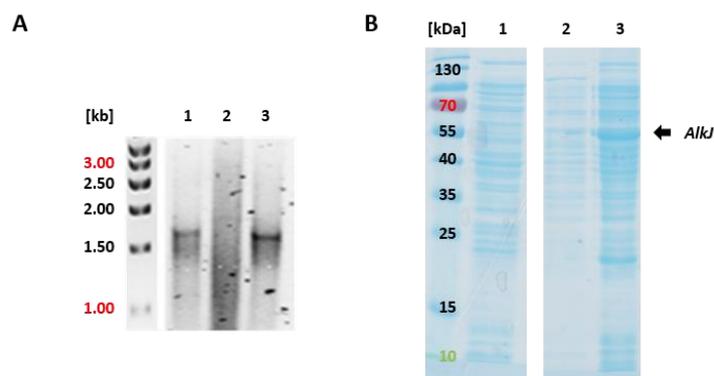


Figure D-5. Successful FastCloning and expression of *AlkJ*. (A) Colony PCR of three *E. coli* TOP10 transformants, AJ1 (1), AJ2 (2), and AJ3 (3). Sanger sequencing of AJ1 and AJ3 plasmid DNA confirmed the integrity of the *alkJ* sequence. (B) SDS-PAGE analysis of CFEs from untransformed *E. coli* BL21(DE3) (1); *AlkJ* expressed from pKA1_alkJ (clone: AJ3) in soluble (2) and insoluble (3) fractions. Protein production: G IV. Sample loading normalized to 10 μ g total protein per lane.

To confirm *AlkJ* activity and to determine the substrate scope, subsequent screenings were performed in RCs under standard screening conditions (Figure D-6).

In the presence of *AlkJ*, especially the substrates **2b**, **4–6b** bearing aliphatic C2 and longer side chains were quickly converted. Surprisingly, **3b** was insufficiently oxidized to aldehyde **3c**. Alcohols with substitutions on the aromatic ring like m-toluoyl alcohol (**7b**) and vanillyl alcohol (**9b**) and the short side chain aldehyde **1b** were moderately converted to the desired aldehydes.

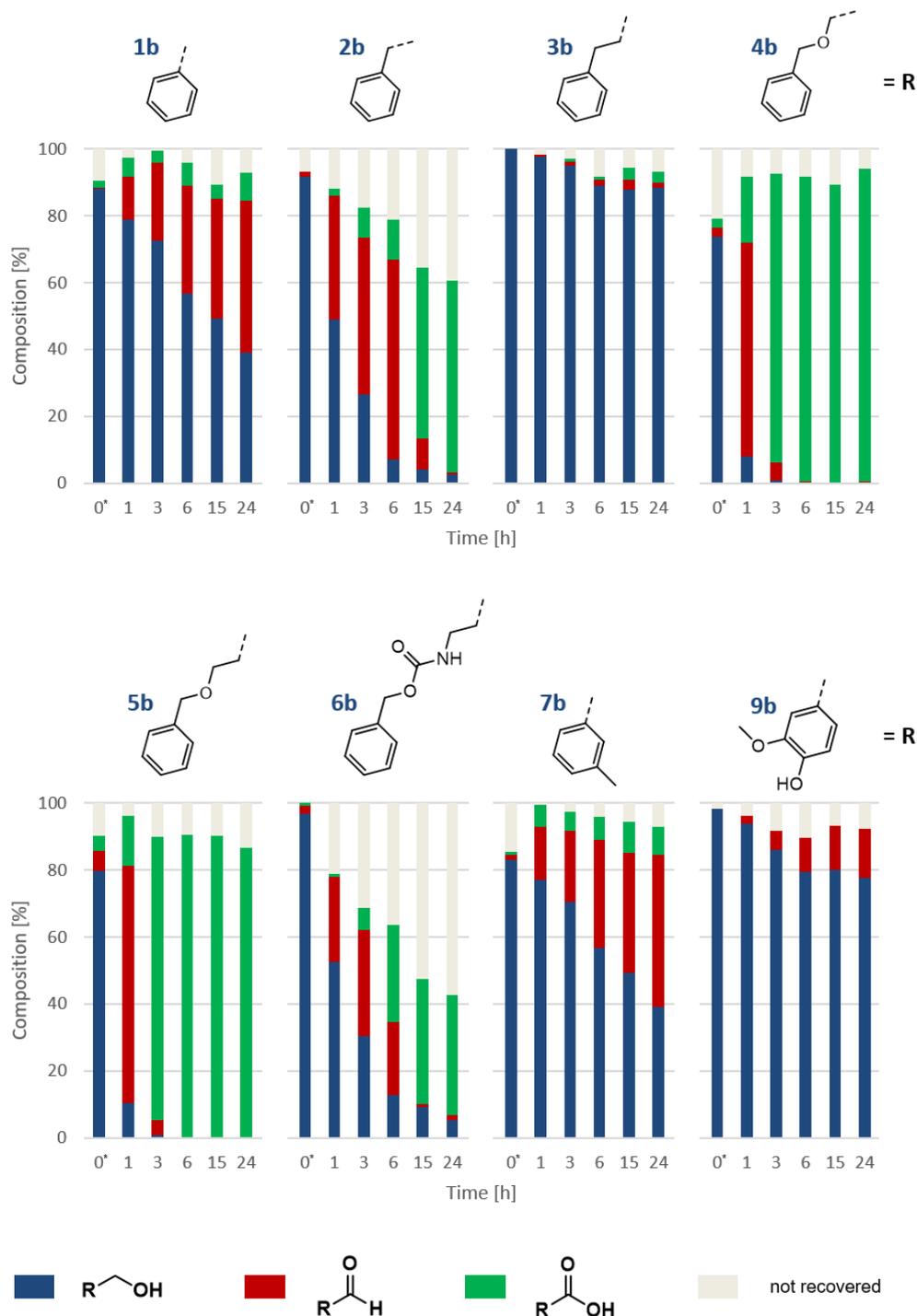


Figure D-6. Substrate profile of *AlkJ*. The primary aromatic alcohols 1–7b and 9b were oxidized to the desired aldehydes. Overoxidation to the corresponding carboxylates was observed in most transformations. Screenings were performed under standard screening conditions (G III.2). Results presented as mean values of biological triplicates ($n = 3$); SD <10% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material at t_0^* due to insufficient mixing and a loss in mass balance (e.g., volatility) at later time points. Parts of this figure were adapted from T. Bayer *et al.* (2017).^[348]

Although the ADH from *P. putida* was reported to oxidize primary (aliphatic) alcohols selectively to the corresponding aldehydes,^[391] the prominent overoxidation of 2c, 4–6c to carboxylic acids 2d, 4–6d could be observed in *E. coli* BL21(DE3) under experimental conditions.^[348] Since endogenous AIDHs can metabolize the produced aldehydes to the carboxylates to reduce oxidative/electrophilic stress,^[281, 472] further RCs were

performed to determine the oxidation capacity of the enzymatic background of *E. coli* BL21(DE3) toward a selection of aromatic alcohols (**Figure D-7**).

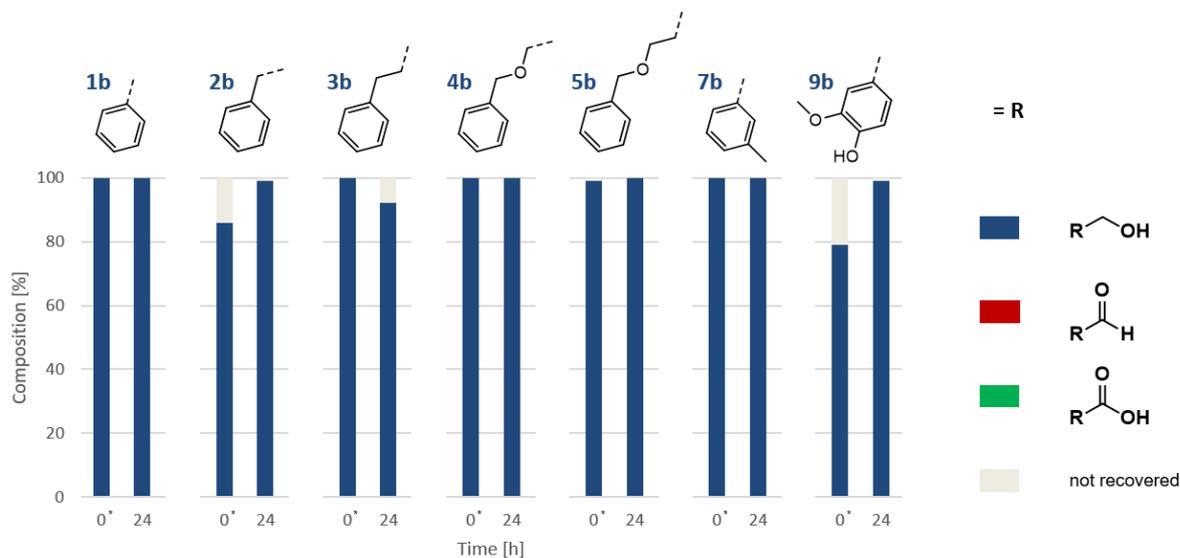


Figure D-7. Primary aromatic alcohols in the cellular environment. None of the primary aromatic alcohols **1–5b**, **7b**, and **9b** was oxidized by the metabolic background of *E. coli* BL21(DE3). Results presented as mean values of biological triplicates ($n = 3$); SD <5% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey); t_0^* sample taken immediately after the addition of 5 mM substrate and mixing.

Additionally, the contribution of *AlkJ* to the oxidation of aldehydes was determined in the metabolic backgrounds of *E. coli* BL21(DE3) and the engineered RARE strain.^[42] Therefore, RCs of both hosts expressing *AlkJ* were challenged with 5 mM of 2-phenyl ethanal (**2c**). Untransformed RCs were prepared and treated accordingly. *E. coli* BL21(DE3) mainly reduced **2c** to the alcohol **2b**, whereas the amount of 2-phenyl acetic acid (**2d**) increased about 2-fold in the presence of *AlkJ*, from $4.3 \pm 0.8\%$ to $9.0 \pm 1.3\%$ (**Figure D-8A**). In contrast, *E. coli* RARE was engineered to minimize the reduction of aromatic aldehydes and **2c** was slowly oxidized to **2d**. The amount of **2d** increased on an average of $3.3 \pm 1.1\%$ in *E. coli* RARE cells expressing *AlkJ* (**Figure D-8B**). Noteworthy, low recoveries may not only be attributed to the volatility of the aldehyde **2c** but crosslinking to proteins, for example, and a reduced cell viability due to the initially high aldehyde concentration. Experiments with the model aldehyde **2c** indicated that both endogenous enzymatic activities and *AlkJ* contribute to the oxidation of aldehydes in the cellular environment.

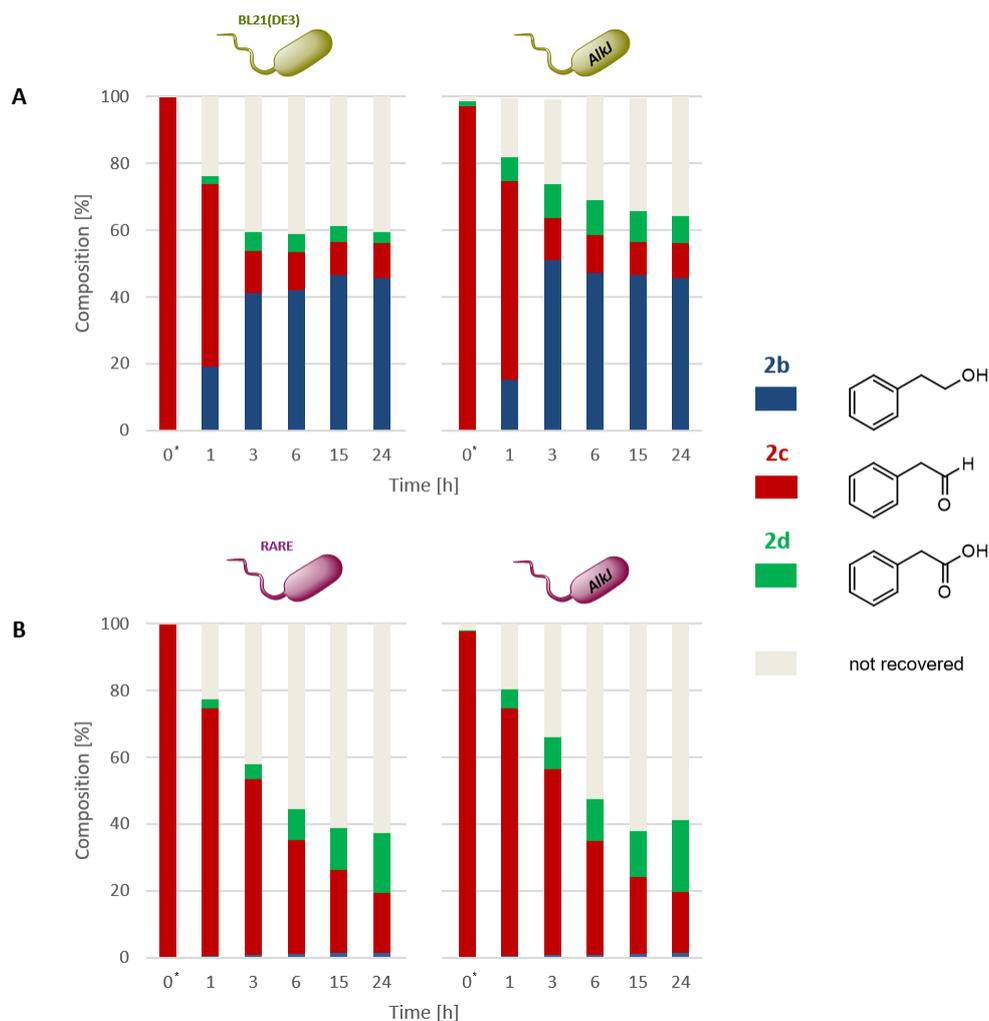


Figure D-8. Contribution of *AlkJ* to the oxidation of aldehydes in different host strains. (A) Untransformed *E. coli* BL21(DE3) mainly reduced the aldehyde **2c** to the corresponding alcohol **2b** (left). RCs expressing *AlkJ* showed slightly increased amounts of the carboxylic acid **2d** (right). (B) Untransformed *E. coli* RARE oxidized **2c** to **2d**, whereas the reduction to the alcohol **2b** was not observed (left). RCs expressing *AlkJ* showed minimally increased amounts of **2d** (right). Results presented as mean values of biological triplicates ($n = 3$); $SD \leq 10\%$ according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey); t_0^* sample taken immediately after the addition of 5 mM **2c** and mixing. Reduced recoveries and the loss in mass balance are attributed to volatility of **2c** and reduced RC viability due to initially high load of aldehyde. Parts of this figure were adapted from T. Bayer *et al.* (2017).^[348]

Finally, control experiments were performed to determine whether the uptake of primary aromatic alcohols through the cellular envelope limits their oxidation by *AlkJ*. Therefore, RCs were incubated in resting cell medium (RCM) supplemented with 1% (v/v) toluene and 5 mM EDTA final concentration. EDTA was added from a 250 mM stock (pH 8.0). RCs were preincubated at 4°C without shaking for 0.5 h, and centrifuged (5 000 rpm, 4°C, 15 min). Cells were resuspended in fresh RCM before 5 mM **2–4b** were added and the (over)oxidation followed by GC/FID as before. However, the oxidation of the tested alcohols was not accelerated and resulted in a similar oxidation pattern as shown in **Figure D-6**. Again, **3b** yielded only traces of **3c** and **3d** (data not shown).

In summary, the ADH *AlkJ* from *P. putida* was identified as an efficient biocatalyst for the oxidation of primary aromatic alcohols **1–2b** and **4–7b** for the *in situ* preparation of aldehydes. *AlkJ* even showed low activity toward **9b**, the reduced form of the industrially important aldehyde vanillin (**9c**; **Figure D-6**). Recently, the group of Li also introduced *AlkJ* as a powerful biocatalyst for synthetic pathway applications.^[43] Under experimental

conditions, target aldehydes were further oxidized to the corresponding carboxylic acid byproducts.^[348] Since *E. coli* does not exhibit endogenous carboxylate reducing activity (Figure D-9), the formation of the thermodynamically favored carboxylic acids represented an irreversible sink for aldehyde intermediates.^[389, 391] This potentially drives the pathway flux into a dead end (Figure D-6), a bottleneck that was approached by the introduction of a reversing enzyme activity and established as a complementing tool to reroute the flux through the *de novo* pathway in this thesis and for future applications (E I.4). Alternatively, byproduct formation could be prevented in a cascade type reaction by coupling the aldehyde forming reaction with a subsequent (fast) enzymatic transformation (e.g., aldolase; D III.1.2 and E I.4).^[348]

D II.3 Investigation of carboxylic acid reductases (CARs) to access the intracellular sink of carboxylate byproducts



As shown experimentally, a variety of aromatic primary alcohols were oxidized *via* aldehyde intermediates to the corresponding carboxylic acids, a common issue of the preparation of aldehydes *in vivo*.^[407] To confirm that *E. coli* does not exhibit endogenous carboxylate reducing activity toward selected carboxylic acids (B I), *E. coli* BL21(DE3) RCs were prepared as before and incubated in the presence of 5 mM carboxylic acids **1–5d**, **7d**, and **9d**. None of the carboxylates was reduced by the metabolic host background (Figure D-9).

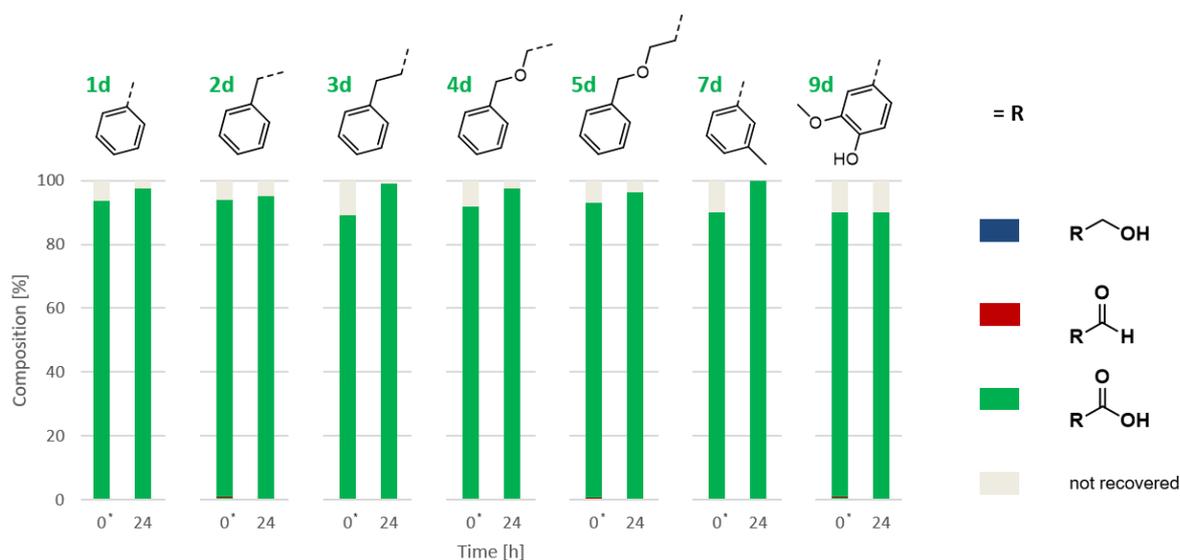


Figure D-9. Aromatic carboxylic acids in the cellular environment. None of the carboxylates **1–5d**, **7d**, and **9d** was reduced by the metabolic background of *E. coli* BL21(DE3). Results presented as mean values of biological triplicates ($n = 3$); $\text{SD} \leq 5\%$ according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey); t_0^* sample taken immediately after the addition of 5 mM substrate and mixing.

Therefore, the irreversible formation of the thermodynamically favored carboxylates, indeed, represented a sink for aldehyde intermediates *in vivo*.^[391] CARs from different species were previously shown to reduce a variety of

carboxylic acid substrates and offered potential to access the intracellular carboxylate sink and to reverse the formation of these byproducts.^[42, 404-405, 482] The unfavored conversion from carboxylates to aldehydes requires the ATP-dependent activation of carboxylates. To be reduced by NADPH, the activated intermediate needs to be shifted from the activation domain to the reduction domain by a phosphopantetheine residue (**Figure C-36**).^[64, 405]

To perform this demanding reduction, the well-studied *CAR_{Ni}* and the CAR from *Mycobacterium marinum* (*CAR_{Mm}*) were cloned by the group of Winkler as described in G VI.3.1.1.^[348, 417] Competent *E. coli* BL21-Gold(DE3) cells were transformed. For subsequent prescreenings, CARs were coexpressed with a phosphopantetheinyl transferase from *E. coli* (*PPTase_{Ec}*)^[483] from pETDuet-1 vectors by autoinduction (**Figure D-10**). *PPTase_{Ec}* posttranslationally modifies *CAR_{Ni}* and *CAR_{Mm}*, which is necessary for their activity.^[405]



Figure D-10. Expression of CAR enzymes. SDS-PAGE analysis of CFEs of untransformed *E. coli* BL21(DE3) (**1**), CFEs containing *CAR_{Ni}* (**2**) and *CAR_{Mm}* (**3**). CARs coexpressed with *PPTase_{Ec}* from pETDuet-1 plasmids. *PPTase_{Ec}* below detection limit under experimental conditions as expected.^[348] Protein production in AIM as in G VI.3.1.1. Sample loading normalized to OD₅₉₀ = 7.0.

Functional expression of both CARs in RCs was tested with **1d** since benzaldehyde is accepted by many CAR enzymes, **2d**, and **7d**. Whereas *CAR_{Ni}* readily converted all three carboxylates to the target aldehydes, *CAR_{Mm}* only accepted **1d** (**Figure D-11**).

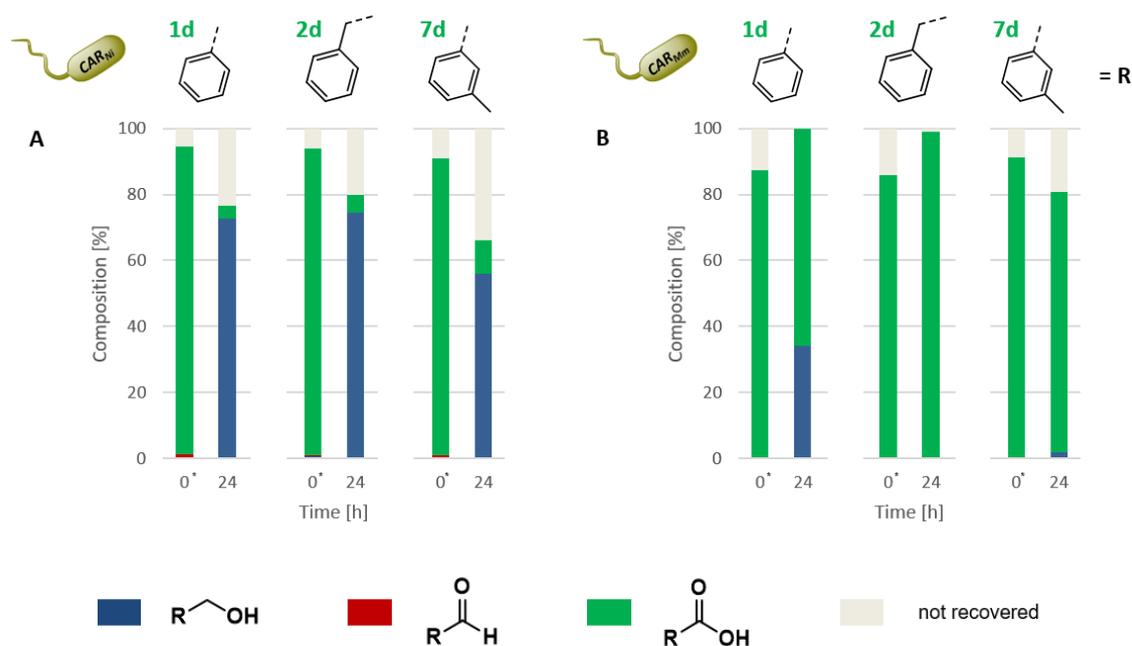


Figure D-11. Biocatalyst prescreening: CAR_{Ni} and CAR_{Mm} . (A) CAR_{Ni} reduced the tested carboxylates, which were further reduced by the enzymatic host background of *E. coli* BL21-Gold(DE3). (B) CAR_{Mm} moderately reduced the standard substrate **1d** but not **2d** and **7d**. RC screenings carried out under standard conditions. Results presented as mean values of duplicates. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey); t_0^* sample taken immediately after the addition of 5 mM substrate and mixing.

Subsequently, RCs coexpressing CAR_{Ni} and $Pptase_{Ec}$ were provided with carboxylic acids **1d–7d** and **9d**. All tested carboxylates were efficiently reduced to the corresponding aldehydes, which were rapidly converted into primary alcohols **1b–7b** and **9b** by the enzymatic host background in response to the oxidative and electrophilic stress induced by the reactive carbonyl group in aldehydes (Figure D-12).^[42, 280-281, 348, 400, 404]

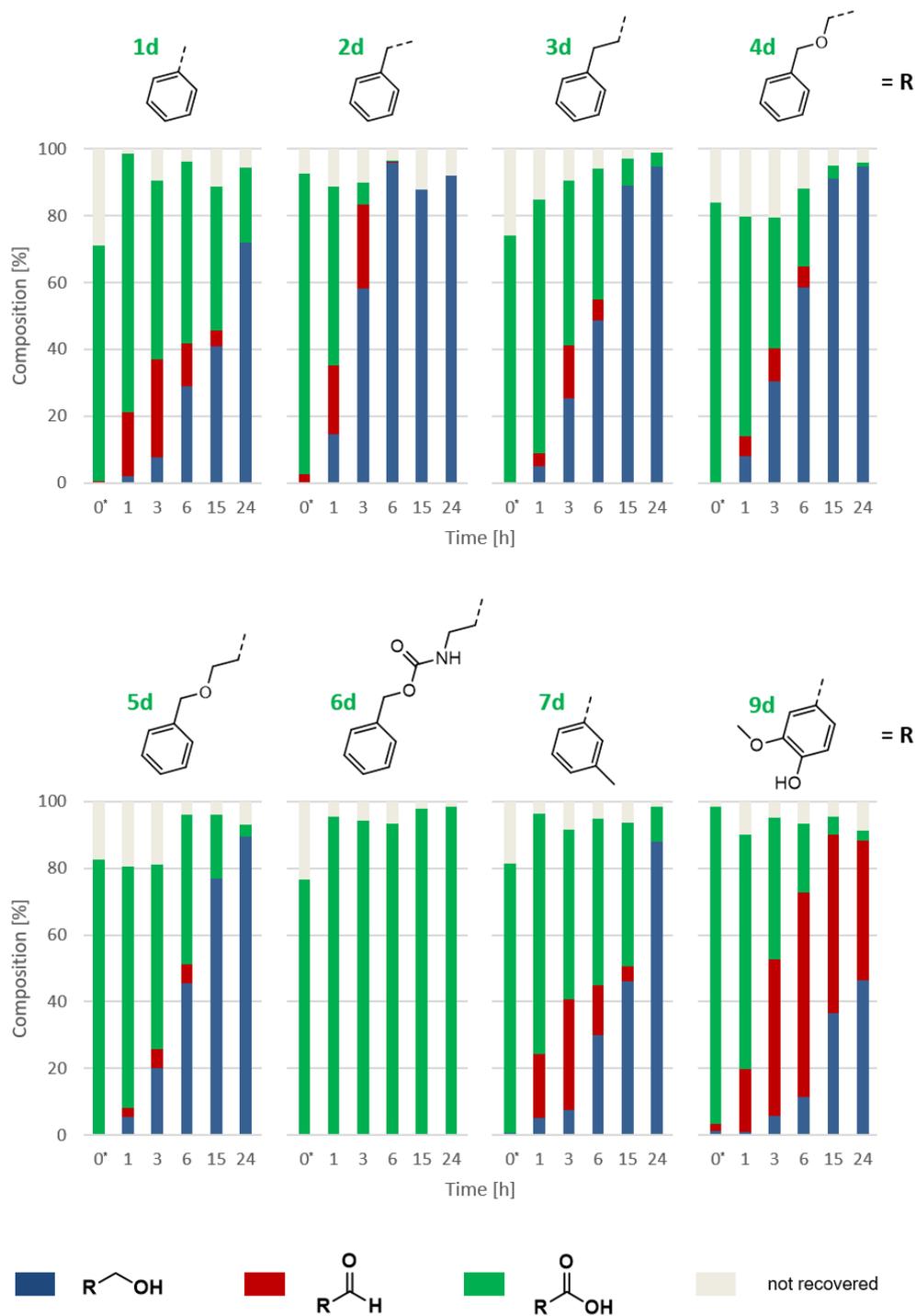


Figure D-12. Substrate profile of CAR_N . The carboxylic acids 1–6d, 7d, and 9d were reduced to the desired aldehydes. Further reduction to the corresponding primary alcohols was observed *in vivo*. Screenings were performed in *E. coli* BL21-Gold(DE3) RCs; to* sample taken immediately after the addition of 5 mM substrate and mixing. Results presented as mean values of biological triplicates (n = 3); SD <10% according to calibrated GC/FID. Error bars omitted for clarity. Bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material due to insufficient mixing and/or the volatility of produced aldehydes. Parts of this figure were adapted from T. Bayer *et al.* (2017).^[348]

To verify that any reduction of target aldehydes can be exclusively attributed to the metabolic background, CAR_N was expressed (and posttranslationally modified) in *E. coli* BL21-Gold(DE3) transformants as before and purified *via* its N-terminal 6x histidine (6xHis) tag by metal affinity chromatography as described in G VI.3.1.2. *In vitro* reductions were performed according to G VI.3.1.3. For NADPH cofactor regeneration, a 6xHis tag purified,

double mutated GDH from *B. subtilis* (GDH_{2xBs} ; $0.05 \text{ mg} \cdot \text{mL}^{-1}$) and 100 mM D-glucose were used (**Figure D-13**).^[332] GDH_{2xmut} expression and purification are described in detail in G VI.7.2.2. The reduction was started by the addition of 0.1 mM NADP^+ and followed by GC/FID as usual.

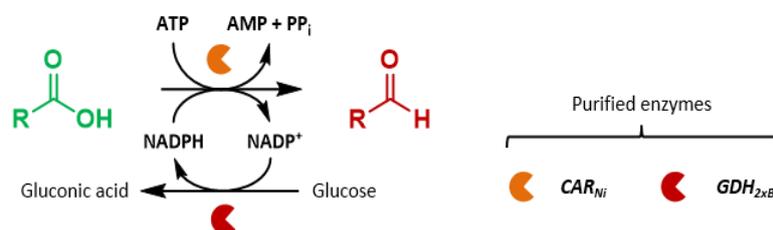


Figure D-13. *In vitro* reduction by CAR_{Ni} and GDH_{2xBs} for NADPH cofactor regeneration. Carboxylates were selectively reduced to aldehydes by purified CAR enzyme and a cofactor recycling enzyme.

Whereas **1d** and **2d** were selectively reduced to the target aldehydes **1c** and **2c**, respectively, cinnamic acid (**8d**) was not reduced under experimental conditions (**Figure D-14A**). This might be due to insufficient amounts of holo- CAR_{Ni} in purified concentrates to reduce poorly accepted substrates. Additionally, the lack of $PPTase_{Ec}$ *in vitro* cannot compensate for a loss of holo- CAR_{Ni} (e.g., inactivation) by modifying another apo- CAR_{Ni} molecule. Since the corresponding alcohol byproducts were not detected and *in vitro* biotransformations starting from the aldehydes **1c** and **2c** did not yield other reaction products (**Figure D-14B**), this experiment confirmed that endogenous enzyme activities exclusively facilitated the reduction of aldehydes *in vivo*.

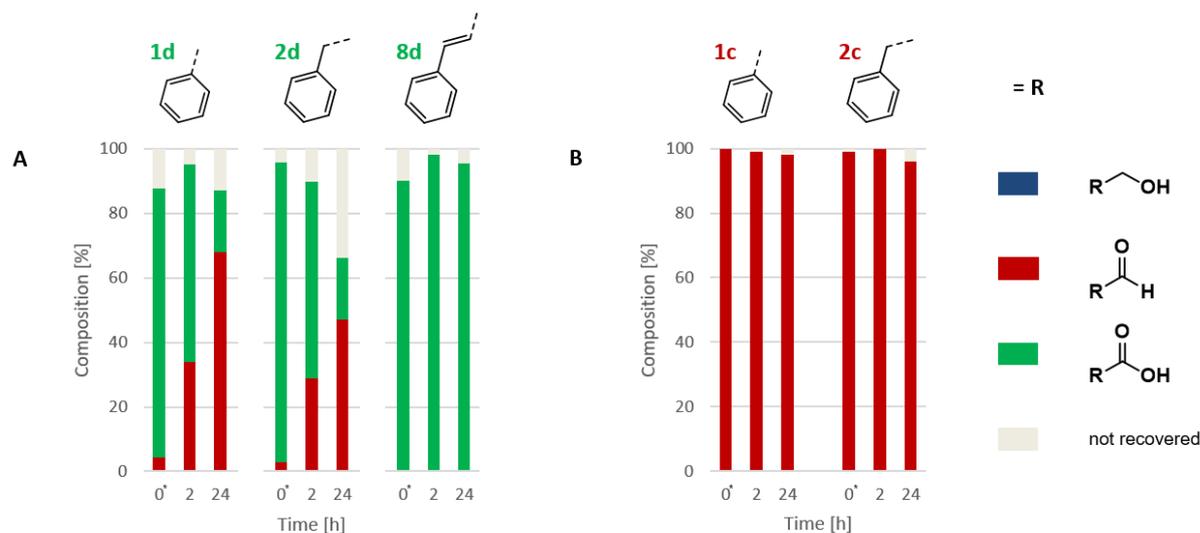
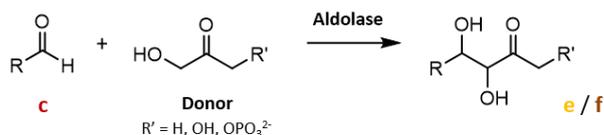


Figure D-14. *In vitro* activity of purified CAR_{Ni} . **(A)** The purified CAR enzyme reduced the aromatic carboxylic acids **1–2d** but not **8d** to the corresponding aldehydes. **(B)** Aldehydes were not further converted in the presence of CAR_{Ni} and the GDH_{2xBs} for NADPH cofactor recycling. Screenings were performed as in G VI.3.1.3; t_0^* sample taken immediately after the addition of NADP^+ and the reaction followed by calibrated GC/FID. Results presented as mean values of duplicates. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey).

Noteworthy, Finnigan *et al.* showed that CAR_{Ni} poorly converted **8d** based on a photometric NADPH consumption assay ($\lambda = 340 \text{ nm}$). Additionally, product inhibitors (e.g., PP_i) were determined for a CAR enzyme from *Mycobacterium phlei* and suggested for other CARs as well.^[64] Potential inhibition of CAR_{Ni} by carboxylic acid substrates and/or aldehyde products are possible and might be also the case for **8d**. However, enzyme inhibition was not further studied in this thesis.

In summary, CAR_{Ni} was confirmed as a versatile biocatalyst for the selective reduction of aromatic carboxylic acids. Since these carboxylates were formed by the overoxidation of cascade intermediates, CAR_{Ni} displays reversing enzymatic activity to reroute the carbon flux from these undesired byproducts to the target aldehydes.^[348] Importantly, *de novo* pathways can also utilize carboxylic acids as valuable starting materials, adding flexibility to convertible substrates in regards to costs and availability.^[348]

D II.4 Carbon framework expansion from aldehydes: Characterization and selection of aldolases



As already discussed, stereoselective C–C bond forming reactions are pivotal in organic synthetic chemistry to both construct and extend carbon frameworks.^[428] Toward the synthesis of chiral polyhydroxylated compounds, biocatalytic retrosynthesis suggested the aldol addition of a donor carbonyl compound to aldehyde acceptors produced by *AlkJ* in the previous cascade step (**Figure D-1**).

This work primarily focused on the implementation of the mutant D-fructose-6-phosphate aldolase of *E. coli*, *Fsa1-A129S*.^[348, 435] The wild type *Fsa1* as well as the engineered aldolase accept different aldehyde acceptors,^[434-435] which makes them suitable biocatalysts to target polyhydroxylated compounds and sugar derivatives. Furthermore, *Fsa1-A129S* was shown to display improved catalytic efficiency toward DHA as the donor substrate in aldol reactions and has been widely applied for synthetic applications *in vitro*.^[432] To transfer this striking advantages *in vivo*, *E. coli* was engineered and transformed into a state-of-the-art microbial cell factory to produce polyhydroxylated compounds.

The aldolases *FruA*, *FucA*, and *RhuA*, have been demonstrated to produce target aldols *in vitro*. All three aldolases display the typically relaxed substrate scope for acceptor aldehydes, but strictly depend on DHAP as the donor. Enzyme cascades involving DHAP-dependent aldolases are generally low yielding, which can be attributed to the lability and the limited availability of DHAP *in vitro* and *in vivo*, respectively. To overcome this bottleneck *in vitro*, DHAP was generated *in situ* by different strategies as discussed earlier (e.g., TIM or *DhaK*). The group of Wever established an alternative route to produce DHAP by phosphorylation of DHA in the presence of PP_i by an acidic phosphatase mutant from *Salmonella enterica* ser. *typhimurium* LT2 (*PhoN-Se V78L*). Subsequently, DHAP was linked to an aldehyde in a concurrent aldolase-mediated condensation reaction to the corresponding aldol adduct. Finally, selective dephosphorylation was catalyzed by the already present *PhoN-Se V78L*.^[484] Importantly, the phosphatase did not show phosphorylation activity for the dephosphorylated aldol. Hence, the irreversibility of this reaction shifted the equilibrium toward the desired aldol products and could also be transferred into a continuous flow reactor system.^[485]

The implementation of a phosphatase was also envisioned by Wei *et al.*, who recently published the first *in vivo* cascade involving the DHAP-dependent aldolases *FruA*, *FucA*, and *RhuA*.^[293] DHAP was provided by the central carbon metabolism of the host and hijacked by aldolase-mediated cascade reaction. The coexpressed phosphatase *YqaB* led to the dephosphorylated aldol adducts (**Figure C-41**). *In vivo* bottlenecks included the poor solubility of different aldolase acceptor aldehydes and was circumvented by using the water-soluble aldehyde 3-trifluoroacetamido propanal, for example. Retroaldol reaction was prevented by a high excess of acceptor aldehydes (100 mM). Although the issues of aldehyde toxicity were not addressed by Wei *et al.*, the synthesis of

several iminosugars in moderate yields of 35% and good diastereoselectivity ($dr = 90:10$) demonstrated the applicability of the metabolic pathway.

To expand pathway complexity *in vivo* and to circumvent artificial donor addition (i.e., hijacking DHAP from the central metabolism), this work also focused on the construction of *de novo* pathways featuring DHAP-dependent aldolases. By implementing the DHAP-dependent aldolases *FucA* and *RhuA*, the stereoconfigurations (3*R*,4*R*) and (3*R*,4*S*) can be obtained, respectively, besides (3*S*,4*R*) produced by *FruA* as well as *Fsa1-A129S* (Figure C-39).^[293, 415] *TagA* yielding the (3*S*,4*S*) diastereomere was not available and, hence, not studied in this thesis. Although employed by the group of Wang,^[293] the maximization of aldol titers of DHAP-dependent aldolase cascades *in vivo* remains an unmet challenge and will be addressed in this thesis.

D II.4.1 Expression of *Fsa1* and *Fsa1-A129S* and *in vitro* characterization of the engineered aldolase

The *fsa1* and the *fsa1-A129S* genes^[432] were subcloned into pET16b and donated by the group of Fessner. Best soluble expression for both aldolases was achieved after transformation of *E. coli* BL21(DE3) in TB medium upon the addition of 0.5 mM IPTG and cultivation at 30°C (200 rpm) for 20 h (Figure D-15).

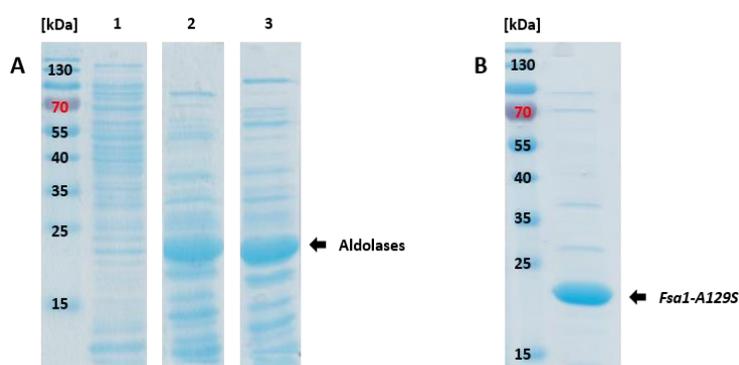


Figure D-15. Expression of (D)HA-utilizing aldolases. (A) SDS-PAGE analysis of CFEs of untransformed *E. coli* BL21(DE3) (1), CFEs containing *Fsa1* (2) and *Fsa1-A129S* (3). Sample loading normalized to $OD_{590} = 7.0$. (B) Semi-purified *Fsa1-A129S* by HS. Sample loading: 10 μ g of total protein according to Bradford assay.

Transformation of *E. coli* JM109(DE3), followed by enzyme expression, did not boost the production of soluble target proteins (data not shown).^[486] Cell lysis was performed in 50 mM glycylglycine (GlyGly) buffer (pH 8.0) in the presence of lysozyme, the aldolases purified by HS (Figure D-15B), and lyophilisates prepared according to G I.15.1. Activity was tested *in vitro* by the aldol addition of the model aldehyde **2c** and DHA (Figure D-16).

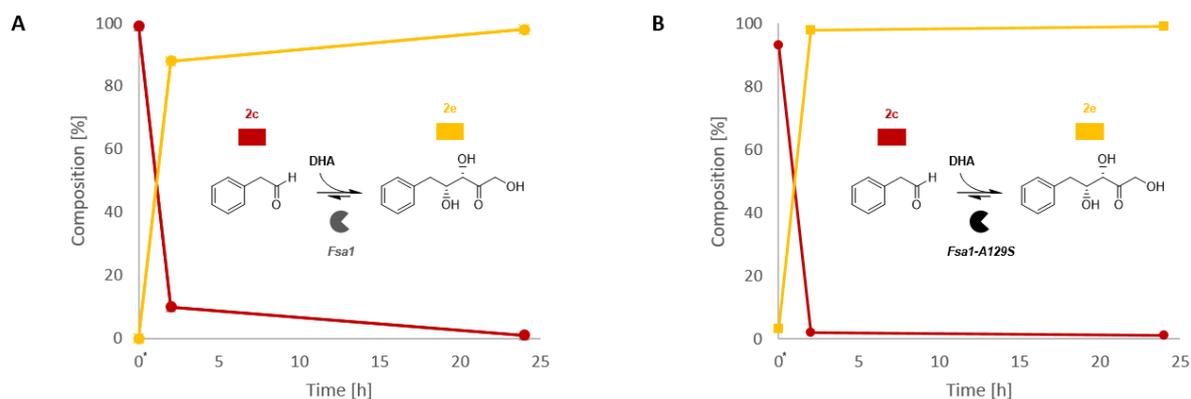


Figure D-16. *In vitro* activity of semi-purified DHA-utilizing aldolases. (A) *Fsa1* fully consumed the aldehyde **2c** (red) to produce the target aldol **2e** (yellow). (B) The mutant aldolase *Fsa1-A129S* fully consumed **2c** in only 2 h and efficiently produced the target aldol **2e**. Screening carried out as in G VI.4.2.3; t_0^* sample taken immediately after the addition of **2c** and the reaction followed by calibrated HPLC with recoveries >95%. Results presented as mean values of duplicates.

Both the wild type *Fsa1* and the *Fsa1-A129S* mutant accepted DHA as donor and the model aldehyde **2c** as acceptor for aldol addition *in vitro*. The higher activity toward DHA was reflected by the slightly slower consumption of **2c** after 2 h reaction time (**Figure D-16**). The high excess of DHA shifted the aldol reaction to completion at longer reaction times by prevention of retro-aldol reaction. The same trend was observed for the donor molecule HA, which gave access to **2f** and related aldol adducts (**Table E-2**).^[487] For subsequent *de novo* pathway assembly and characterization, *Fsa1-A129S* was further studied in course of this work.^[348, 435]

D II.4.2 Cloning, expression, and *in vitro* characterization of DHAP-dependent aldolases

The aldolase genes *fruA* from *E. coli*, the *fucA* and *rhuA* from *E. coli* were subcloned into pKK223-3 vectors (referred to as pKK in the following) and kindly provided by the group of Fessner.^[486] Competent *E. coli* BL21(DE3) cells were transformed and the aldolases successfully expressed following the protocol in G VI.4.1.1 (**Figure D-17**).

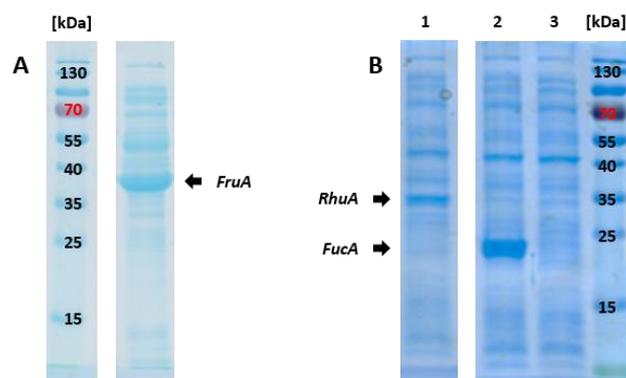


Figure D-17. Expression of DHAP-dependent aldolases. SDS-PAGE analysis of CFEs containing (A) *FruA*; (B) *RhuA* (1) and *FucA* (2); CFE from untransformed *E. coli* BL21(DE3) (3) for comparison. Aldolases expressed from pKK vectors (G VI). Sample loading normalized to 10 μ g total amount of protein per lane.

Initial studies *in vitro* were performed by T. Wiesinger with *FucA* and *RhuA* to access (3*R*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e_{FucA}**) and (3*R*,4*S*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e_{RhuA}**), respectively, which are

complementary to previously synthesized **2e** with (3*S*,4*R*) configuration (**Figure D-16**). Relative configurations were confirmed by ¹³C- and ³¹P-NMR by T. Wiesinger as well (data not shown).^[487]

Biotransformations were performed in duplicates either in a one pot fashion by simultaneously adding CFEs containing *FucA* (or *RhuA*) and a phosphatase from *Shigella flexneri* (*PhoN-Sf*) or by adding the aldolase and *PhoN-Sf* sequentially (**Figure D-18**). The characterization and selection of phosphatases (e.g., *PhoN-Sf*) will be discussed in the next section (D II.5). *In vitro* reactions were carried out with CFEs containing *FucA* and *PhoN-Sf* and the model aldehyde **2c**. The addition DHAP started the reaction. The experimental set up is outlined in G VI.4.1.3. The consumption of **2c** was followed by calibrated GC/FID, whereas the formation of the target aldol **2e_{FucA}** was monitored by calibrated HPLC (**Figure D-18**). Proteins were precipitated prior to HPLC measurement as before.

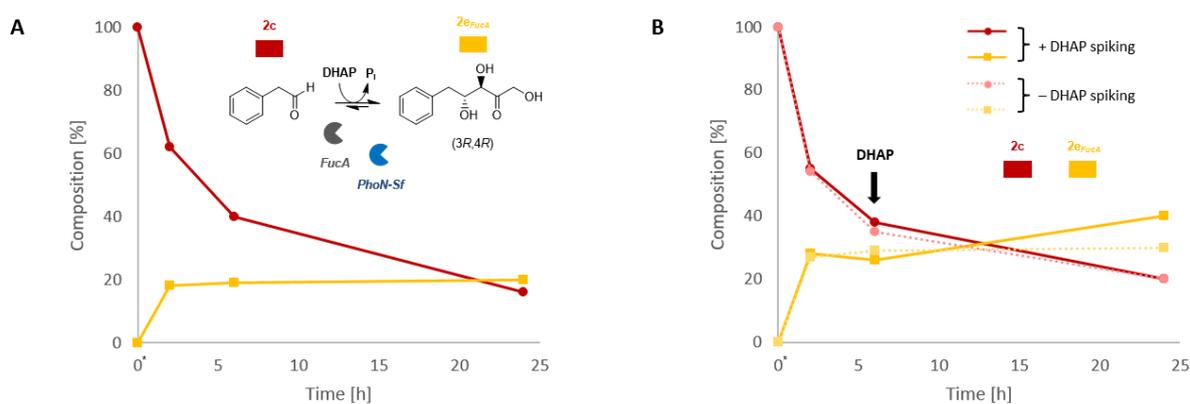


Figure D-18. *In vitro* activity of *FucA*. (A) In the presence of *FucA* and *PhoN-Sf*, aldehyde **2c** (red) was consumed and the target aldol **2e_{FucA}** (yellow) produced. (B) Repeated experiment with spiking of 1 mM DHAP after 6 h reaction time (full lines) and without spiking of DHAP (dashed lines). Spiking slightly increased the yield of **2e_{FucA}**. Aldol production performed as in G VI.4.1.3 with 1.7 eq **2c** (5 mM), and DHAP (A: 1.0 eq, 3 mM; B: 1.3 eq, 4 mM). Reactions followed by calibrated HPLC. Results presented as mean values of duplicates. The figure was adapted from T. Wiesinger (2017).^[487]

In the presence of *FucA* and *PhoN-Sf*, **2c** and DHAP were successfully converted to the target **2e_{FucA}** after dephosphorylation (*dr* = 70:30).^[487] However, **2e_{FucA}** yields plateaued between 20–30% after 2 h (**Figure D-18**) and could be increased by 10% upon the addition of 1 mM DHAP after 6 h reaction time (**Figure D-18B**). A similar experiment employing the addition of 1 mM **2c** after 6 h did not shift the equilibrium toward the desired aldol adduct (data not shown). This indicates that DHAP hydrolyzes under experimental conditions by *PhoN-Sf*, for example, or that *FucA* is inactivated by product inhibition. Methylglyoxalate, a decomposition product of DHAP, and inorganic phosphates are known strong inhibitors for aldolases;^[431, 488-489] the latter accumulate due to the dephosphorylating activity of *PhoN-Sf* in the reaction mixture (**Figure D-18A**). Nonetheless, these bottlenecks might not even occur *in vivo* since DHAP is rather stable in the cellular environment, intracellular DHAP levels are adjusted by host cell responses, and inorganic phosphates recycled.^[293] Substrate inhibition by **2c** could be excluded *in vitro* since experiments at lower substrate loading (1 mM) yielded similar amounts of **2e_{FucA}** as demonstrated by T. Wiesinger.^[487]

FucA was not only shown to retain full activity in the presence of different organic solvents (e.g., ACN, DMF),^[487] regarding the enzymatic synthesis of reference aldol compounds and the solubility of aldehyde acceptors, *FucA* was the DHAP-dependent aldolase of choice for subsequent *in vivo* pathway construction. In this context, the *fucA* gene was cloned into pKA1, which features a P15A ORI compatible with standard pET vectors, for example, that feature a ColE1 ORI (**Table C-1**).^[6] Subcloning employed an adapted FC procedure (G II.2.1) and is described in detail in G VI.4.1.2. *In vivo* assembly depended on matching homologous overhangs (29–34 bp homology)

introduced during PCR amplification of the *fucA* insert and the pKA1 backbone. Successful assembly was verified by *NcoI* restriction enzyme digestion of plasmid DNA from putative positive clones (Figure G-18B), and the integrity of the *fucA* sequence finally confirmed by Sanger sequencing (H I.1.5.3). Competent *E. coli* BL21(DE3) was transformed with pKA1_*fucA* (clone: BWA4) and the production of *FucA* was compared to cells harboring the parent pKK_*fucA* plasmid (Figure D-19A).

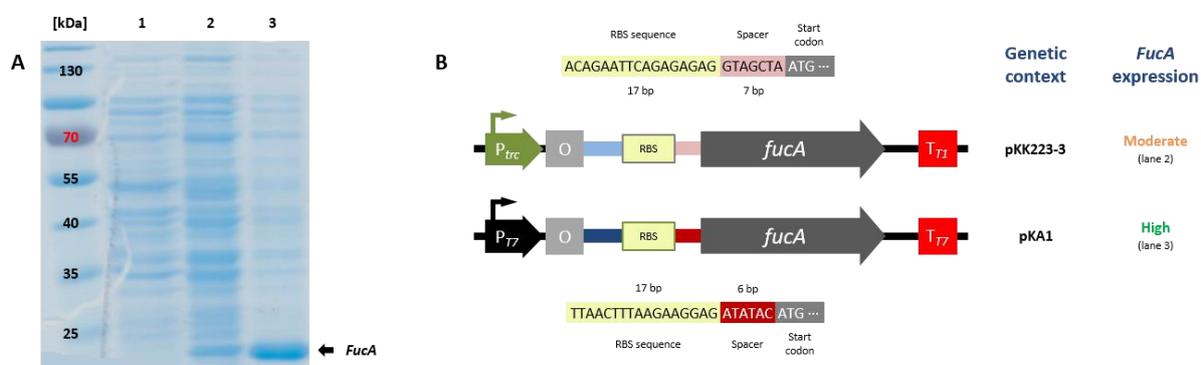


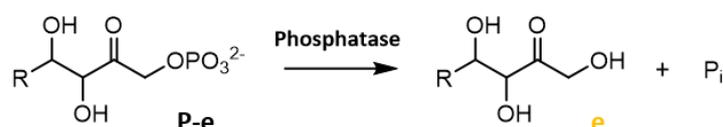
Figure D-19. Context dependency influences *FucA* expression. (A) SDS-PAGE analysis of CFEs from untransformed *E. coli* BL21(DE3) (1) and transformants expressing *FucA* from the parent pKK plasmid (2) and the newly assembled pKA1 vector (3). Protein production according to G VI. Sample loading normalized to OD₅₉₀ = 7.0. (B) Comparison of the genetic context in pKK_*fucA* (top) and pKA1_*fucA* (bottom) around the *fucA* gene. Transcription is influenced by promoters (P), operators (O), and terminators (T); translation efficiency is tuned by the RBS, the spacer sequence (red) between the RBS and the ATG start codon, and the sequence upstream of the RBS (shades of blue).^[6]

Cells harboring the newly assembled pKA1_*fucA* plasmid, produced significantly higher amounts of soluble *FucA* compared to pKK_*fucA* transformants under optimized conditions for single enzyme expression (Figure D-19A; see also G IV). *FucA* production was also increased under different expression conditions including cultivation media (e.g., M9-N* or LB-Miller medium), reduced expression temperature (20°C), and different modes of induction (e.g., autoinduction). Protein production is not only influenced by expression conditions, the genetic context adjacent to the gene to be expressed can have profound influence.^[6, 490] Analysis of the genetic context around the inserted *fucA* gene revealed that the RBS in pKA1_*fucA* (TTAACTTTAAGAAGGAG) was optimized for the translation in *E. coli* BL21(DE3) and other T7 expression systems (Figure D-19B). The RBS was followed by an AT-rich spacer sequence (6 bp). In contrast, the RBS in pKK_*fucA* (ACAGAATTCAGAGAGAG) was different from the optimized one in the pKA1 vector, explaining the reduced amounts of *FucA* produced. Furthermore, the spacer sequence had a higher GC content and was one basepair longer, which also affects translation efficiency. Sequences upstream (not shown) and downstream (shades of blue) of the promoter can influence transcription and transcription/translation, respectively.^[6] Both systems feature different promoters but with comparable strength *in vivo* (Figure D-19B).^[491-492]

Summing up, *FucA* represented a suitable DHAP-dependent aldolase for subsequent pathway assembly *in vivo* to access the aldol **2e_{FucA}** with a (3*R*,4*R*) configuration. By coupling the aldolase reaction with the non-specific phosphatase *PhoN-Sf*, the model substrate **2c** and DHAP were linked and the intermediate dephosphorylated successfully, demonstrating the applicability of this two-step cascade (Figure D-18A). The addition of DHAP slightly boosted aldol formation, suggesting that elevated intracellular concentrations of DHAP might be beneficial for pathway applications *in vivo* (Figure D-18B).^[487] The slightly lower *dr* = 70:30 compared to a *dr* = 90:10 reported in the literature can be explained by the employment of CFEs,^[487] which contain solutes of the cytosol and host enzymes, probably, endogenous aldolases (e.g., *FruA* and *RhuA*). Consequently, natively expressed aldolases can lead to the formation of diastereomeric byproducts – not only *in vitro*.

The construction of pKA1_ *fucA* greatly enhanced the production of *FucA* under various expression conditions (Figure D-19A) and could be explained by the optimized genetic context for the expression in *E. coli* BL21(DE3) (Figure D-19B). Context dependency influences the transcription and translation of single genes and can be used to optimize whole metabolic pathways by tuning genetic regulatory elements such as the RBS and its adjacent sequences.^[6] Furthermore, *FucA* produced from the newly assembled vector was applied in biotransformations (coupled with *PhoN-Sf*) on preparative scale to synthesize **2e_{FucA}** by T. Wiesinger.^[487]

D II.5 Last but not least: Characterization and selection of wild type and engineered phosphatases



Bacterial non-specific acidic phosphohydrolases (i.e., phosphatases; PhoNs) belong to group of enzymes secreted as soluble periplasmic proteins. Alternatively, they can be retained as membrane-associated lipoproteins. Phosphatases dephosphorylate a variety of structurally unrelated organic phosphoesters including nucleotides, sugar phosphates, and phytic acid. Hence, phosphatases are involved in the acquisition of inorganic phosphate and organic byproducts as carbon source.^[493]

Biocatalytically, phosphatases played a minor role but have gained significant attention with the introduction of DHAP-dependent aldolases for C–C coupling reactions. Enzymatic cascades employed the *FruA*, *FucA*, *RhuA*, and *TagA* to link acceptor aldehydes and DHAP and the dephosphorylating activities of phosphatases from various bacteria including *Salmonella* sp.,^[484] *Shigella flexneri* (*S. flexneri*),^[485] and *E. coli*.^[293] In this thesis, two wild type phosphatases from *S. flexnerii* (*PhoN-Sf*) and *E. coli* (*YqaB*) and the engineered *PhoN-Se V78L* were recombinantly expressed in *E. coli* BL21(DE3) and further characterized.

PhoN-Se V78L was previously engineered to increase its phosphorylating activity toward DHA. The single mutant showed increased DHAP production in the presence of PP_i and was also more active than the wild type *PhoN-Se* and *PhoN-Sf* in an aldolase cascade reaction *in vitro* (pH 6.0).^[484] Since the enhanced DHA phosphorylating activity of *PhoN-Se V78L* could be beneficial for the desired application *in vivo*, the gene was ordered from GeneArt™ and subcloned into pET26b(+) utilizing *NdeI/HindIII* restriction sites as described in G VI.5.1.1 with the improved Florida cloning procedure (G II.1). The mutant phosphatase gene was published with an N-terminal signal peptide (SP; MKSRYLVFFLPLIVA) for transport of the protein into the periplasmic space. In addition, the gene was ordered with a C-terminal 3xFLAG fusion tag (DYKDHDGDYKDHDIDYKDDDDK) for subsequent purification.^[494-495] In contrast to the original FLAG tag (DYKDDDDK),^[496] the limit of detection is greatly improved with 3xFLAG tag and has been used to perform cellular localization studies by immunofluorescence, for example. The C-terminally fusion of the 3xFLAG tag aimed at the set-up of an ultra-sensitive detection and quantification method by enzyme-linked immunosorbent assay (ELISA). However, this project line could not be followed due to time constraints and will not be further discussed in this thesis.

In the following, *PhoN-Se V78L* containing the SP and the fusion tag will be referred to as *PhoN-Se V78L_{fl}*. To study the impact of both the SP and the fusion tag on the activity of the phosphatase, primers were designed to amplify the gene without the SP (*PhoN-Se V78L_{3xFLAG}*) and without the 3xFLAG tag (*PhoN-Se V78L_{SP}*). All inserts were successfully amplified by PCR (Figure G-22) and subcloned into pET26b(+) as before. Sanger sequencing confirmed intact sequences (H I.1.6). Competent *E. coli* BL21(DE3) cells were transformed with the plasmids encoding the phosphatase variants and successfully expressed as described in G VI.5.1.1 (Figure D-20). To

determine the influence on the cellular localization of *PhoN-Se V78L* variants, proteins were released from *E. coli* in the periplasmic space by osmotic shock following the protocol used by the Wever group (Figure D-20).^[497]

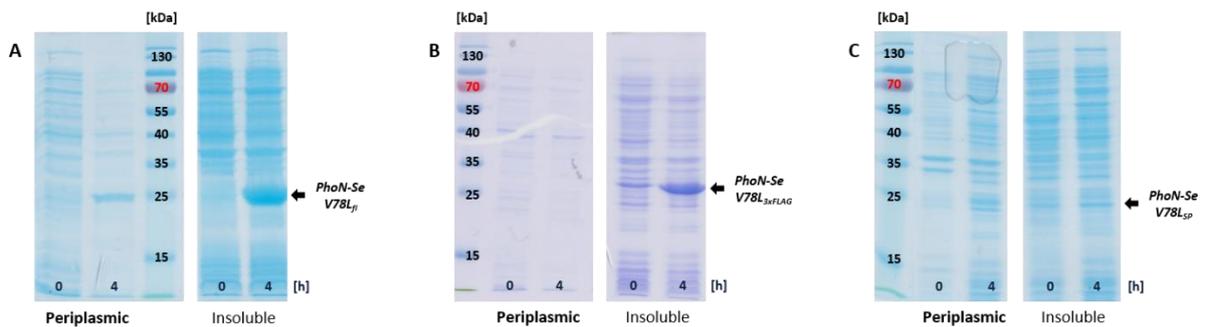


Figure D-20. Cellular localization of *PhoN-Se* variants. Phosphatase variants were released from the periplasmic space by osmotic shock before (0 h) and after induction (4 h). Periplasmic and insoluble (including cytosolic) fractions containing (A) *PhoN-Se V78L_{fl}*, (B) *PhoN-Se V78L_{3xTAG}* missing the SP, and (C) *PhoN-Se V78L_{SP}* lacking the 3xFLAG tag. Protein production: G VI.5.1.1. Osmotic shock: G VI.5.1.2. Sample loading normalized to $OD_{590} = 7.0$.

Phosphatase variants bearing the SP were transported into the periplasmic space. After release by osmotic shock, *PhoN-Se V78L_{fl}* and *PhoN-Se V78L_{SP}* could be detected in the corresponding fractions as soluble proteins according to SDS-PAGE analysis (Figure D-20A and Figure D-20C, respectively). *PhoN-Se V78L_{3xTAG}* lacking the SP was not released by osmotic shock and was exclusively found in the remaining fraction containing most cytosolic and insoluble proteins such as membrane proteins (Figure D-20B). Interestingly, variants C-terminally fused to the 3xFLAG tag showed increased expression levels (Figure D-20A–B), whereas *PhoN-Se V78L_{SP}* was only poorly expressed (Figure D-20C). As a side note, the *PhoN-Se V78L_{SP}* resembles the engineered *PhoN-Se V78L* as published by van Herk *et al.* in 2009.^[484] Phosphatase activity was confirmed for *PhoN-Se V78L_{fl}* and *PhoN-Se V78L_{SP}* by 5-bromo-4-chloro-3-indoyl phosphate (BCIP) cleavage assay (Figure D-21A). Upon cleavage of the phosphate group, 5-bromo-4-chloro-3-indoxyl is oxidized by atmospheric oxygen to form the blue dye 5,5'-dibromo-4,4'-dichloro indigo (Figure G-24).^[498] Colonies with enhanced phosphatase activity show a 'blue phenotype' on agar plates (Figure D-21). This functional screening was used by Sarikhani *et al.* to identify a non-specific acidic phosphatase in *P. putida*, for example^[499]

Commonly, phosphorylated molecules (e.g., sugar phosphates, DHAP, PP_i , or BCIP) cannot cross the envelope of *E. coli* and other bacterial cells in either direction by passive diffusion. Therefore, exclusively cells overexpressing phosphatases in the periplasmic space (e.g., *PhoN-Se V78L_{fl}*, *PhoN-Se V78L_{SP}*, and *PhoN-Sf*) gave a blue phenotype in the BCIP assay. The activity of *PhoN-Se V78L_{3xFLAG}*, which lacks the SP, is, thus, restricted to the cytosol. Colonies expressing this phosphatase variant were not stained; neither were *E. coli* BL21(DE3) cells transformed with the empty pET26b(+) vector (Figure D-21).

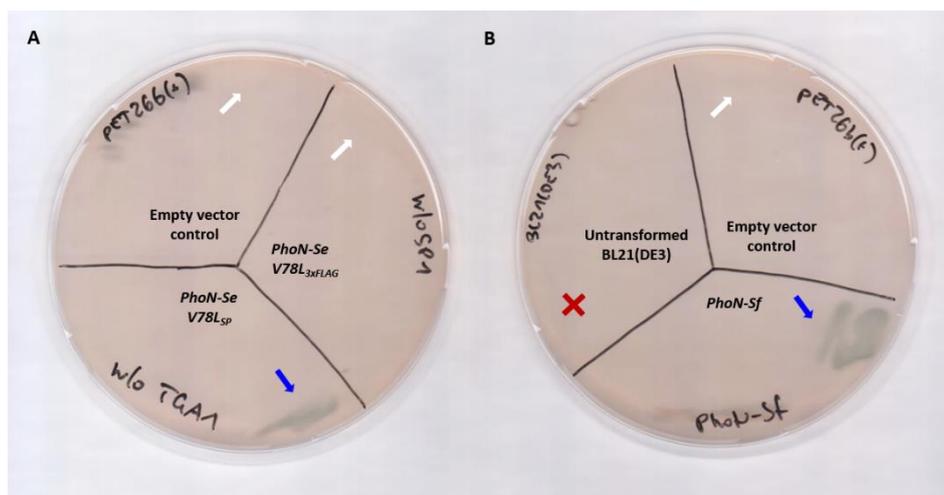


Figure D-21. Functional screening of phosphatases by BCIP assay. (A) Cells expressing *PhoN-Se V78L_{sp}* in the periplasm show a ‘blue phenotype’ (blue arrows). Colonies with cytosolic *PhoN-Se V78L_{3xFLAG}* and empty vector transformants were not stained (white arrows). (B) Cells expressing wild type *PhoN-Sf* in their periplasmic space. *E. coli* BL21(DE3) did not grow on agar plates supplemented with Kan as expected (red cross). BCIP assay according to G VI.5.4.2.^[499]

The second phosphatase characterized in this thesis was the wild type *PhoN-Sf*,^[497] which was also applied in an aldolase-coupled cascade *in vitro*.^[484-485] *PhoN-Sf* was ordered from GenScript™ in a pET26b(+) vector utilizing *NdeI/HindIII* restriction sites. The sequence was checked by Sanger sequencing after transformation of *E. coli* DH5 α and plasmid re-isolation (H I.1.6.5). *PhoN-Sf* contained an N-terminal SP (MKRQLFTLSIVGVFSLNTFA) guiding the phosphatase into the periplasmic space. The cellular localization of *PhoN-Sf* in *E. coli* BL21(DE3) transformants was confirmed by osmotic shock and SDS-PAGE analysis as before. Overall expression of *PhoN-Sf* peaked after 3 h of induction (**Figure D-22**). Amounts decreased in periplasmic fractions at longer expression times, suggesting regulatory responses by the host to the heterologous production of *PhoN-Sf* (e.g., toxicity). To circumvent the tedious enzyme preparation by osmotic shock, CFEs were prepared after cell lysis by sonication as described in G I.5. Compared to *PhoN-Se V78L_{sp}*, *PhoN-Sf* showed increased and stable expression levels even under atypical phosphatase cultivation conditions such as decreased expression temperatures $\leq 30^{\circ}\text{C}$ (**Figure D-23A**), which is advantageous since the simultaneous production of multiple pathway enzymes in the same host cell is routinely performed at lower temperatures.^[6, 43, 348] In contrast, *PhoN-Se V78L* variants very poorly express at temperatures $< 37^{\circ}\text{C}$ (**Figure D-23B**).

PhoN-Sf activity was confirmed by BCIP cleavage assay (**Figure D-21B**). Like the functional screening of esterases by the hydrolysis of pNPA, CFEs were prepared as described in G I.5 and activities toward p-nitrophenyl phosphate (pNPP) tested (**Figure D-23C**). For subsequent coupling of the DHAP-dependent aldolase reaction catalyzed by *FucaA*, *PhoN-Sf* was successfully employed *in vitro* to dephosphorylate phosphorylated aldol products yielding **2e_{Fuca}** (**Figure D-18**). *PhoN-Se* variants were not tested *in vitro* due to varying expression levels in *E. coli* BL21(DE3), especially of *PhoN-Se V78L_{sp}* (**Figure D-23B**), which is in accordance with R. Wever.^[500]

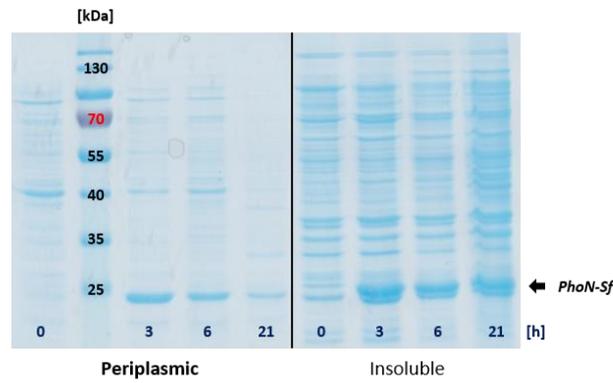


Figure D-22. Time-resolved cellular localization of *PhoN-Sf*. The phosphatase was released from the periplasmic space by osmotic shock before (0 h) and after 3–21 h after induction. The periplasmic fraction mainly contained *PhoN-Sf* with the highest expression levels at 3 h. Protein production: G VI.5.2.1. Sample loading normalized to $OD_{590} = 7.0$.

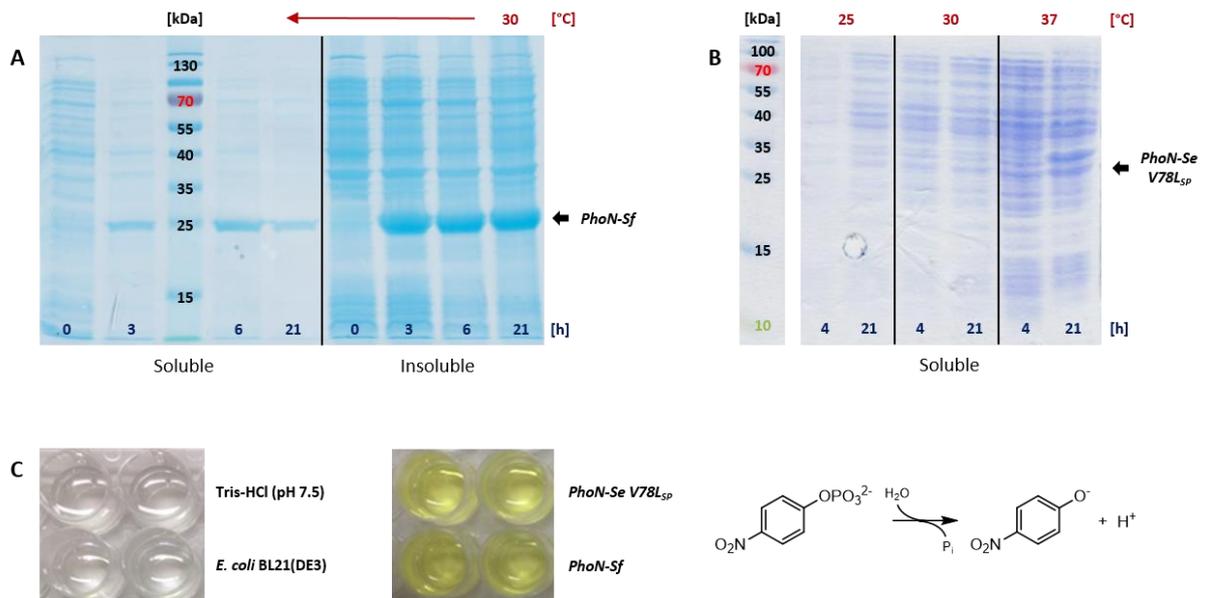


Figure D-23. Successful expression and functional testing of *PhoN-Sf*. (A) SDS-PAGE analysis of soluble and insoluble fractions containing *PhoN-Sf* after lysis by sonication. (B) For comparison, *PhoN-Se V78Lsp* CFEs were prepared accordingly. Sample loading normalized to $OD_{590} = 7.0$. (C) Photometric assay monitoring the hydrolysis of pNPP in buffer (top left), CFE from untransformed *E. coli* BL21(DE3) (bottom left), in the presence of *PhoN-Se V78Lsp* (expressed at 37°C; top right), and *PhoN-Sf* (expressed at 30°C; bottom right); P_i = phosphate. Assay performed in triplicates according to G VI.5.4.1.

The third phosphatase characterized in this thesis was the *E. coli* phosphatase *YqaB*,^[501] which was implemented in an aldolase-coupled cascade *in vivo*.^[293] *YqaB* was ordered from GenScript™ in a pCDFDuet-1 vector utilizing *NcoI/BamHI* restriction sites. The sequence was checked by Sanger sequencing after re-isolation of plasmid DNA from *E. coli* DH5α transformants (H I.1.6.6).

YqaB could be successfully produced in *E. coli* BL21(DE3) at different expression temperatures with the highest soluble amounts produced at 20°C after 21 h. Expression at low temperatures was also beneficial in terms of greatly reduced yields of target protein in insoluble fractions (Figure D-24A). Contrary to the phosphatases *PhoN-Se* and *PhoN-Sf* from *S. enterica* and *S. flexneri*, respectively, *YqaB* is a cytosolic phosphatase and, thus, does not

contain a SP and was not released by osmotic shock.^[502] The hydrolysis of pNPP confirmed the activity of CFEs containing *YqaB*. CFEs of untransformed *E. coli* BL21(DE3) only showed minimal endogenous phosphatase activity (**Figure D-24B**). CFEs were prepared as before.

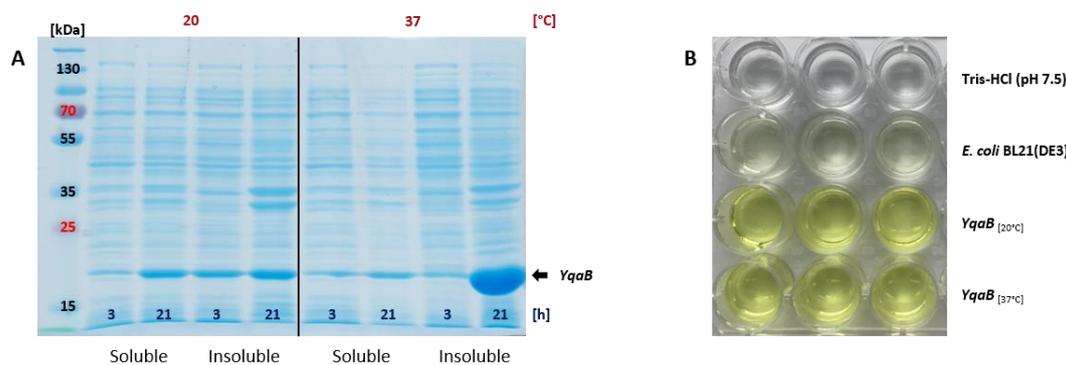


Figure D-24. Successful expression and functional testing of *YqaB*. (A) SDS-PAGE analysis of soluble and insoluble fractions containing *YqaB* after expression at different temperatures and cell lysis by sonication. Sample loading normalized to OD₅₉₀ = 7.0. (B) Photometric assay monitoring the hydrolysis of pNPP in buffer (top), CFE from untransformed *E. coli* BL21(DE3) (top center), in the presence of *YqaB* (expressed at 20°C; bottom center), and *PhoN-Sf* (expressed at 37°C; bottom). Assay performed in triplicates as before.

In summary, phosphatases from three different species were recombinantly expressed in *E. coli* BL21(DE3) and their phosphatase activity was confirmed by functional screening assays (BCIP and pNPP assay). Furthermore, the secretion of *PhoN-Se V78L* variants and the wild type *PhoN-Sf* guided by SPs was confirmed by protein fractions collected after osmotic shock (**Figure D-20** and **Figure D-22**, respectively). The export into the periplasmic space is reminiscent of the metabolic function of these phosphatases in the corresponding host organisms.^[493] However, *PhoN-Se V78L* variants were not further studied due to unsatisfying expression levels.^[500] Furthermore, DHA phosphorylation could not be reproduced *in vitro*. Biological triplicate experiments were conducted in cooperation with S. Milker. Phosphorylation activity was tested over a wide pH range (4.0–8.0) and an excess of PP_i as the phosphate donor; *PhoN-Sf* was used as reference phosphatase, which showed no increased production of DHAP either under experimental conditions (data not shown). This is in disagreement with the results reported by the Wever group.^[484]

On the other hand, CFEs containing *PhoN-Sf* were successfully used in set ups employing *FucA* (or *RhuA*) besides the phosphatase to convert the model aldehyde **2c** and DHAP into the dephosphorylated aldol product **2e_{FucA}** *in vitro* (**Figure D-18**). The *E. coli* phosphatase could be stably expressed in the cytoplasm and displayed enhanced activity toward pNPP (**Figure D-24**). Consequently, the periplasmic *PhoN-Sf* and the cytosolic *YqaB* were selected as biocatalysts for pathway assembly *in vivo*. The different cellular localization of *PhoN-Sf* and *YqaB* allowed to study the influence on the dephosphorylation of non-natural substrates *in vivo*, synthetic pathway performance, and the overall impact on the engineered system (see D III.2 and E I.6). As already mentioned, *YqaB* was previously applied in a synthetic pathway in living cells also expressing the complementary DHAP-dependent aldolases *FruA*, *FucA*, and *RhuA* to hijack DHAP from the host and to produce dephosphorylated aldol products.^[293]

D III Assembly of *de novo* pathways

D III.1 Assembly of a synthetic mini-pathway consisting of *AlkJ* and the mutant aldolase *Fsa1-A129S*

Based on biocatalytic retrosynthetic analysis, the oxidation of primary aromatic alcohols and the subsequent C–C bond forming aldol reaction with (D)HA are the key steps to form the desired aldol adducts (**Figure D-1**). The identification of *AlkJ* and *Fsa1-A129S* as suitable biocatalysts for this two-step transformation was described in the previous sections (D II.2.2 and D II.4.1, respectively). In the following, pathway assembly, introduction into the host *E. coli* BL21(DE3), and pathway validation *in vivo* will be described.

The combination of target pathway enzymes in modules has provided major advantages and has been successfully applied to construct many different (synthetic) pathways.^[43-44, 67, 75] The heterologous expression of whole metabolic pathways imposes an inherently high metabolic burden on the host and can impair growth rates, cell viability, and the flux through the artificial pathway.^[6, 435] Thus, the coproduction of multiple pathway proteins from one (or a few) expression vector not only reduces the (plasmid) burden on the host;^[6] it provides the possibility to increase cascade complexity by adding modules encoded on compatible vectors.^[43] Therefore, to reduce the metabolic burden *ex ante* and to expand pathway complexity in the future, the *alkJ* and the *fsa-A129S* gene were ought to be coexpressed from a single plasmid.

D III.1.1 *De novo* mini-pathway assembly and characterization

The pKA1 vector, which is a pACYC derivative, was already used to subclone the *alkJ* gene and functionally express it in *E. coli* BL21(DE3) cells harboring the pKA1_alkJ plasmid (**Figure D-5**). pKA1 features a P15A ORI that allows for the stable cotransformation of many routinely used vectors including pCDF, pET, and pRSF (**Table C-1**).^[6] Furthermore, pKA1 provides the optimal genetic context for P_{T7} based transcription and translation in *E. coli* BL21(DE3), JM109(DE3), and similar strains. Expression levels of *AlkJ* were sufficient for the oxidation of primary alcohols **1–2b** and **4–7b** to the corresponding carboxylic acid byproducts *via* target aldehydes **1–2c** and **4–7c** (**Figure D-6**). Subcloning of the *fucA* gene from the pKK223-3 parent into the pKA1 vector greatly increased the production of *FucA* due to the optimal genetic context (**Figure D-19**). Finally, pKA1 is a medium copy number vector bestowing a defined burden onto the host.

Consequently, SLIC methods were applied to assemble the linear *fsa1-A129S* fragment (parent: pET16b_*fsa1-A129S*) and the previously constructed pKA1_alkJ backbone (parent: pKA1_alkJ). The detailed cloning procedure is given in G VII.1 for all constructs described below. All primers were designed with long homologous overhangs (36–40 bp) to facilitate efficient recombination and assembly of the two linear fragments *in vivo*.^[87] For the amplification of the *fsa1-A129S* insert, different *fwd* primers were designed to include either the RBS only or the RBS and an individual P_{T7}, giving rise to genetic arrangements of *alkJ* and *fsa1-A129S* in operon (OPE) and pseudo-operon (POP) configuration (**Figure D-25A** and **Figure D-25B**, respectively). In a first round of cloning, the OPE plasmid (pOPE) was constructed by SLICE,^[97] assembling the pKA1_alkJ:*fsa1-A129S* construct, in which expression of the two genes is controlled by only one P_{T7} in front of the *alkJ* gene and one T_{T7} downstream of the aldolase-coding region (**Figure D-25A**).^[435] Secondly, the POP plasmid (pPOP) was constructed, assembling pKA1_alkJ::*fsa1-A129S*, in which the two genes are under the control of individual T_{T7} promoters. Both genes share a common T_{T7} (**Figure D-25B**).^[348, 435] Consecutively, two vectors with monocistronic arrangements (MON) were constructed by molecular cloning. Terminator sequences were placed downstream of the stop codon of the *alkJ* gene and upstream of the P_{T7} of the aldolase-coding region by utilization of the unique *BamHI* recognition site between the *AlkJ* and *Fsa1-A129S* coding regions. Two different bidirectional terminators were inserted:

B0011 and B00014. B0011 is a terminator derived from the *luxICDABEG* operon^[503] of *Vibrio fischeri* and forms a single hairpin loop, whereas the synthetic B0014 is a rather strong terminator consisting of two stem loops formed by B0012 and B0011 (Figure D-26).

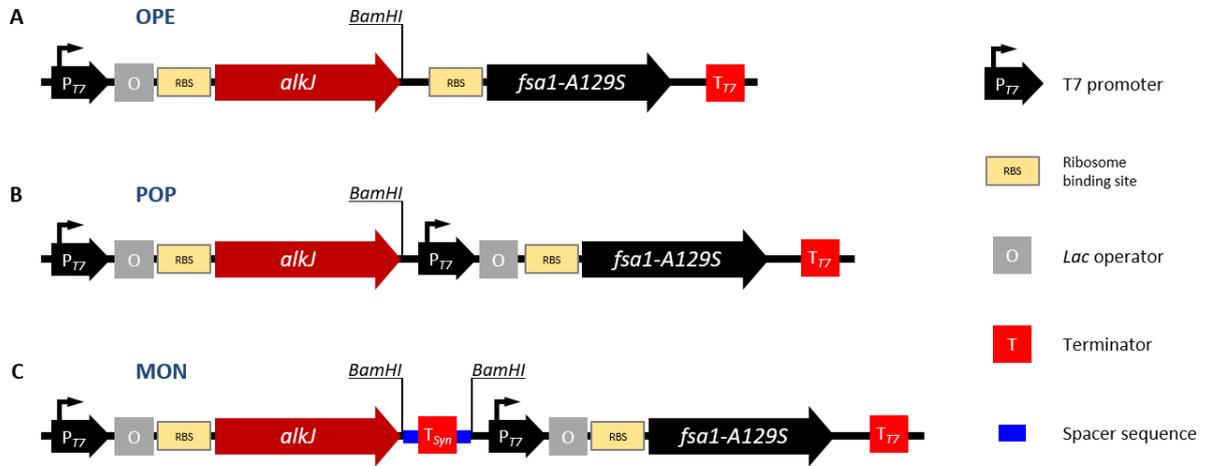


Figure D-25. Mini-pathway designs and genetic context. (A) The operon (OPE) configuration features one P_{T7} under the control of a *lac* operator sequence and one T_{T7} . (B) The pseudo-operon (POP) has an additional P_{T7} controlling *fsa1-A129S* expression. (C) Both genes feature their own P_{T7} . Synthetic terminators (T_{Syn}) are inserted at the *BamHI* restriction site and function as insulator between the two ORFs; T_{Syn} flanked by short spacer sequences on both sites.

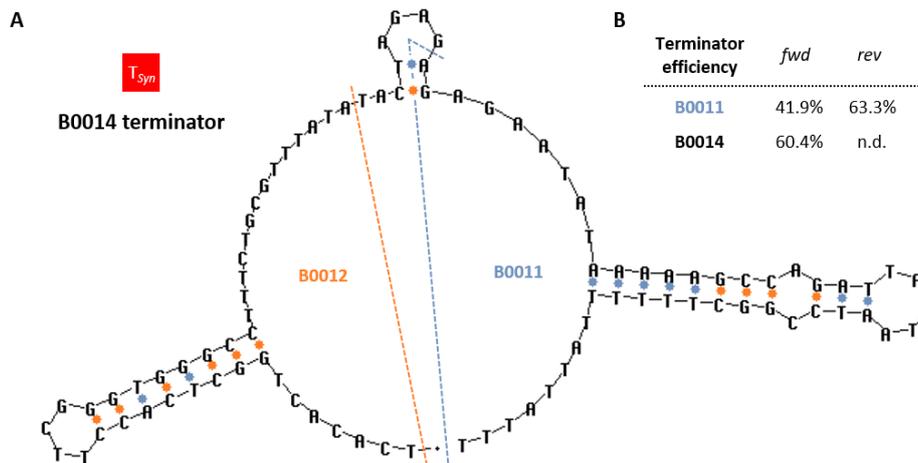


Figure D-26. B0014: a synthetic bidirectional terminator. (A) B0014 consists of two single stem loop terminator sequences, B0012 and B0011. The latter is derived from the *luxICDABEG* operon of *Vibrio fischeri*.^[504] (B) Terminator efficiencies determined by the [%] reduction in fluorescence (CFP or YFP) in cells harboring genetic constructs containing the terminator sequences in forward (*fwd*) or reverse (*rev*) orientation. Fluorescence was measured by flow cytometry. The figure was adapted from http://parts.igem.org/Part:BBa_B0014 (2003); terminator efficiency data were produced by J. Kelly (2007) and retrieved from the open access database OpenWetWare: http://www.openwetware.org/wiki/Cconboy:Terminator_Characterization/Results; n.d. = not determined.

The synthetic terminator (T_{Syn}) was designed by R. Shetty and its sequence retrieved from the open access Registry of Standard Biological Parts: http://parts.igem.org/Part:BBa_B0014.^[504] The B0014 sequence was ordered from GenScript™ with short flanking spacer sequences (5'-GGCTGCTAAC-3') and *BamHI* restriction sites. B0014 was delivered in a standard pUC57 vector, which was used as PCR template (G VII.1.3). The insertion of B0011 and B0014 gave rise to the plasmids pMON1 (pKA1_alkJ::B0011::fsa1-A129S)^[435] and pMON4

(pKA1_alkJ::B0014::fsa1-A129S), respectively. Cloning of B0011 and B0014 resulted in a 26 bp and a 16 bp spacer sequence, respectively (**Figure D-25C**). The length of spacer sequences in the final constructs include the remaining *Bam*HI restriction site. The bidirectionality of the T_{Syn} was crucial to achieve transcriptional stop independent of the terminator orientation upon insertion, which is difficult to control with only one *Bam*HI restriction site in place (**Figure D-25**).

The construction of the *AlkJ/Fsa1-A129S* mini-pathway in different genetic architectures allows to balance enzyme coproduction by additional genetic regulatory elements and their impact on cellular growth and expression levels. Whereas the modifications of promoters and the RBS are well-established to tune protein production, terminators have been widely neglected.^[6, 505] Recently, Mairhofer *et al.* designed a T_{Syn} and combined it with two well-studied transcriptional terminators, T_{T1} and T_{T7} . Combination of three terminators reduced transcriptional read-through to 1%, in other words, enhanced terminator efficiency to 99% according to RNA analysis by chip-based capillary electrophoresis. The analysis of mRNA provides a more accurate tool to determine termination efficiency than coupling to the recombinant production of fluorescent proteins, for example (**Figure D-26B**). In fermentation experiments with *E. coli* HMS174(DE3), the improved termination signal led to a significant decrease in plasmid copy numbers and increased the total protein yield, thereby, enhancing the overall fermentation process.^[505]

The constructed MON plasmids in this thesis contain the target mini-pathway, in which expression should occur independently since the pathway genes, *alkJ* and *fsa1-A129S*, are controlled by individual promoters and terminators. The T_{Syn} should act as an insulator to separate the two genes contextually. As with the two-plasmid system, *E. coli* BL21(DE3) pOPE and pPOP transformants, cells harboring pMON1 and pMON4 plasmid were subject to growth studies in different media and subsequent expression studies. The yield of soluble *Fsa1-A129S* produced by the different engineered systems was determined both qualitatively and semi-quantitatively by SDS-PAGE and quantification of semi-purified aldolase after heat shock by Bradford assay.

Therefore, competent *E. coli* BL21(DE3) cells were transformed with either the two plasmids pKA1_alkJ and pET16b/*fsa1-A129S* or one of the vectors (pOPE, pPOP, pMON1, and pMON4) for the coproduction of pathway enzymes. While untransformed *E. coli* BL21(DE3) and cells carrying an additional plasmid burden did not show different growth behavior in rich medium (LB-Miller medium; **Figure D-27A**), (subtle) differences could be determined by monitoring bacterial growth in minimal medium (M9-N* medium; **Figure D-27B**).

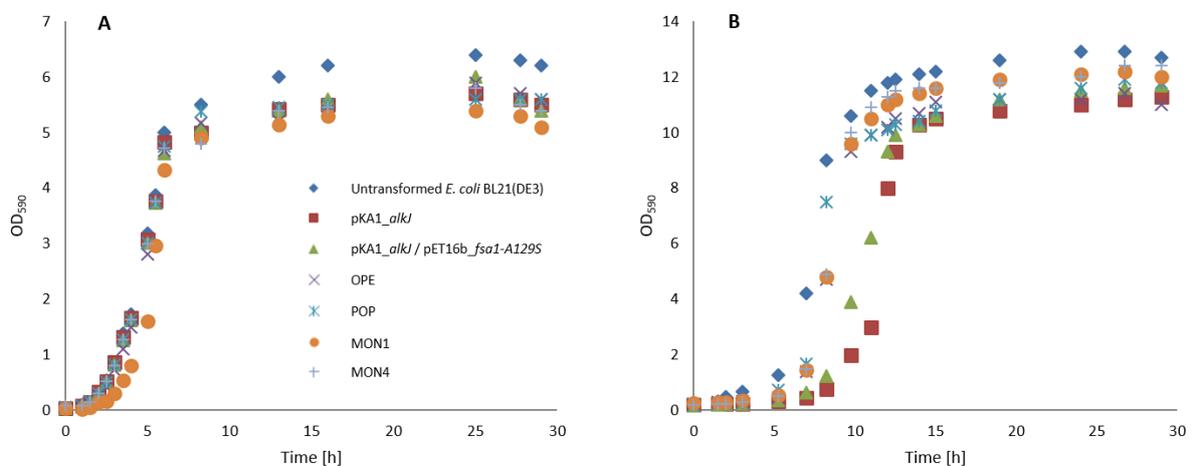


Figure D-27. Growth studies in complex and minimal medium. (A) Growth in (rich) LB-Miller medium of (engineered) *E. coli* BL21(DE3) cells. (B) Subtle growth differences in (minimal) M9-N* medium in cells harboring different genetic constructs of the *AlkJ/Fsa1-A129S* mini-pathway. Cellular growth studies were performed in duplicates. Cultivation conditions: Inoculation with 1% (v/v) preculture, 37°C (200 rpm) for 30 h.

Unburdened *E. coli* BL21(DE3) showed short initial lag phases; *E. coli* cells containing the POP plasmid grew equally fast. The two-plasmid system grew slower due to the increased plasmid burden, which is depicted by the maximal growth rate and prolonged lag phases. The same was true for pKA1_alkJ transformants (Table D-4).^[435]

Table D-4. Cellular growth and *Fsa1-A129S* expression from different genetic constructs

(Engineered) host	Growth rate [h ⁻¹]	t _(growth max) [h]	Fold-increase of soluble <i>Fsa1-A129S</i>
Untransformed <i>E. coli</i> BL21(DE3)	0.66	7.6	-
pKA1_alkJ	0.56	11.6	-
pKA1_alkJ / pET16b_fsa1-A129S	0.55	11.5	1.00
OPE	0.70	8.8	0.50 ± 0.32
POP	0.76	7.7	1.10 ± 0.12
MON1	0.66	8.8	0.94 ± 0.29
MON4	0.69	8.8	1.12 ± 0.21

Cellular growth studies were performed in baffled flasks in duplicates; cultivation in M9-N* medium after inoculation with 1% (v/v) preculture at 37°C (200 rpm). Fold-increase of soluble *Fsa1-A129S* measured by Bradford assay of semi-purified aldolase by HS; expression/HS performed in biological triplicates and presented as mean values ± SD. Fold increase normalized to [g] dry cell weight. The table was adapted from T. Wiesinger *et al.* (2017) and updated.^[435]

For further characterization, expression studies were performed with all *E. coli* BL21(DE3) transformants under the optimized conditions for the two-plasmid system in M9-N* medium as described in G VII.1.4.1. Enzyme production was monitored over time (0–20 h) by SDS-PAGE analysis (Figure D-28; for time course, see Figure G-40).

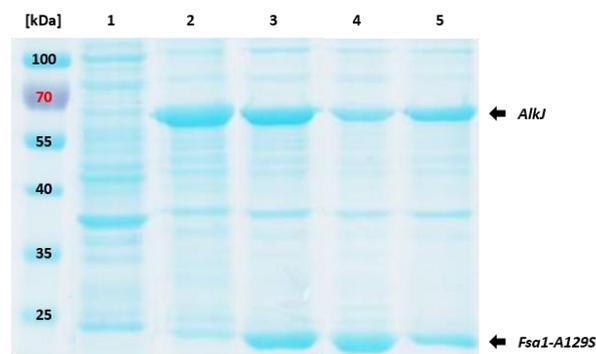


Figure D-28. Mini-pathway expression studies. SDS-PAGE analysis of whole cells of untransformed *E. coli* BL21(DE3) (1), harboring the OPE plasmid (2), POP (3), MON4 (4), and cotransformants harboring pKA1_alkJ and pET16b_fsa1-A129S. Protein production as in G VII.1.4.1. Sample loading normalized to OD₅₉₀ = 7.0. The figure was adapted from T. Wiesinger *et al.* (2017).^[435]

In all engineered *E. coli* BL21(DE3) cells, *AlkJ* was readily produced and exclusively found in insoluble fractions as it is a membrane-associated protein (Figure D-28).^[43] The production of *Fsa1-A129S* was strongly influenced by the genetic context and the different pathway architectures. The incorporation of regulatory elements (e.g., P_{T7} and T_{Syn}) lead to different expression levels of the aldolase as determined by protein quantification after the preparation of CFEs and HS purification as described in G I.15.1. The amounts of *Fsa1-A129S* produced from pOPE, pPOP, pMON1, and pMON4 were compared to the two-plasmid system consisting of pKA1_alkJ and

pET16b_ *fsa1-A129S* (Table D-4). *Fsa1-A129S* was poorly produced from pOPE simply because the distance between the P_{T7} in front of the *alkJ* gene and the aldolase coding region was too long (Figure D-28).^[435] The production of soluble *Fsa1-A129S* was improved in all constructs featuring an individual P_{T7} controlling the downstream aldolase ORF (POP, MON1 and MON4 plasmids). However, pPOP not only yielded slightly higher amounts of the target aldolase (Table D-4 and Figure D-28); based on the beneficial growth behavior (Figure D-27B and Table D-4), pPOP transformants of *E. coli* BL21(DE3) were used for subsequent mini-pathway validation (Figure D-29).^[348, 435]

D III.1.2 Successful production of polyhydroxylated compounds *in vivo*

First, pathway validations were performed with 5 mM of the model substrate **2b** and the influence of different donor concentrations of monomerized DHA (0–20 eq) tested in RCs under standard screening conditions (Figure D-29B). In the absence of the donor molecule, the cytotoxic aldehyde intermediate **2c** accumulated after 3 h reaction time. After 15 h, almost complete overoxidation to the corresponding carboxylic acid **2d**, mediated by the endogenous enzymes and, to a minor extent by *AlkJ*, was observed as expected (Figure D-29B; see also Figure D-6). At a moderate excess of DHA (5 eq = 25 mM), the desired aldol adduct **2e** was quickly produced. However, after longer reaction times, retro-aldol reaction formed the intermediate aldehyde **2c**, which was further oxidized to the carboxylic byproduct **2d**. Larger DHA excesses (10–20 eq = 50–100 mM) significantly shifted the equilibrium of the aldol reaction toward the product **2e** with yields up to 95% after 2 h reaction time. Retro-aldol reaction and subsequent carboxylate formation was suppressed in the presence of 20 eq DHA. This contrasts with the reaction with 5 eq DHA, in which the freely available aldehyde intermediate was almost fully converted to **1c**, directing the carbon flux into a dead end.^[348, 435]

Satisfyingly, the constructed mini-pathway consisting of two metabolically non-related enzymes *AlkJ* and *Fsa1-A129S* produced the target aldol **2e** (Figure D-29B).^[348] From the previous characterization of *AlkJ* in RCs, it was already known that the primary aromatic alcohols **1–9b** can freely enter host cells (Figure D-6). This was also confirmed in the context of pathway validation with the model substrate **2b** (Figure D-29B). Control experiments to determine whether the permeability of the cellular envelope of *E. coli* influences uptake of DHA and the overall pathway performance, RCs were incubated with 1% (v/v) toluene and 5 mM EDTA. Pretreatment conditions were already outlined in D II.2.2. However, higher permeability did not increase aldehyde, respectively, aldol production (Figure D-29C). This double confirms sufficient uptake of alcohol substrates such as **2b** to be efficiently converted inside the cell. Furthermore, the uptake of DHA is not limiting either, which is as expected since DHA can also be utilized as a carbon source by *E. coli*.^[506-507]

In summary, complementary (advanced) cloning techniques led to the assembly of plasmids for the coexpression of a *de novo* pathway consisting of *AlkJ* and *Fsa1-A129S*, two metabolically non-related enzymes from *P. putida* and *E. coli*. Newly constructed vectors differed in the number and the type of regulatory elements (e.g., P_{T7} and T_{syn}), which translated into improved cellular growth in minimal medium M9-N* (Figure D-27B) and increased *Fsa1-A129S* yields in *E. coli* BL21(DE3) transformants (Table D-4). After successful expression of the mini-pathway from the POP plasmid (Figure D-29A) using either M9-N* medium or autoinduction medium (AIM), pathway functionality was demonstrated by the production of the target aldol **2e** from the model substrate **2b**. The highest yield of 95% according to calibrated HPLC was achieved in the presence of the DHA donor in high excess (20 eq) in 2 h reaction time (Figure D-29B). The DHA excess pushed the equilibrium toward the aldol adduct **2e**, slowing down the retro-aldol reaction and subsequent overoxidation of the reactive aldehyde intermediate **2c** to the carboxylate byproduct **2d**.^[348, 435] Analysis of supernatants by HPLC revealed that all cascade compounds, **2b–e** and DHA, can freely pass the cell membrane. Permeabilization of the cellular envelope by EDTA and toluene did not enhance intermediate or product formation (Figure D-29C). Therefore, primary aromatic alcohols were finally confirmed as suitable substrates to synthesize polyhydroxylated compounds *in vivo*.

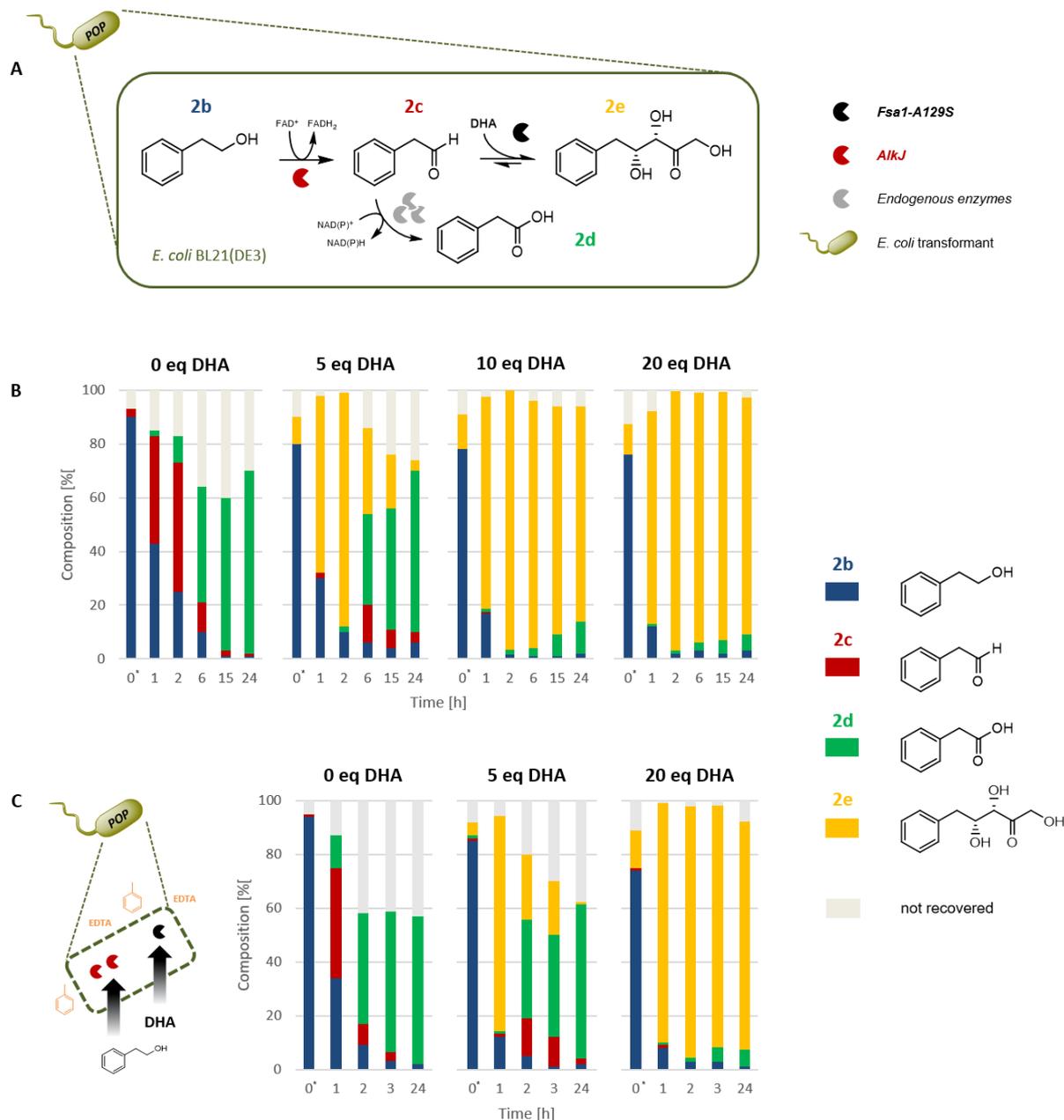
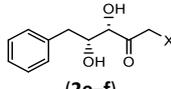
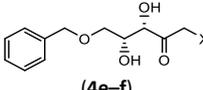
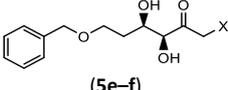
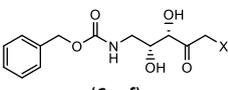


Figure D-29. Mini-pathway validation, DHA concentration and uptake screening. (A) Synthetic mini-pathway *in vivo* starting from the alcohol substrate **2b** and extracellularly added DHA to produce **2e**. Mainly endogenous host enzymes lead to the undesired **2d**. (B) *E. coli* BL21(DE3) expressing *AlkJ* and *Fsa1-A129S* pPOP produce **2e** from **2b**. Production of **2e** monitored at varying DHA concentrations (0–20 eq = 0–100 mM) over time. (C) Pretreated cells (1% (v/v) toluene, 5 mM EDTA, 4°C, 0.5 h) were used to determine if increased permeability influences substrate uptake; t_0^* samples taken immediately after the addition of 5 mM **2b**. Results presented as mean values of biological triplicates ($n = 3$); SD <10% and <5% according to calibrated GC/FID and HPLC, respectively. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), aldol (yellow), and not recovered material (grey). Reduced recovery of material at t_0^* due to insufficient mixing and a loss in mass balance (e.g., volatility) at later time points. Parts of (B) were adapted from T. Bayer *et al.* (2017).^[348]

Based on the results presented above, the genetically improved POP plasmid and the engineered *E. coli* system promised to be a useful whole cell biocatalyst to synthesize a variety of polyhydroxylated compounds and sugar derivatives.^[435] The *in vivo* system was exploited to produce a whole palette of aldol adducts from alcohols **2b**, **4–6b** *via* aldehyde acceptors **2c**, **4–6c** with DHA and HA as donor molecules yielding **2e**, **4–6e** and **2f**, **4–6f**,

respectively (**Table D-5**). Whereas *Fsa1-A129S* shows a higher activity toward DHA, it displays slightly reduced affinity for HA as the donor substrate compared to the *Fsa1* wild type enzyme.^[432] Nonetheless, as with DHA, a high excess of HA (20 eq) was used to shift the equilibrium toward the target aldols. For preparative scale experiments, the mini-pathway was expressed in POP transformants of *E. coli* BL21(DE3), which were cultivated in M9-N* medium. Protein production was induced at OD₅₉₀ = 0.5 with 0.5 mM IPTG at 25°C, 150 rpm. Biotransformations were run under optimized RC conditions (OD₅₉₀ = 10.0 in RCM, 5 mM substrate, 100 mM donor, and 5% (v/v) ACN as cosolvent; 25°C, 250 rpm).^[348, 435, 487]

Table D-5. Isolated yields of target aldol compounds without optimized product isolation

Products	Substrate	Donor	Isolated yields [%] ^[a]	
			This work ^[b]	Literature ^[c]
 (2e-f)	2b	DHA (e)	28	46 ^[432]
		HA (f)	40 ^[d]	48 ^[432]
 (4e-f)	4b	DHA (e)	37	28 ^[432]
		HA (f)	42	71 ^[432]
 (5e-f)	5b	DHA (e)	18	n.a.
		HA (f)	21	n.a.
 (6e-f)	6b	DHA (e)	35	79 ^[492]
		HA (f)	32	n.a.

^[a] Isolated compounds contain 10% water based on ¹H-NMR experiments in MeOH-*d*₄ and DMSO-*d*₆.^[435, 487]
^[b] Reactions were quenched with MeOH, centrifuged, the supernatants concentrated, and dissolved in MeOH. Product solubility was monitored by TLC, purification done by preparative HPLC.^[487] ^[c] Isolated yields from *in vitro* preparations employing *Fsa1-A129S* lyophilisates^[444] or *HL-ADH*, *NOX*, and *Fsa1-A129S*.^[508]
^[d] Isolation by centrifugation, extraction of the supernatant with CH₂Cl₂ or EtOAc, and purification by preparative HPLC.^[487] X = OH (e) or H (f); n.a. = not available. Isolated yields adapted from T. Wiesinger *et al.* (2017).^[435]

Although all substrates were converted to the corresponding aldol adducts in up to 95% yield according to calibrated HPLC analysis, the isolated yields of the two-step aldol production *in vivo* were inferior to the isolated yields of single-step transformations *in vitro* (**Table D-5**).^[435, 444, 508] Acyclic aldol compounds are temperature sensitive and highly water soluble.^[435] Hence, their isolation and purification are particularly challenging and limit yields.^[293, 436, 444] Consequently, to translate the excellent yields of the *de novo* pathway as determined by complementary GC and HPLC analysis into isolated yields, the downstream processing and product isolation were optimized. This led to the establishment of an easy to apply solid phase extraction (SPE) protocol, which was investigated by T. Wiesinger in detail.^[487] Application of the SPE purification achieved isolated aldol yields of up to 91% (**Table E-2**)^[435] and will be briefly discussed in E 1.5 as one strategy to optimize synthetic pathway performance.

Apart from the challenging purification of polyhydroxylated compounds, retro-aldol reaction is a main issue of the application of aldolases *in vivo*. *In vitro* set ups often shift the equilibrium toward target products by simply increasing one of the substrate concentrations. One-step aldol reactions can either increase the aldehyde acceptor or the donor concentrations based on cost and availability. The high excess of one reactant pushes the

equilibrium in the direction of the concentration gradient (i.e., toward the product), conveniently, preventing retro-aldol formation. However, the cytotoxicity of aldehydes does not allow high concentrations to push the equilibrium. This issue could be overcome by the *in situ* production of aldehydes from primary aromatic alcohols by *AlkJ*. As demonstrated above, high excesses of the donor molecule DHA resulted in efficient production of aldols. Nonetheless, this approach is not atom efficient in terms of incorporation of DHA into the target compound. Additionally, excess DHA concentrations probably decrease at prolonged reaction times since DHA can be metabolized by *E. coli* and serve as additional carbon source.^[506] Retro-aldol reaction is profoundly accelerated in the cellular host environment because of several reasons:

The products of retro-aldol reaction are the (cytotoxic) aldehyde and DHA. Aldehydes are rapidly reduced to the corresponding alcohols *in vivo*. In the presence of *AlkJ* oxidizing primary alcohols, endogenous AIDHs detoxify reactive aldehyde species by the irreversible oxidation to the corresponding carboxylic acids. This competing side reaction withdraws cascade intermediates from the equilibrium, directing the *de novo* pathway flux into a dead end, especially at low DHA donor concentrations (**Figure D-29B**).^[348] Solutions to this bottleneck will be discussed in chapter E.

Furthermore, complementary strategies from genetic engineering and synthetic biology were used in this thesis to optimize the presented mini-pathway (and other enzyme cascades) and will also be discussed in chapter E.

D III.2 Assembly of a *de novo* pathway featuring *AlkJ*, the DHAP-dependent aldolase *FucA*, and *PhoN-Sf*

Since *FucA* retained enzymatic activity in the presence of different organic solvents (e.g., ACN) that facilitate the solubility of substrates and the DHAP-dependent aldolase and the phosphatase *PhoN-Sf* were shown to produce the aldol **2e_{FucA}** *in vitro* (**Figure D-18**), the two enzymes were chosen for pathway construction and to access the (3*R*,4*R*)-aldol product (**2e_{FucA}**).

In assemblies of vectors for the coproduction of *AlkJ* and *Fsa1-A129S*, the arrangement of both genes in a pseudo-operon was beneficial for mini-pathway expression in *E. coli* BL21(DE3) host cells.^[435] Hence, the construction of two target pKA1 backbone vectors containing the *alkJ* and the *fucA* genes in pseudo-operon configuration and, ultimately, the modular pathway assembly will be described in the following.

D III.2.1 Synthetic DHAP-dependent aldolase pathway assembly and characterization

The first 'enzymatic core module' to be assembled was pKA1_alkJ::fucA (**Figure D-31**). Construction was performed by an adapted FC procedure employed by J. Reiterlehner.^[3] Primers contained 26–31 bp homologous overhangs for directed assembly *in vivo*. The *fucA* insert was amplified using the pKK223-3_3_fucA template. The primers (POPFUC1A fwd and POPFUC1A rev) specifically amplified the aldolase coding region (G VII.2.1). Regulatory elements including the P_{T7} and the RBS were encoded on pKA1_alkJ::fsa1-A129S (pPOP), which served as PCR template to amplify the pKA1 backbone including the *alkJ* gene. This cloning strategy simply replaced the *fsa1-A129S* gene by the *fucA* gene and the detailed FC procedure is given in G VII.2.1. Assembly was confirmed by *BglI* control digestion yielding a single 7.7 kb fragment for the target pKA1_alkJ::fucA vector (**Figure D-30A**), whereas the parent pPOP contained an additional *BglI* restriction site and gave to fragments of 4.3 kb and 3.4 kb (**Figure D-30A**, lane 6). The integrity of the *fucA* sequence in pKA1_alkJ::fucA (clone #1) was finally confirmed by Sanger sequencing. Chemically competent *E. coli* BL21(DE3) cells were transformed with the newly constructed plasmid to perform enzyme expression studies and, subsequently, biotransformations to confirm enzymatic activities.

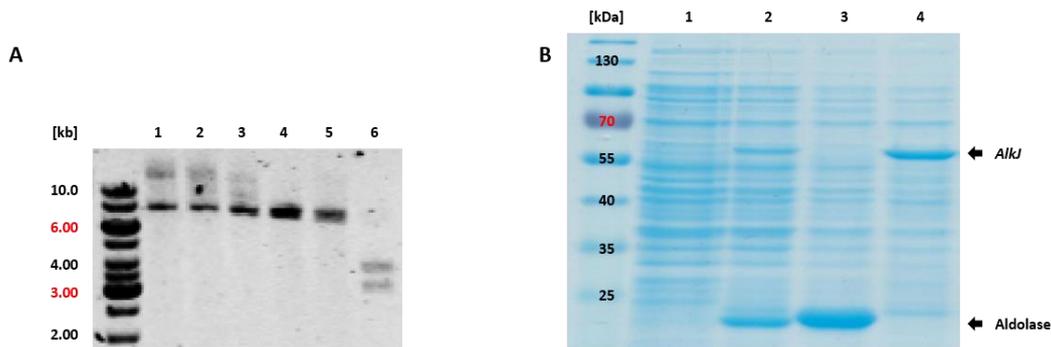


Figure D-30. Assembly controls for pKA1_alkJ::fucA. (A) *Bgl*II control digestion of correctly assembled plasmids re-isolated after FC (1–5) and the parent pKA1_alkJ::fucA (pPOP; 6). (B) SDS-PAGE analysis of whole cell samples of untransformed *E. coli* BL21(DE3) (1), pPOP transformants coexpressing *AlkJ* and *Fsa1-A129S* (2), pKA1_fucA transformants expressing *FucA* (3), and cells harboring the newly assembled plasmid (4). Controls were as expected (1–3) but *FucA* was not produced from pKA1_alkJ::fucA (4). Sample loading normalized to $O_{D590} = 7.0$. The figure was adapted from J. Reiterlehner (2017).^[3]

Protein production was performed in AIM for 24 h adapted from G VI.3.1.1. Whereas control expressions of *FucA* from pKA1_fucA and coproduction of *AlkJ* and *Fsa1-A129S* from pPOP were successful, the newly assembled pKA1_alkJ::fucA exclusively overexpressed *AlkJ* (Figure D-30B). As noted, Sanger sequencing confirmed both the presence and the integrity of the inserted *fucA* gene. *In silico* alignments of the gene sequences in the parent pKK223-3_fucA, the previously assembled pKA1_fucA, and the target pKA1_alkJ::fucA resulted in 100% sequence identity for the gene of interest. However, careful analysis of the genetic context revealed an unintended insertion of 7 bp into the spacer region between the RBS and the ATG start codon of the *fucA* gene in pKA1_alkJ::fucA (Figure D-31B). The insertion also contained an ATG start codon optimally positioned to the RBS. The following four nucleotides caused a frame shift resulting in a premature stop codon and, ultimately, an untranslated ORF (Figure D-31B).^[3]

To reinstall the optimal contextuality for *fucA* gene expression, deletion of undesired nucleotides was endeavored by J. Reiterlehner and the application of the Q5® Site-Directed Mutagenesis Kit from NEB.^[3] Briefly, sense and antisense primers were designed by the free NEBaseChanger® online tool. The optimal annealing temperature ($T_a = 56^\circ\text{C}$) was calculated by the NEBaseChanger® and the PCR was performed accordingly. Subsequent kinase-ligase-*DpnI* (KLD) reaction was followed by transformation of competent DH5 α cells. Plasmid DNA from five putative positive clones post mutagenesis were isolated and sent for sequencing. Sequencing indicated the complete deletion of the *alkJ* coding region in all clones and SDS-PAGE analysis confirmed the sequencing result (Figure D-32). One explanation for the deletion of the whole ORF is the identical sequences upstream of the *alkJ* and the *fucA* gene. The regulatory and adjacent sequences assembled on pKA1_alkJ::fucA originated from a pACYC-derived vector (parent: pPOP = pKA1_alkJ::fucA) and pET16b (parent: pET16b_fsa1-A129S). Both vectors are optimized for the T7 RNA polymerase based expression and translation in host cells with the (DE3) genotype, hence, providing a very similar genetic environment. As a result, the primers designed by the NEBaseChanger® were not specific enough to differ between the spacer regions in front of the two genes. Ultimately, the sequence identity renders primer design, be it for another round of SLIC with the same templates or Q5 mutagenesis, very difficult.^[3, 6]

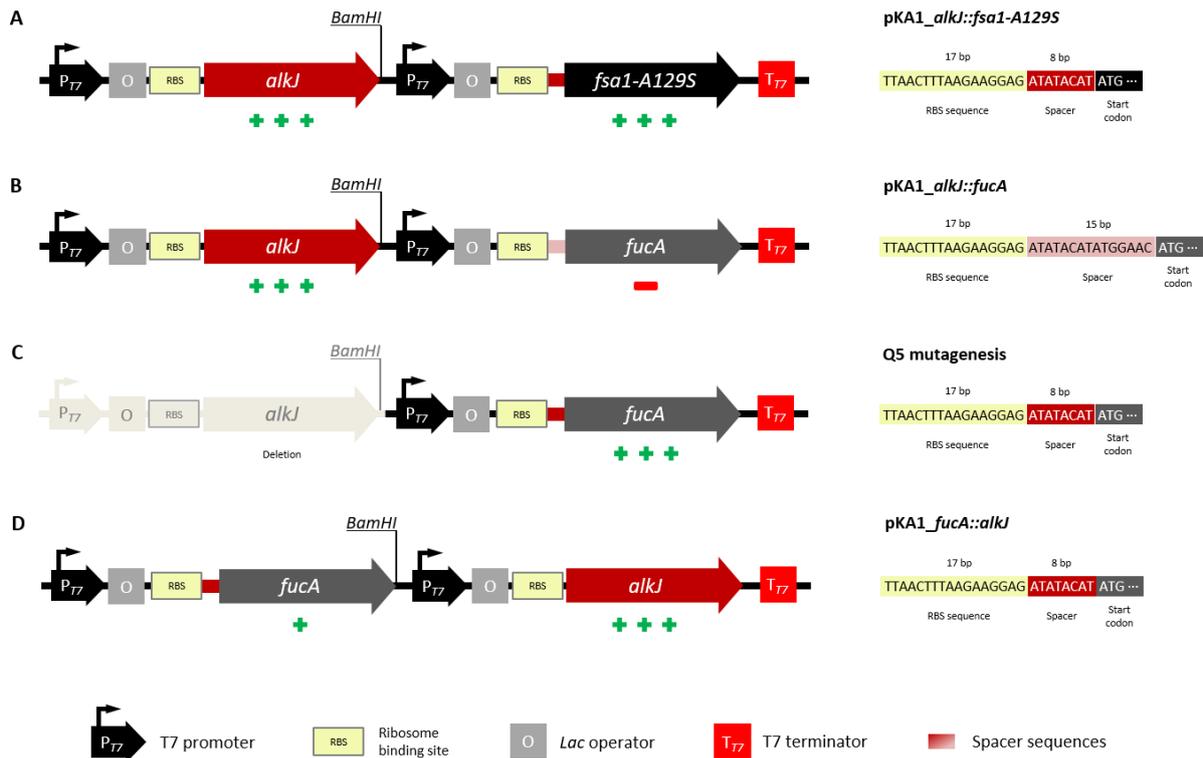


Figure D-31. Contextuality of the *fucA* gene in different constructs for pathway expression. (A) The backbone the parent POP plasmid used to construct *pKA1_alkJ::fucA*. Cloning strategy aimed at replacing the *fsa1-A129S* by the target *fucA* gene. (B) The unintended insertion of 7 bp in the spacer region resulted in a frameshift and abolished *fucA* expression. (C) Q5 site-directed mutagenesis led to the unexpected deletion of the *alkJ* coding region. (D) Alternative arrangement of the *alkJ* and the *fucA* gene in POP configuration. On the right: Simplified schemes of different constructs. Lengths of ORFs do not reflect actual length. Expression levels according to SDS-PAGE analysis are indicated with high (+++), low (+), and no expression (-) below the corresponding ORF. On the right, genetic context of the *fucA* gene in each construct.

One solution was already intended and aimed at the construction of the second enzymatic core modul in pseudo-operon configuration, *pKA1_fucA::alkJ*, and will be outlined below. Another straight forward route longing to construct *pKA1_alkJ::fucA* employed another round of FC. Other than the cloning approach described above, primer were designed to amplify the *pKA1* backbone and the *fucA* gene from the previously constructed *pKA1_fucA* (D II.4.2).^[509] The *alkJ* insert including its *P_{T7}* and RBS was amplified depending on the initially assembled *pKA1_alkJ*.^[348] Unfortunately, assembly of the two linear fragments was not successful, neither with FC nor the application of SLiCE due to unknown reason (see G VII.2.1).

Alternatively, a two-step cloning strategy could be applied. In the first cloning step, the *fucA* gene can be amplified with primers containing *NcoI/XhoI* restriction sites using *pKK223-3_fucA* as PCR template. Subcloning of the insert into *pET16b* will result in the optimal genetic context for *fucA* gene expression. Subsequently, SLiC can be performed to generate the target *pKA1_alkJ::fucA* from scratch. Due to time constraints, this cloning strategy could not be pursued.

However, *FucA* was readily produced from the plasmid generated by mutagenesis in *E. coli* BL21(DE3) transformants (Figure D-32). Sanger sequencing of the *fucA* coding region finally confirmed the successful deletion of the target 7 bp insertion (Figure D-31C).^[3]

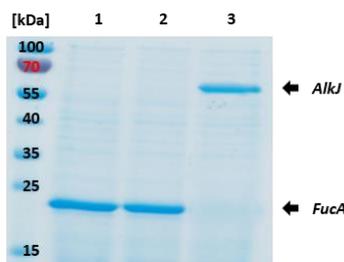


Figure D-32. Characterization of pKA1_alkJ::fucA post site-directed mutagenesis using the Q5® system. SDS-PAGE analysis of whole cell samples of *E. coli* BL21(DE3) transformed with pKA1_fucA (1), cells harboring the mutated plasmid (2), and cells containing the unmutated plasmid (3). Protein production in AIM for 24 h. Sample loading normalized to $O_{D590} = 7.0$. The figure was adapted from J. Reiterlehner (2017).^[3]

As noted above, the second vector to be constructed contained the two genes in pseudo-operon but the order of the two genes was inverted. The previously assembled pKA1_fucA vector (clone: BWA4)^[509] served as template for backbone amplification. The *alkJ* insert including the P_{T7} and RBS was amplified using the pKA1_alkJ plasmid.^[348] Assembly of the two linear fragments was achieved by FC as described in G VII.2.2. Successful assembly of pKA1_fucA::alkJ (Figure D-31D) was confirmed by *NcoI* control digestion yielding the expected DNA restriction pattern with one 4.6 kb and one 3.3 kb fragment (Figure D-33A). Sanger sequencing confirmed the *alkJ* sequence (clone: JRE2-1; H I.1.9.7).^[3]

E. coli BL21(DE3) transformants harboring pKA1_fucA::alkJ were cultivated in AIM as before. SDS-PAGE analysis showed moderate overexpression of *AlkJ* and a weak band corresponding to the size of *FucA* (Figure D-33B).^[3] *AlkJ* activity was confirmed in RCs under standard screening procedures with the primary aromatic alcohols 2b and 4-5b (Figure D-34). All model substrates were oxidized to the target aldehydes 2c and 4-5c. In contrast to other plasmids producing *AlkJ*, none of the substrates were fully consumed (see Figure D-6 and Figure D-29 for comparison). Overoxidation to the corresponding carboxylic acids 2d and 4-5d was observed under experimental conditions as expected.

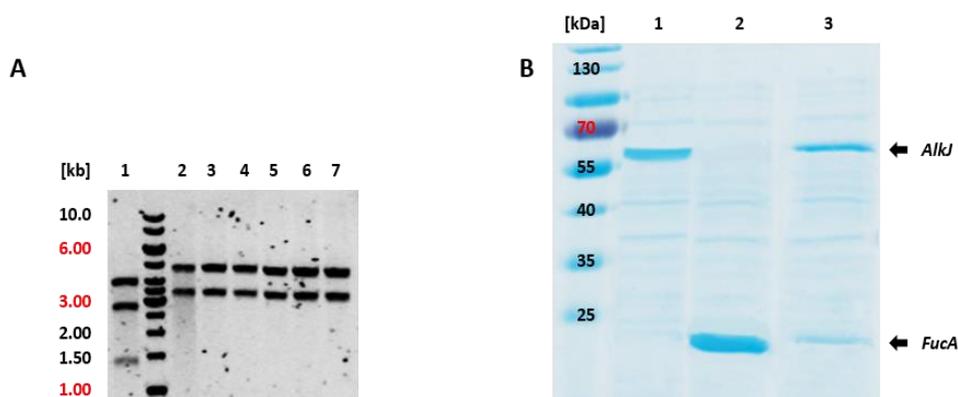


Figure D-33. Coproduction of *AlkJ* and *FucA* from pKA1_fucA::alkJ. (A) *NcoI* control digestion of pPOP as control (1) and correctly assembled plasmids re-isolated after FC (2-6). (B) SDS-PAGE analysis of whole cell samples of *E. coli* BL21(DE3) transformed with pKA1_alkJ (1), pKA1_fucA (2), and the newly assembled pKA1_fucA::alkJ coproducing both enzymes (3). Protein production as before. Sample loading normalized to $O_{D590} = 7.0$. The figure was adapted from J. Reiterlehner (2017).^[3]

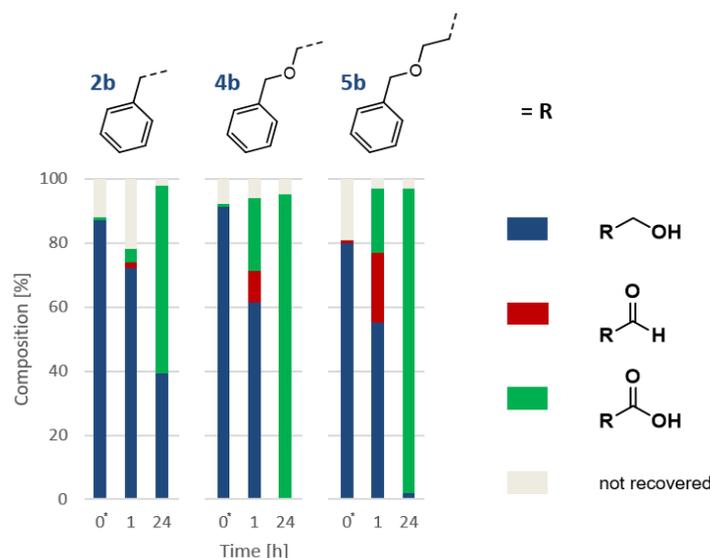


Figure D-34. Functional screening of *AlkJ* in the *pKA1_fucA::alkJ* construct. The primary aromatic alcohols **2b** and **4–5b** were oxidized to the desired aldehydes. Overoxidation to the corresponding carboxylates was observed in all transformations. RC screenings were performed under standard conditions; t_0^* sample taken immediately after the addition of 5 mM substrate and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <15% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material at t_0^* due to insufficient mixing. The results for **2b** and **4b** by courtesy of J. Reiterlehner (2017).^[3]

For *de novo* pathway assembly, competent *E. coli* BL21(DE3) cells were cotransformed with the newly assembled *pKA1_fucA::alkJ* and *pET26b(+)_phoN-Sf*. Stable plasmid maintenance is conveyed by the compatible ORIs, P15A and ColE1. Cotransformants were selected on agar plates supplemented with Cam and Kan and coproduction of *AlkJ*, *FucA*, and *PhoN-Sf* tested in AIM (**Figure D-35**).

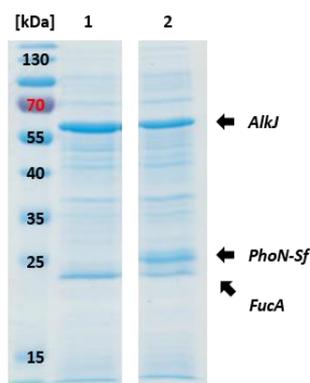


Figure D-35. Pathway expression in AIM. SDS-PAGE analysis of whole cell samples of *E. coli* BL21(DE3) transformed with *pKA1_fucA::alkJ* coexpressing *AlkJ* and *FucA* (**1**) and cotransformants also expressing *PhoN-Sf* from *pET26b(+)_phoN-Sf* (**2**). Protein production in AIM for 18 h. Sample loading normalized to $OD_{590} = 7.0$.

Since the construction of one pathway core module (**Figure D-31D**) was achieved by SLIC and all three pathway enzymes, the ADH *AlkJ*, the DHAP-dependent aldolase *FucA*, and the phosphatase *PhoN-Sf* were successfully coexpressed in *E. coli* BL21(DE3) (**Figure D-35**), the engineered *E. coli* was subjected to experiments for pathway validation.

D III.2.2 Unsuccessful production of (3*R*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e_{FucA}**) *in vivo*

First pathway validation experiments employed RCs under standard screening conditions (G III.2). Briefly, *E. coli* BL21(DE3) cotransformed with pKA1_ *fucA*::*alkJ* and pET26b(+)_ *phoN-Sf* (referred to as the AFucP strain; **Figure D-36**) was cultivated in AIM and expression was confirmed by SDS-PAGE as before (see **Figure D-35** for comparison). The reaction was started by the addition of **2b**. Samples were taken for GC and HPLC analysis as described in G III.1 and G III.2, respectively. Analyses did not show consumption of **2b** in the AFucP strain. In a parallel experiment, *E. coli* BL21(DE3) transformed with pPOP produced the target aldehyde **2c** with the highest concentration detected after 2 h reaction time (**2c**: 44.2±4.5%) according to calibrated GC/FID. The aldehyde was further oxidized to **2d**, being the main product after 24 h (**2b**: 17.1±1.6%, **2c**: n.d., **2d**: 77.8±1.7%, recovered material: 94.8±1.7%). This suggested an inactive ADH for the *in situ* preparation of the aldehyde intermediate **2b** in RCs of the engineered AFucP strain.

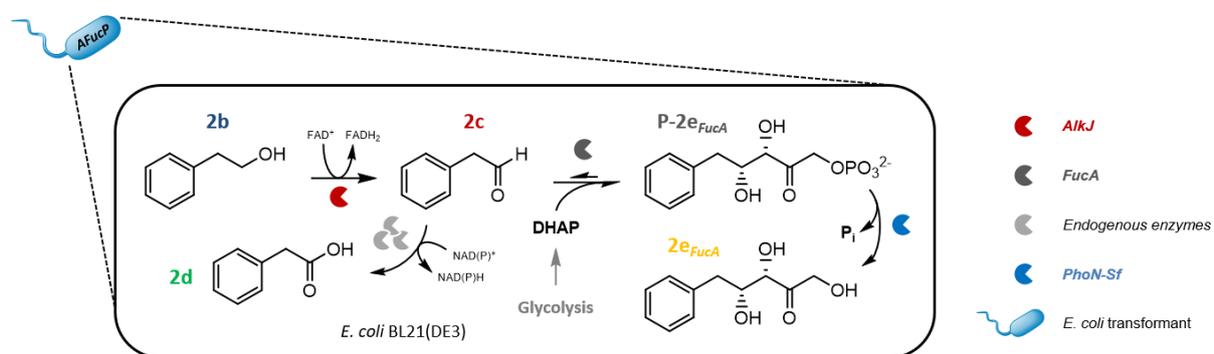


Figure D-36. Synthetic DHAP-dependent pathway scheme. The *de novo* pathway consists of *AlkJ*, which oxidizes **2b** to **2c** by *AlkJ* and the DHAP-dependent aldolase *FucA*, catalyzing the aldol reaction with the glycolytic DHAP to form the phosphorylated intermediate **P-2e_{FucA}**. Finally, *PhoN-Sf* irreversibly dephosphorylates the intermediate yielding the target aldol **2e_{FucA}**. Endogenous host enzymes interfere with the flux through the artificial pathway and produce the carboxylate **2d**.

As this was in remarkable contrast to the functional testing of *AlkJ* in RCs harboring the pKA1_ *fucA*::*alkJ* plasmid (**Figure D-34**), the immediate attempt to synthesize the desired aldol **2e_{FucA}** was to circumvent the oxidation step and entering the cascade at the intermediate aldehyde stage. Therefore, RCs were challenged with **2c**. Since the presence of high aldehyde concentrations can impair cell viability,^[280] the screening was started with **2c** at standard (5 mM) and reduced final concentrations (1 mM). However, no formation of aldol adduct **2e_{FucA}** was detected over time (0–24 h). Furthermore, **2c** was reduced to a lesser extent in RCs expressing the whole pathway than in a control experiment employing pPOP transformants as shown for 5 mM initial concentration of **2c** in **Figure D-37**.

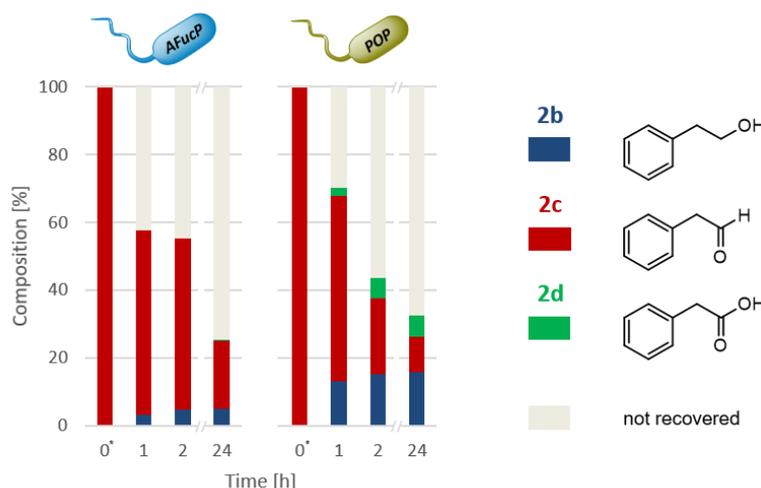


Figure D-37. DHAP-dependent pathway validation starting from aldehyde substrates. *E. coli* BL21(DE3) expressing *AlkJ*, *FucA*, and *PhoN-Sf* (AFucP) did not transform **2c** into the dephosphorylated aldol product **2e_{FucA}**. Additionally, **2c** was poorly detoxified by the metabolic host background suggesting reduced cell viability. In transformants heterologously expressing *AlkJ* and *Fsa1-A129S* (POP), **2c** was mainly converted to the corresponding alcohol **2b** by endogenous enzyme activities. Screenings performed under standard conditions; t_0^* sample taken immediately after the addition of 5 mM **2c** and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <10% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material due to volatility, binding to biomolecules (e.g., proteins),^[400] and reduced cell viability.

The addition of aldehydes such as **2c** induces stress responses in living cells.^[280, 471] These responses are complex and involve the adaption of gene expression and the upregulation of enzymatic activities to reduce the oxidative/electrophilic burden.^[281, 510] Hence, RCs might not be capable to respond appropriately since the RCM lacks a nitrogen source and restricts bacterial reproduction, growth, and the biosynthesis of proteins (including detoxifying enzymes). In experiments challenging RCs with the addition of **2c**, differently engineered strains exhibited different detoxification responses (**Figure D-37**). Whereas **2c** was metabolized mainly to the alcohol **2b** (15.1±3.2%) and minorly to the carboxylate **2d** (1.5±1.2%) after 2 h reaction time in pPOP transformants, formation of the nontoxic **2b** was 3.2-fold reduced in the AFucP strain (**2b**: 4.7±1.2%, **2d**: n.d.). This indicates that cells expressing the DHAP-dependent aldolase pathway consisting of *AlkJ*, *FucA*, and *PhoN-Sf* are probably less viable than the less burdened strain overproducing only two enzymes, *AlkJ* and *Fsa1-A129S*, from one plasmid.

To reduce pathway complexity and the innately high metabolic burden, *E. coli* BL21(DE3) RCs expressing the DHAP-dependent aldolase from pKA1_ *fucA* and *PhoN-Sf* from pET26b(+) were prepared (referred to as FucP strain). SDS-PAGE analysis confirmed protein production prior to starting the functional screening. RCs were challenged with 5 mM or 1 mM **2c**. Again, no production of **2e_{FucA}** could be detected over time (0–24 h) according to calibrated HPLC.

Wei *et al.* recently established a synthetic pathway consisting of a DHAP-dependent aldolase (i.a., *FucA* from *Thermus thermophilus* HB8) and the *E. coli* phosphatase *YqaB*. Their pathway assembly involved subcloning of both genes into pCDFDuet-1, which resulted in the pCDFDuet-1_ *fucA*::*yqaB* construct. Optimization of the synthetic pathway was based on the water-soluble acceptor aldehyde 3-trifluoroacetamido propanal (3-TFAP) and condensation with intracellular DHAP to form the corresponding aldol adduct.^[293]

Protein production was performed in LB-ECAM. The engineered strain was grown to OD₅₉₀ in the exponential phase (OD₅₉₀ > 1.8). Protein production was induced by IPTG (1 mM, 30°C). Importantly, biotransformations were

started with 20 mM aldehyde and additional 20 mM aldehyde fed at later time points.) These optimized conditions yielded 12.8% (3*R*,4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one (*dr* = 87:13).^[293]

Motivated by Wei *et al.*, who demonstrated the functionality of their DHAP-dependent aldolase pathway *in vivo*, fermentation and aldol production conditions were adapted where possible. However, preliminary cultivation studies in LB-ECAM showed significantly reduced bacterial growth (e.g., untransformed *E. coli* BL21(DE3), AFucP and FucP strain) and did not exceed OD₅₉₀ > 1.5. Hence, LB-ECAM was discarded as cultivation medium. The use of LB medium (supplemented with the appropriate antibiotics) and 1 mM IPTG to induce protein expression resulted in the successful coproduction of *FucA* and *PhoN-Sf* (**Figure D-38A**). The increased expression temperature (30°C) and IPTG concentration also led to high amounts of protein in insoluble fractions (**Figure D-38B**).

After expression, cells were transferred into 8 mL reaction vials and biotransformations performed at OD₅₉₀ = 6.0–7.0 but otherwise standard screening conditions as described in G III.2. After the addition the aldehyde **2c** (1 mM, 5 mM, or 10 mM; V_{total} = 2.0 mL), samples were taken for GC and HPLC analysis. The increase to 10 mM **2c** was difficult due to poor solubility of the substrate. Since *E. coli* BL21(DE3) cell viability was impaired in the presence of higher amounts of ACN (as cosolvent to facilitate substrate solubility) and concentrations of **2c** as low as 0.1 mM restricted bacterial growth to half of the maximal OD₅₉₀ value, aldehyde concentrations higher than 10 mM were not tested (**Figure D-39**). Ultimately, the desired aldol **2e_{FucA}** was not produced under the adapted conditions (**Table D-6**).

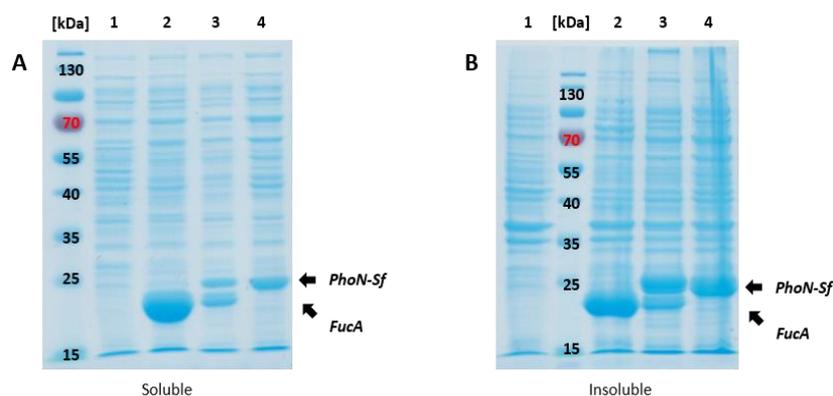


Figure D-38. IPTG-induced protein production in the FucP strain. (A) SDS-PAGE analysis of soluble protein fractions of untransformed *E. coli* BL21(DE3) (**1**), pKA1_*fucA* transformants (**2**), the FucP strain coexpressing *FucA* and *PhoN-Sf* (**3**), and pET26b(+)_*phoN-Sf* transformants (**4**). (B) SDS-PAGE analysis of insoluble fractions; loading scheme as in (A). Protein production in LB-Miller medium but otherwise according to Wei *et al.* (2015).^[293] Sample loading normalized to 10 µg total amount of protein per lane.

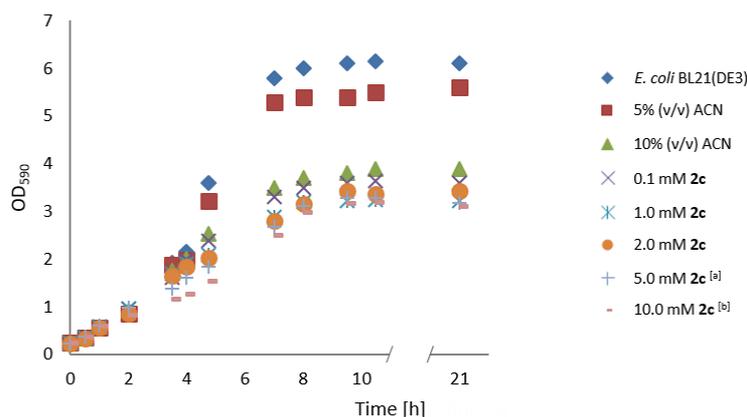


Figure D-39. Bacterial growth in the presence of organics. Growing *E. coli* BL21(DE3) cells were challenged with solely 5–10% (v/v) ACN or different concentrations of aldehyde **2c** (0.1–10 mM) and 5% (v/v) ACN as the cosolvent. Addition of organics at OD₅₉₀ > 1.0. ^[a] Insolubility of **2c** detected after 2 h; ^[b] insolubility detected upon addition of **2c**.

Despite approximation to the conditions suggested by Wei *et al.* and testing the FucP strain, differences between the two systems remained that might be crucial. One inherent difference is based on the particular aldehyde compounds to be converted *in vivo*. The Wang group mainly transformed terminally substituted propanals, most of them bearing a $-\text{NHCOCF}_3$ group that facilitated solubility of substrates in water. Unfortunately, their standard substrate 3-TFAP only became available at the end of this thesis and, thus, could not be tested. The synthesis of 3-TFAP was performed by T. Wiesinger.^{Wiesinger, 2017 #924} Additionally, the analysis of 3-TFAP, its corresponding aldol adduct, and potential byproducts (e.g., 3-trifluoroacetamido propanoic acid) would have required a different HPLC analytics,^[293] which would not have been possible to synthesize and establish, respectively, within the given time frame.

Regarding the choice of aldehyde, the water-soluble 3-TFAP and derivatives might not induce oxidative/electrophilic stress responses in *E. coli* as pronounced as the aldolase acceptor aldehydes **2c** and **4–6c**. Furthermore, Wei *et al.* fed their system with up to 40 mM of 3-TFAP to shift the equilibrium toward the desired aldol product. Since aldehyde toxicity is an unmet challenge of *in vivo* biocatalysis,^[280] their preparation *in situ* – as intended by the implementation of *AlkJ* – is highly preferred and the overall concentration in the system should be low. This can be achieved either by increasing the amount of active *FucA* in the system to convert the toxic aldehyde intermediates or by a ‘intracellular reservoir’ of aldehydes that contains reactive aldehyde species at viable concentrations, yet freely available for the subsequent aldol reaction.^[348] The latter will be discussed as one optimization strategy in E 1.4 and balancing enzyme stoichiometry to enhance the flux through the synthetic pathway in E 1.6. However, increasing the production of *FucA* (or any other pathway enzyme) can impose an additional metabolic burden and might reduce cell viability and, consequently, overall pathway performance. As experimental data indicate, the metabolically burdened AFucP strain is less robust than the POP strain if challenged with aldehydes, for example (Figure D-37). Although these two strains are unequally burdened with AFucP expressing three pathway enzymes from two plasmids and POP expressing the mini-pathway consisting of *AlkJ* and *Fsa1-A129S* from one plasmid, the overproduction of the nonnative *PhoN-Sf* phosphatase might be toxic to the host cell. The strong overexpression of recombinant enzymes can evoke a general stress response in *E. coli*.^[6, 58, 279] Indications to a cellular response resulting in the downregulation of *PhoN-Sf* in periplasmic (Figure D-22) and cytosolic protein fractions (Figure D-23A) at longer expression times do not necessarily relate to the toxicity of *PhoN-Sf*. Nonetheless, subtle differences in bacterial growth can be seen under expression conditions (Figure D-40).

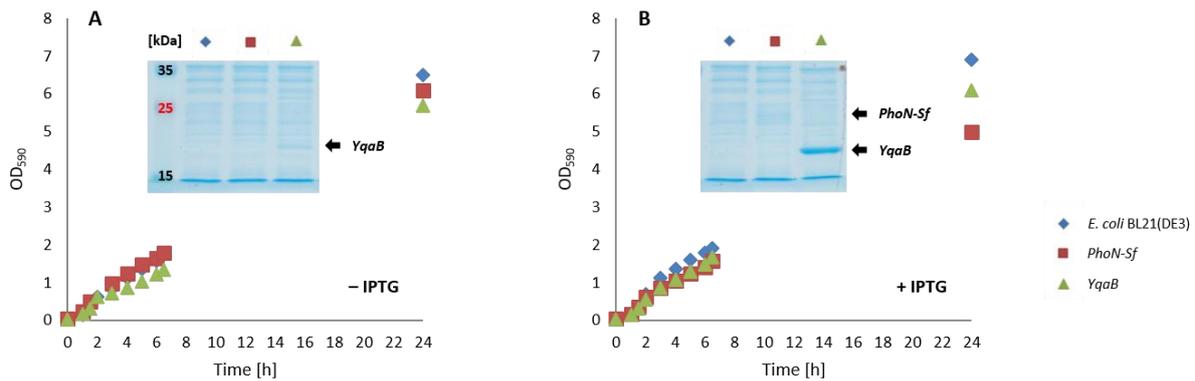


Figure D-40. Bacterial growth under phosphatase expression conditions. (A) Growth of untransformed *E. coli* BL21(DE3) (◆), pET26b(+)*_phoN-Sf* transformants (▲), and pCDFDuet-1*_yqaB* transformants (■) in the absence of IPTG. (B) Phosphatase expression upon IPTG addition. Growth studies performed in duplicates. Inserts: SDS-PAGE analysis of whole cell samples collected after 24 h. Sample loading normalized to $OD_{590} = 7.0$.

Cultures were grown at 37°C (200 rpm) until $OD_{590} > 0.5$ was reached. IPTG was added, if applicable, and the temperature switched to 20°C for 24 h. In the absence of IPTG, leaky expression of the *E. coli* *YqaB* was observed after 24 h cultivation time, which is reflected in the slightly lower OD_{590} value (Figure D-40A). The addition of IPTG induced phosphatase production and increased the burden on host cells, which is reflected in the lower final OD_{590} values for *PhoN-Sf* and *YqaB* transformants. The final $OD_{590} = 6.1$ of cells burdened by the high yielding production of *YqaB* is slightly lower than the final $OD_{590} = 6.9$ of untransformed *E. coli* BL21(DE3) cells (Figure D-40B). In contrast, *PhoN-Sf* was only weakly expressed under experimental conditions according to SDS-PAGE analysis. The final $OD_{590} = 5.0$ was significantly lower suggesting a negative influence of *PhoN-Sf* production on the whole cell system.

The application of the nonnative *PhoN-Sf* is another difference between the FucP strain constructed in this thesis and the system established in the Wang group, who used *YqaB* from *E. coli* instead. This difference might contribute to the – so far – unsuccessful production of the target aldol **2e_{FucA}** *in vivo*.

Nonspecific acid phosphatases share a conserved active site with mammalian glucose-6-phosphatases (G6Pase).^[511] Tanaka *et al.* examined the phosphorylation of glucose and dephosphorylation of G6P catalyzed by *PhoN-Sf* (and *PhoN-Se_{SP}*). Their findings suggest that *PhoN-Sf* regiospecifically phosphorylates glucose to G6P in the presence of PP_i ($K_M = 5.3$ mM at pH 6.0). On the other hand, the K_M value for G6P is much lower ($K_M = 0.02$ mM at pH 6.0). Although this study does not reflect the neutral to slightly basic pH of the cytoplasm in *E. coli* and the determined K_M values may not reflect the catalytic efficiency of *PhoN-Sf in vivo*, a low K_M indicates optimal use of small substrate levels in the environment. Hence, the lower K_M for G6P resembles the native function of nonspecific acid phosphatases, which is the tight binding of phosphorylated compounds and, subsequently, cleavage of the ester bond to acquire inorganic phosphate and the organic byproduct. The access to alternative nutrient sources certainly provided evolutionary advantages. In synthetic pathway applications, however, the unbalanced overproduction of *PhoN-Sf* might interfere with the central carbon metabolism, consequently, reducing cell viability by the undesired dephosphorylation of phosphorylated metabolites such as G6P. (As a side note, the intracellular concentration of G6P is 0.08 mM under steady state conditions in *E. coli*.^[512])

On the other hand, the intracellular accumulation of phosphorylated aldol adducts may be toxic to host cells or, at least, a burden. As noted above, the phosphatase should selectively dephosphorylate the target aldol products under physiological conditions with little, ideally without any interference to other phosphorylated metabolic intermediates.^[293] As demonstrated by Wei *et al.*, the irreversible dephosphorylation and the secretion the

resulting aldols out of the cell offer to shift glycolysis and the aldol reaction toward the formation of product. To follow this strategy, *YqaB* was previously characterized *in vitro* (Figure D-24). Furthermore, expression and the influence on cellular growth compared to *PhoN-Sf* and the unburdened *E. coli* BL21(DE3) whole cell system was studied (Figure D-40). Pathway implementation of *YqaB* and validation of the optimization potential will be discussed in E 1.6.

The last initial variation from the conditions applied by Wei *et al.* was the use of RCs and its limitation already discussed. Wei *et al.* employed *E. coli* BL21(DE3) cells in their late exponential phase (LEP). An on-demand feed of glucose compensated its consumption and guaranteed a constant supply of DHAP, the donor for *FucA* mediated aldol reaction, *via* glycolysis. Optimization strategies targeting the intracellular DHAP pool will be addressed in E 1.6.1 and included the *in situ* production DHAP from an external non-glucose carbon source.

D III.2.3 Interim summary I

Although both engineered strains, AFucP and FucP, successfully expressed all target pathway enzymes according to SDS-PAGE analysis (Figure D-35 and Figure D-38A), substrates (**2b** or **2c**) were not converted under various conditions and the target aldol product **2e_{FucA}** not produced (Table D-6).

Table D-6. Summary: Unsuccessful production of the aldol **2e_{FucA}**

Engineered strains ^[a]	Expression conditions	Expression levels ^[b]			Screening conditions	Aldol production	Associated literature
		<i>AlkJ</i>	<i>FucA</i>	<i>PhoN-Sf</i>			
	AIM ^[c]	+++	+	+	RCs ^[d, e] (5 mM 2b)	n.d.	This work
	AIM ^[c]	+++	+	+	LEPs ^[e, f] (5 mM 2b)	n.d.	This work
	AIM ^[c]	+++	+	+	RCs ^[d] (1 / 5 mM 2c)	n.d.	This work
	LB-ECAM / Wei <i>et al.</i> ^[g]	n.a.	n.a.	n.a.	n.a. ^[h]	n.a.	This work / Wei <i>et al.</i> ^[293]
	LB-Miller / Wei <i>et al.</i> ^[g]	n.a.	++	++	RCs ^[d] (1 / 5 mM 2c)	n.d.	This work / Wei <i>et al.</i> ^[293]
	LB-Miller / Wei <i>et al.</i> ^[g]	n.a.	++	++	LEPs ^[f] (1 / 5 mM 2c)	n.d.	This work / Wei <i>et al.</i> ^[293]

^[a] AFucP harboring pKA1_*fucA::alkJ* and pET26b(+)_*phoN-Sf*, FucP harboring pKA1_*fucA* and pET26b(+)_*phoN-Sf*. ^[b] Expression levels according to SDS-PAGE analysis. ^[c] AIM inoculated with 0.2% (v/v) preculture (37°C, 150 rpm, 4 h), expression at 20°C, 150 rpm, 18 h. ^[d] RCs in RCM (OD₅₉₀ = 10.0), 5 mM substrate, 5% (v/v) ACN, 25°C, 250 rpm. ^[e] No *AlkJ* activity in AFucP. ^[f] Late exponential phase cells (LEPs; OD₅₉₀ = 5.0–7.0), 1% (w/v) glucose, 5 mM substrate, 5% (v/v) ACN, 25°C, 250 rpm. ^[g] Medium inoculated with 1% (v/v) preculture (37°C, 200 rpm, OD₅₉₀ > 1.5), expression with 1 mM IPTG at 30°C, 200 rpm, 12 h. ^[h] Experiment terminated due to low maximal OD₅₉₀ ≤ 1.5 in LB-ECAM (37°C, 200 rpm); expression media were supplemented with Cam and Kan; n.a. = not applicable, n.d. = not detected.

Not only remains the cytotoxicity of aldehydes a major obstacle for biotransformations in RCs (Figure D-37), the presence of aldehyde even at low concentrations is deleterious to growing cells, as well (Figure D-39). Furthermore, experimental data suggest detrimental effects of the heterologous expression of the *de novo* pathway enzyme *PhoN-Sf* on cell viability. The enzymatic background in RCs expressing *PhoN-Sf* fails to counter the aldehyde burden by reducing the reactive aldehyde species to the nontoxic primary alcohols. Contrary, RCs bearing a similar metabolic burden from overexpressing the mini-pathway consisting of *AlkJ* and *Fsa1-A129S*, but not the phosphatase, detoxify aldehydes to the corresponding alcohols and carboxylic acids (Figure D-37). Furthermore, literature precedence points to the possibility of an undesired dephosphorylation activity of

PhoN-Sf.^[511] The dephosphorylation of phosphorylated metabolites, potentially including G6P, might not only explain the reduced cellular viability; it fatally disturbs glycolysis and the continuous production of DHAP, the donor for *FucA* mediated aldol formation. In comparison to *E. coli* BL21(DE3) overexpressing the native *YqaB*, cells heterologously expressing *PhoN-Sf* showed also reduced growth (**Figure D-40**). This substantiates the increased burden from *PhoN-Sf* expression on host cells.

E Optimization strategies for synthetic pathways *in vivo*

This chapter comprises various strategies to enhance the flux through artificial enzyme cascades in living cells and a refined SPE purification, which maximized isolated aldol yields.

E I Optimization of pathway performance by complementing flux enhancement strategies

In past decades, synthetic routes with increasing complexity have been realized by cascade-type reactions to produce value-added chemicals.^[6-8, 277, 454] The cooperative effect of multiple biocatalysts with their inherently high chemo-, regio-, and stereoselectivity in combination with the omission of intermediate isolation has increased product yields of *in vitro* and whole-cell-mediated cascades.^[9, 348] Both approaches strictly require a balanced carbon flux through the synthetic pathway to meet performance metrics such as high product titers. As noted earlier, *in vivo* approaches offer advantages including the recycling of cofactors and the simultaneous production of all pathway elements.^[6]

Despite these obvious benefits, the implementation of artificial metabolic routes in living cells can be highly challenging since control over the carbon flux can be complicated in many aspects. The introduction of nonnative enzymes potentially interferes with the metabolic environment in host cells, which, in turn, might primarily impair pathway functionality *in vivo*.^[6, 10, 47] Reciprocal interactions between the host and heterologous pathways or unexpected interactions between different synthetic genetic elements can be summarized as context dependency or context effects.^[6, 490]

In the previously assembled mini-pathway consisting of *AlkJ* and *Fsa1-A129S*, both enzymes are synthesized by the cellular transcription/translation machinery but enzyme production had to be adjusted to increase *Fsa1-A129S* yields. A minimal change in the genetic context by the introduction of an individual P_{T7} upstream of the *fsa1-A129S* gene sufficed to improve aldolase production (**Figure D-25**). With increasing pathway complexity, overall balancing of enzyme stoichiometry becomes increasingly important to prevent the accumulation of (harmful) cascade intermediates, which impedes the flux through the synthetic pathway and reduces host viability.^[6, 348] In this regard, the *in situ* production of aldehydes by *AlkJ* as intermediates for the subsequent aldol formation is a very suitable example. The metabolic background of *E. coli* converts reactive aldehyde intermediates further to the corresponding carboxylic acids. In the 'host context' of *E. coli*, these cytotoxic intermediates are metabolized (i.e., detoxified) to reduce the oxidative and electrophilic stress imposed by the reactive carbonyl group in aldehydes.^[280-281] Other hazardous chemicals might not only be subject to metabolism but can also be actively transported across the cellular barrier.^[513] However, competing side reactions such as the formation of carboxylates potentially interfere with *de novo* pathway performance (**Figure D-29A**) and decrease aldol product yields (**Figure D-29B**).^[348]

Regarding toxicity, the heterologous expression of enzymes not only contributes to the metabolic burden on the host; the production of a target pathway enzyme (above a certain level) can lead to adverse and detrimental effects on cell growth and viability caused by the nonnative enzymatic activity, for example (Figure D-40). This might be the case for the phosphatase *PhoN-Sf* by the unintended dephosphorylation of phosphorylated metabolic intermediates (Figure D-36).^[6, 52, 492, 511]

Methodologies from metabolic and genetic engineering, systems biology, and synthetic biology, efficiently removed many of the bottlenecks described above and led to optimized enzyme cascades *in vivo*.^[6, 36-37, 348] Well-known genetic regulatory elements (e.g., promoters, RBS) to balance enzyme production were complemented by substrate channeling strategies as noted earlier in this thesis.^[6, 349] A synthetic protein scaffolds, for example, was applied to recruit tagged pathway enzymes. The number of interaction domains provided by the scaffold dictated enzyme stoichiometry and toxic cascade intermediates were channelled along the scaffold, which enhanced the pathway flux (Figure C-37).^[283] Alternative strategies to increase pathway fluxes are the engineering of host cell genomes by targeted gene KO. Different approaches have long been used to minimize, even eradicate, competing endogenous host enzyme activities to reroute the carbon flux of natural metabolic pathways and to improve synthetic pathway performance (Figure C-28).^[34, 42, 44, 406, 514]

The methods described above are well-established and proved their applicability beyond doubt. However, single optimization approaches may not suffice to tackle different bottlenecks and contextual issues simultaneously to significantly enhance the flux through the artificial pathway. Moreover, bottlenecks might not be easy to detect, particularly when genetic devices or pathways considered to be successfully assembled fail to function. To identify and solve such 'context bottlenecks', complementing strategies were applied to optimize the previously assembled enzymatic cascades (Figure D-29A and Figure D-36) and will be discussed in the following.

E I.1 Synthetic pathway contextuality and interactions between *de novo* genetic elements

Host cells and their metabolic networks had long periods of coevolution and led to optimized endogenous processes that provide robust and effective cellular functions for survival. In contrast, heterologous pathways are designed from metabolically (and contextually) unrelated biological parts and their assembly performed isolated from the target cellular host context. Hence, synthetic pathways have not had the advantage of evolutionary optimization.

Genetic context effects have already been encountered in previous chapters. One example was the construction of pKA1_alkJ::fucA. Whereas *AlkJ* was functionally expressed, *FucA* production was abolished due to an unintended insertion of 7 bp into the spacer region between the RBS and the ATG start codon of the *fucA* gene (Figure D-31B). The application of a site directed mutagenesis procedure not only deleted the undesired 7 bp and restored the optimal context for *FucA*; the whole *alkJ* gene was deleted (Figure D-31C). In contrast, both enzymes could be produced from the pKA1_fucA::alkJ containing the two genes in inversed order (Figure D-31D). However, in the genetic context for coexpression (Figure D-33B), *FucA* is less efficiently synthesized compared to the context of single enzyme expression in pKK223-3_fucA and pKA1_fucA (Figure D-19A).

In the pKA1_fucA::alkJ plasmid, both genes feature identical regulatory sequences (P_{T7} , RBS, and spacer sequence). Nonetheless, the genetic context is different since these sequences control DNA sequences of different compositions, precisely, the *fucA* and the *alkJ* gene. The RBS sequence, for example, determines the efficiency of translation. In turn, the sequence to be translated influences RBS efficiency.^[6]

Generally, the activity of regulatory and expressed sequences composed on the same DNA molecule can be influenced by their specific ordering. Such linkages may facilitate structural interactions on DNA or mRNA level.

Like domains on the same polypeptide or subunits within a protein complex are only active when adopting their native conformations and functional quaternary structures (i.e., are placed in the right structural context), synthetic DNA molecules may depend on defined conformations to exhibit a certain function.^[490]

One undesigned interaction that has not been described so far in this thesis occurred with the construction of mini-pathway arrangements in MON configurations, which colocalized T_{Syn} sequences with the *alkJ* and the *fsa1-A129S* genes (**Figure D-25C**). According to SDS-PAGE, both enzymes were readily produced in *E. coli* BL21(DE3) harboring pMON1 or the pMON4 (**Figure D-28**). However, functional testing of *AlkJ* revealed that substrates **2b** and **4b** were not or hardly oxidized to **2c** and **4c**, respectively (**Figure E-3B and C**).

Analyzing at the genetic context, the terminator sequences B0011 and B0014 were placed between the coding regions of the ADH and the aldolase. Cloning of B0011 in the pMON1 (pKA1_alkJ::B0011::fsa1-A129S)^[435] resulted in a 16 bp spacer sequence; cloning of B0014 in the pMON4 (pKA1_alkJ::B0014::fsa1-A129S) also contained 16 bp flanking the terminator sequence (**Figure D-25C**). The spatial proximity of the T_{Syn} sequences to the C-terminus of the *alkJ* gene might cause unwanted interactions between the two genetic elements. Since Sanger sequencing confirmed the integrity of the *alkJ* gene sequence, interactions between the 3'-end of the *alkJ* gene and the T_{Syn} sequences must translate into variations in the protein sequence that lead to decreased activity toward **2b** and **4b**. Since B0011 and B0014 form stable hairpin structures, their rigidity in the context of the surrounding DNA sequence might influence local DNA conformation that cause the T7 RNA polymerase to fall off, prematurely stopping transcription.^[515] Secondly, the composition of nucleotides may result in T7 RNA polymerase slippage.^[516-517] Thirdly, the strong secondary structure of the synthesized mRNA molecule incorporates nucleotides in 3' coding region of the *alkJ* gene, terminating translation rather dependent on the structural context than a freely accessible stop codon. Translation termination without a dedicated stop codon was recently suggested for eukaryotes but might occur in prokaryotes as well, especially in artificial genetic devices.^[518] Nonetheless, all three scenarios would result in a truncated C-terminus of *AlkJ*. Truncations can only concern a few amino acids since SDS-PAGE analysis shows a protein band corresponding to the expected size of *AlkJ* (**Figure D-28**).

To study whether local interactions between the 3'-end of the *alkJ* gene and B0014 undermine the function of *AlkJ*, PCR primer pairs were designed to increase the length of spacer sequences flanking the T_{Syn} . Primers included spacers of 50 bp and 100 bp, giving rise to pMON5 (pKA1_alkJ::B0014₅₀::fsa1-A129S) and pMON6 (pKA1_alkJ::B0014₁₀₀::fsa1-A129S), respectively (**Figure E-1**). Primers also encoded *BglIII* restriction sites, whereas the target pPOP vector was digested with *BamHI* to insert the B0014 variants. By utilization of the *BamHI/BglIII* isocaudamer, ligation of target vector and B0014 insert resulted in a scar sequence that could not be cleaved by either of the original restriction enzymes.^[6] Subcloning was performed by C. Wokurek and the detailed procedure is given in G VII.1.5.1.^[519]

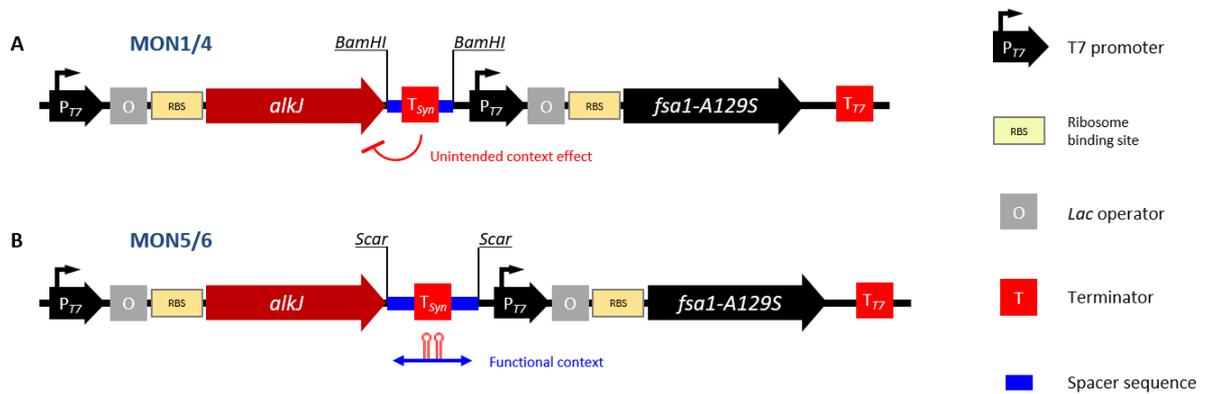


Figure E-1. Contextuality of the *alkJ* gene and a synthetic terminator sequence. (A) The monocistronic configurations MON1 and MON4 contain T_{Syn} sequences (B0011 and B0014) between the *alkJ* and the *fsa1-A129S* coding regions with short spacers of 16 bp and 26 bp, respectively. *AlkJ* activity is strongly impaired in these configurations. (B) Increasing the spacer sequences to 50 bp and 100 bp in MON5 and MON6, respectively, provides a context for functional expression of *AlkJ*. Length of ORFs does not reflect gene length.

Insertion was confirmed by *NcoI* control digestion after isolation of plasmid DNA from putative positive clones after transformation (Figure G-42). Competent *E. coli* BL21(DE3) cells were transformed with the newly assembled plasmids. SDS-PAGE analysis confirmed the production of *AlkJ* and *Fsa1-A129S* (Figure E-2).

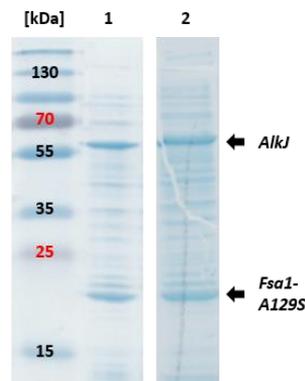


Figure E-2. Coproduction of *AlkJ* and *Fsa1-A129S* from pMON5 and pMON6. SDS-PAGE analysis of whole cell samples of *E. coli* BL21(DE3) transformants harboring the MON5 (1) or the MON6 plasmid (2). Protein production described in G VII.1.5.2. Sample loading normalized to $OD_{590} = 7.0$.

Subsequently, RCs were prepared and *AlkJ* functionality was tested with **2b** and **4b** under standard screening conditions (Figure E-3E–F). RCs expressing *AlkJ* from pPOP were used as assay control (Figure E-3A) and RCs producing inactive *AlkJ_{trnc}* (see E I.1) as negative control (Figure E-3B).

In assay controls, **2b** and **4b** were oxidized to the corresponding aldehydes and further to **2d** and **4d** as expected, whereas *AlkJ_{trnc}*, which misses a 774 bp portion close to the N-terminus (H I.1.10.4) and is inactive toward primary aromatic alcohols tested in this thesis (B I), did not oxidize **2b** and **4b**. *AlkJ* produced in the context of pMON1 and pMON4 did not or poorly convert the tested alcoholic substrates. *AlkJ* activity could be restored in the context of pMON5 and pMON6. *E. coli* BL21(DE3) RCs harboring the MON6 plasmid oxidized **2b** and **4b** as efficiently as the assay control (**Figure E-3B** and **Figure E-3F**, respectively). This suggests that, indeed, physical interactions and/or compositional effects between the 3'-end of the *alkJ* gene and B0014 translate into the production of a dysfunctional *AlkJ* protein.

In summary, the unexpected reduction of *AlkJ* activity when produced from pMON1 and pMON4 could be traced back to an unintentional context effect arising from spatially colocalizing synthetic genetic elements (**Figure E-1A**). The expansion of spacer sequences flanking the artificial B0014 terminator probably prevented local interactions between the *alkJ* gene and B0014 (**Figure E-1B**), thereby, restoring *AlkJ* activity in *E. coli* BL21(DE3) as determined by the oxidation of the standard substrates **2b** and **4b**.

The presence of T_{Syn} sequences had a negligible impact on expression levels of the downstream *fsa1-A129S* gene according to SDS-PAGE analysis (**Figure D-28**) and the quantification of soluble *Fsa1-A129S* (**Table D-4**). Future studies will determine the performance of the mini-pathway in monocistronic configurations toward the production of target aldol adducts. The example above illustrates that local context dependency can have far-reaching impacts on *de novo* pathway performance. To ensure the functionality of designed genetical devices, the elucidation of context effects needs to be integrated into the characterization of biological parts these devices are made of and will remain a challenge in synthetic biology.^[490, 520]

E I.2 Cofactor recycling for *in vitro* (and *in vivo*) cascades by the engineered *GDH_{2xBs}*

Industrial biocatalytic processes regularly employ cofactor dependent enzymes (e.g., ADHs, oxidases) to carry out synthetically useful but demanding reactions.^[6, 30, 316] Such applications only became feasible by the implementation of cofactor recycling systems. GDH (EC 1.1.1.47) is commonly used to regenerate NAD(P)H from the oxidized precursors NAD(P)⁺ with concomitant oxidation of inexpensive glucose^[521] (e.g., **Figure C-22** and **Figure C-24**) due to the high specific activity toward NAD(P)⁺.^[12] As nictotinamide cofactors are expensive, industrial (*in vitro*) application of cofactor recycling enzymes are mainly driven by cost.^[332]

Upon the introduction of artificial pathways performing redox transformations in living cells, the host cell supplies and regenerates cofactors.^[6] However, the utilization of carboxylic acids as substrates in the CRAS strain, for example, might not only be impaired by an unbalanced pathway enzyme stoichiometry; the ATP- and NADPH-dependent reduction of carboxylates by *CAR_{Ni}* (**Figure C-36**)^[64] might be simply too demanding at initially high carboxylate loading (see E I.4).

To address this potential bottleneck, the characterization of an engineered GDH from *B. subtilis* (*GDH_{Bs}*) will be described in the following. As noted above, GDH is an efficient NAD(P)H regenerator and widely applied in redox biocatalysis.^[316] Moreover, the hands-on production and purification of a robust cofactor regeneration system can certainly reduce costs in the laboratory of the Mihovilovic group and make biocatalytic (*in vitro*) transformations independent from the supply of commercial GDHs.

E I.2.1 Purification and characterization of two thermostable GDH mutants

GDH candidates were the double mutated *GDH_{2xBs}* (E170K and Q252L) and the seven times mutant *GDH_{7xBs}* (P45A, N46E, F155Y, E170K, V227A, W230F, and Q252L) from *B. subtilis*. The Bommarius group kindly donated the pET28a(+) plasmids containing the corresponding genes. Both engineered GDHs were reported to exhibit

increased thermostability and organic solvent tolerance compared to the wild type *GDH_{Bs}*. Both GDH mutants contained an N-terminal 6xHis tag for subsequent purification (H 1.1.7).^[332]

The GDH mutants were expressed in *E. coli* BL21(DE3) transformants according to Vázquez-Figueroa *et al.* without modification of the enzyme production protocol (**Figure E-4**).

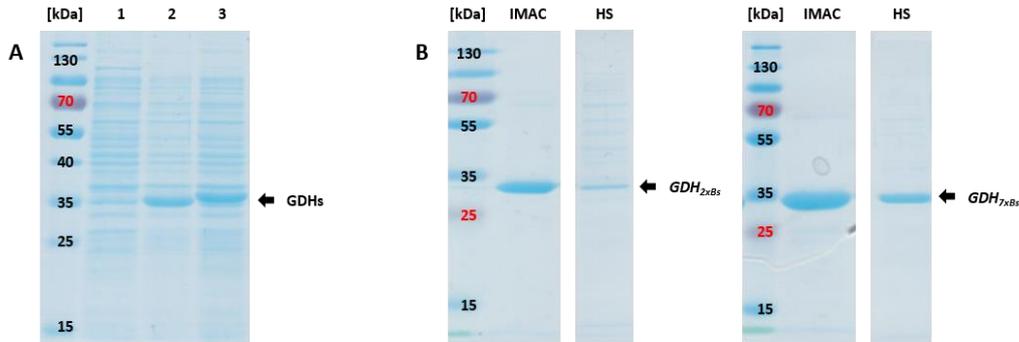


Figure E-4. Expression and purification of engineered GDHs from *B. subtilis*. (A) SDS-PAGE analysis of whole cell samples of untransformed *E. coli* BL21(DE3) (1), *GDH_{2xBs}* (2), and *GDH_{7xBs}* transformants (3). Protein production as in G VI.7.1. Sample loading normalized to OD₅₉₀ = 7.0. (B) Purification of both GDHs by HisTrap™ purification (IMAC) under unoptimized conditions or HS. Sample loading: *GDH_{2xBs}* and *GDH_{7xBs}* (5 µg) for IMAC and OD₅₉₀ = 7.0 for HS. (B) was adapted from P. Wolf (2017).^[522]

Although *GDH_{7xBs}* yielded higher amounts upon IPTG-induced expression according to SDS-PAGE, initial photometric NADP⁺ consumption assays showed reduced activity for *GDH_{7xBs}* after both purification by heat shock (HS) and immobilized metal ion affinity chromatography (IMAC) (**Figure E-5A**). *GDH_{2xBs}*, too, showed significantly lower activity after HS (**Figure E-5A**). Consequently, P. Wolf performed the optimization of isolation, purification, and the preliminary evaluation of storage conditions for *GDH_{2xBs}*.^[522]

In a nutshell, cell lysis by sonication as described in G I.5 in 50 mM Tris-HCl (pH 7.5) containing of 20% (w/v) glycerol was utile regarding yields and the activity of *GDH_{2xBs}*. Purification by IMAC and subsequent elution was also performed in the presence of glycerol as described in G VI.7.2.2. Combined fractions were snap-frozen in liquid N₂ and the samples lyophilized. For comparison, *GDH_{2xBs}* was lyophilized without glycerol as supplement. The resulting glycerol concentrates were stored at either -20°C or 4°C; the lyophilizate was stored at 4°C (**Figure E-5B**). For subsequent activity measurements, *GDH_{2xBs}* stocks were 1:5-diluted in 50 mM Tris-HCl (pH 7.5), the lyophilizate was dissolved in buffer, and the total amount of protein determined by Bradford assay (G I.7). Photometric assays were carried out in duplicates as described in G VI.7.3 (**Figure E-5B**).

Based on NADPH absorbance assay, glycerol assisted the preservation of *GDH_{2xBs}* activity during cell lysis, purification, and lyophilization.^[523-524] Lyophilization without glycerol as protectant resulted in a complete loss of activity. Contrary, glycerol concentrates could be stored at different temperatures and retained activity. The activity of *GDH_{2xBs}* was slightly higher in dilutions prepared from the glycerol concentrated stored at 4°C (**Figure E-5B**). Fresh dilutions were prepared from the concentrates after five days and activity compared to previously prepared 1:5 dilutions, which had been stored at 4°C and -20°C, respectively. The *GDH_{2xBs}* activity of dilutions stored at 4°C showed comparably high activities after five days, whereas freezing at -20°C and thawing reduced *GDH_{2xBs}* activity (data not shown). This suggests that storage of glycerol concentrates and dilutions (containing approximately 20% (v/v) glycerol) is superior to storage at -20°C.

To test the applicability of *GDH_{2xBs}* as cofactor regenerator in enzymatic reduction reactions, ene-reductase *YqjM*^[525] and *CAR_{Ni}* were produced in *E. coli* BL21(DE3) and *E. coli* BL21-Gold(DE3) transformants, respectively. *YqjM* was expressed from the pHT_*yqjM* plasmid^[526] and *CAR_{Ni}* (and the *Pptase_{Ec}*) from the previously

constructed pETDuet-1_*ppptase_{Ec}::car_{Ni}*.^[348] Protein production was performed according to G IV. Both enzymes contained N-terminal 6xHis tags for subsequent purification by IMAC as described in the corresponding subsections.

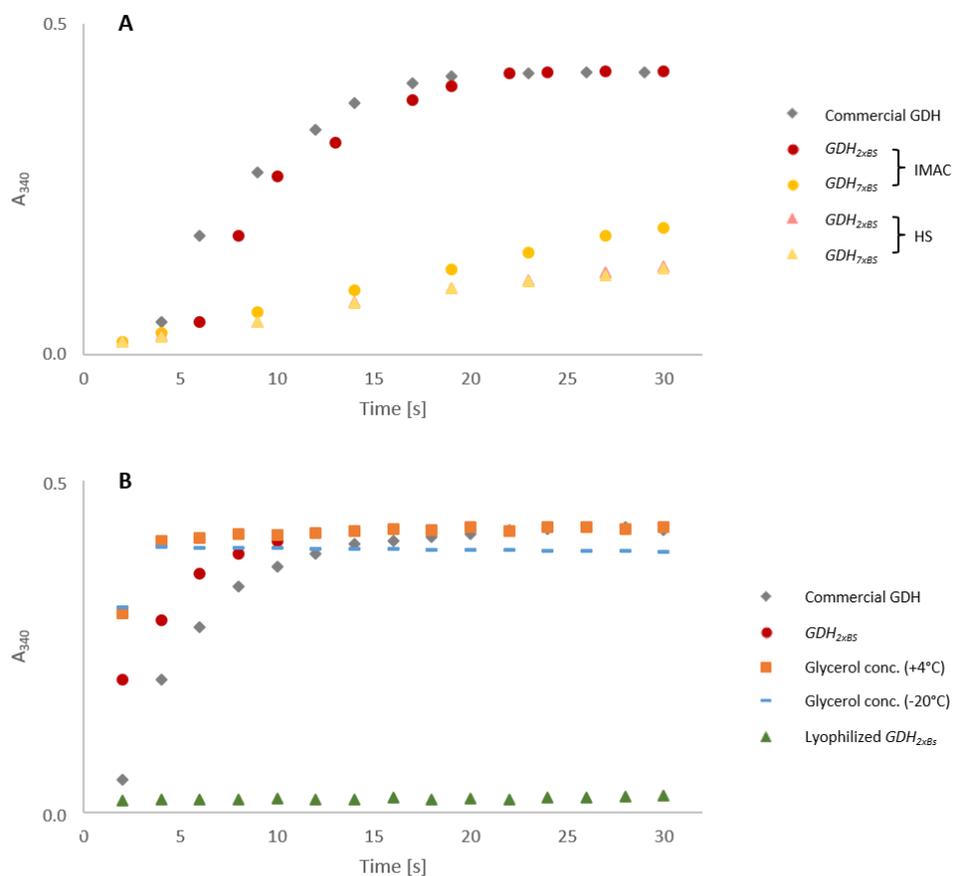


Figure E-5. Activity assays of differently purified GDHs. (A) Activity comparison of a commercially available GDH from *Pseudomonas* sp. (◆) and GDH mutants. GDH mutants were either purified by IMAC (●) or by HS (▲). The GDH_{2xBS} showed similar activity as the commercial GDH under experimental conditions. (B) GDH_{2xBS} was purified as before (●) without glycerol (▲) or in the presence of glycerol. Combined pure fractions were lyophilized. Resulting glycerol concentrates were stored at 4°C (■) or -20°C (-). Activity monitoring based on NADPH absorbance ($\lambda = 304$ nm; G VI.7.3).^[332] Results presented as mean values of duplicates. The figure was adapted from P. Wolf (2017).^[522]

E 1.2.2 Cofactor dependent reductions of double bonds and carboxylic acids *in vitro*

The asymmetric reduction of α , β -unsaturated C=C bonds creates up to two new chiral carbon centres and, thus, is an interesting transformation in organic synthesis. The biocatalytic addition of a hydride onto a C=C bond bearing an electronwithdrawing group (e.g., aldehydes, ketones, carboxylic acids and derivatives, or nitriles) is catalyzed by enoate reductases (EREDs), which belong to the old yellow enzyme (OYE) family. Enzymes in this group of NAD(P)H dependent flavoproteins enzymes are typically involved in oxidative stress responses.^[527-529]

YqjM from *B. subtilis* also belongs to the OYE family. It noncovalently binds the cofactor FMN, which is reduced by NADPH, and transfers electrons from the flavin α , β -unsaturated C=C bonds.^[528] Since cyclohexenone (**10**) and 2-methyl cyclohexenone (**12**) are known standard substrates for *YqjM*,^[528] they were subjected to reduction *in vitro*.

YqjM was successfully expressed, isolated in the presence of 0.1 mM FMN to assist protein folding, purified by IMAC, and concentrated (**Figure E-6**). Both substrates, **10** and **12**, were efficiently reduced *in vitro* in the presence of purified *YqjM* and *GDH_{2xB5}* for cofactor recycling and yielded >95% of cyclohexanone (**11**) and 2-methyl cyclohexanone (**13**), respectively, in 2 h reaction time according to GC/FID (**Figure E-7**). In control experiments employing a commercial GDH, **10** and **12** were reduced as fast (data not shown).

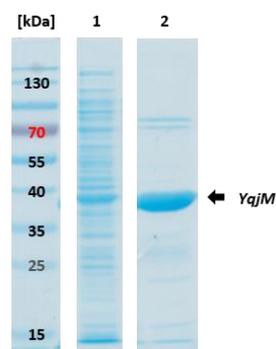


Figure E-6. Expression and purification of *YqjM*. SDS-PAGE analysis of CFEs from *E. coli* BL21(DE3) expressing *YqjM* from pHT_*yqjM* (**1**) and concentrate of *YqjM* after purification by IMAC (**2**). Protein production as in G VI.8.1. Purification of *YqjM* by HisTrap™. Sample loading normalized to 10 µg total amount of protein per lane.

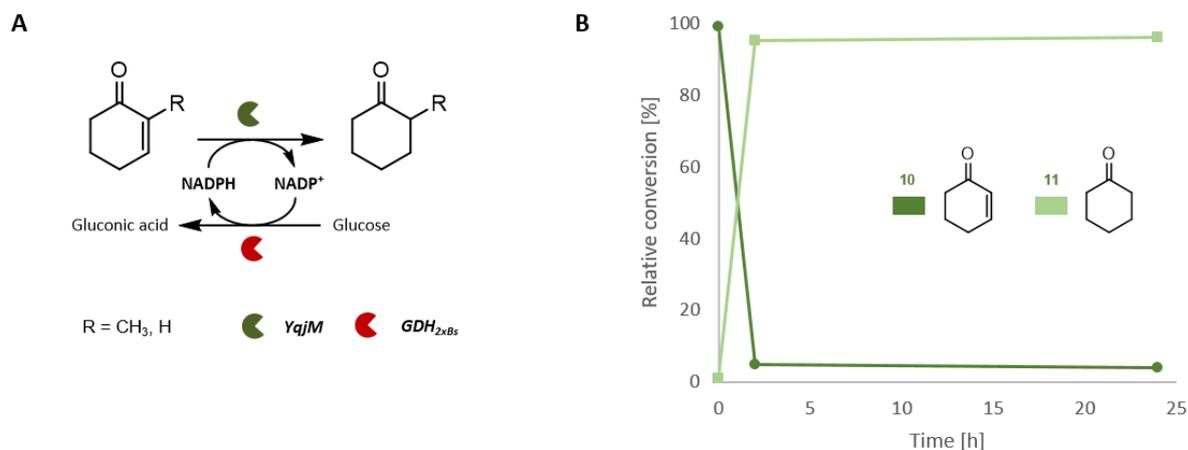


Figure E-7. *In vitro* reduction of C=C bonds by *YqjM*. (A) Reduction of the α , β -unsaturated compounds by *YqjM* to the corresponding saturated cyclohexanones with cofactor recycling. (B) Exemplary *in vitro* reduction of **10** to **11**; **12** to **13** performed equally efficient (data not shown). Screenings were carried out according to G VI.8.3 with purified *GDH_{2xB5}*. Control transformation with commercial GDH not shown for clarity. Results presented as mean values of duplicates.^[522]

The applicability of the recycling system could already be demonstrated in a more demanding reduction as described in D II.3 (**Figure D-13**). *CAR_{Ni}* was purified by IMAC and concentrated as described in G VI.3.1.2 (**Figure E-8**). *In vitro* reductions were carried with adaption of the protocol suggested by Finnigan *et al.*^[64] and NADPH cofactor regeneration performed as for the reduction of C=C bonds by *YqjM*. The carboxylates **1c**, a standard substrate for CARs, **2d**, and **8d** were tested. Whereas **1d** and **2d** were specifically reduced to the corresponding aldehydes **1c** and **2c**, respectively, **8d** was not under experimental conditions (**Figure D-13**).

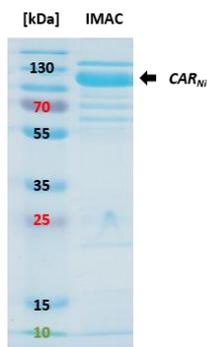


Figure E-8. Purification of CAR_{Ni} by Ni affinity chromatography. SDS-PAGE analysis of concentrate mainly containing CAR_{Ni} after purification. Bands of smaller sizes probably correspond to apo- CAR_{Ni} . Protein production in AIM (G VI.3.1.1). Purification of CAR_{Ni} by HisTrap™. Sample loading normalized to 10 μ g total amount of protein per lane. The figure was adapted from P. Wolf (2017).^[522]

Regarding the optimization of *in vivo* cascades employing CARs, for example, GDH_{Bs} and the double mutant offer potential to achieve a redox neutral reaction set up. However, the coproduction of CAR_{Ni} , $PPTase_{Ec}$, and GDH_{2xBs} was not feasible since target genes are encoded by vectors containing the same ORI, hence, belong to the same incompatibility group (Table C-1).^[6] The gdh_{2xBs} gene could be easily subcloned into the multiple cloning site (MCS)-1 of pCDFDuet-1, for example, utilizing *NcoI/HindIII* restriction sites that were also used to construct the parent pET28a(+) $_{gdh_{2xBs}}$.

In summary, after the identification of suitable conditions to produce and store GDH_{2xBs} (Figure E-5), the engineered GDH was shown to be as efficient as a commercial GDH from *Pseudomonas* sp. to regenerate NADPH, driving the reduction of α , β -unsaturated cyclohexenones (Figure E-7) and aromatic carboxylic acids *in vitro* (Figure D-13).

E I.3 Utilization of host strains with engineered metabolic backgrounds

In the previous section, *YqjM* was applied for the efficient reduction of α , β -unsaturated compounds **10** and **12** *in vitro* (Figure E-7). However, major obstacles occur if similar reductions are performed *in vivo*.^[6, 141] This is due to endogenous enzyme activities that can interfere with the asymmetric reduction of C=C bonds and yield mixtures of chiral compounds with undesired configuration.^[527-529] For example, *S*-(+)-carvone (**14**) and *R*-(-)-carvone (**16**) were reduced to (2*R*,5*S*)-dihydrocarvone (**15**) and (2*R*,5*R*)-dihydrocarvone (**17**), respectively, by the background of *E. coli* BL21(DE3) (Figure E-9). The *nemA* gene encoding N-ethylmaleimide reductase (*NemR*) was suggested by N. Oberleitner as competing endogenous activity, diverging the pathway flux from synthetic cascades containing the EREDs xenobiotic reductase B (*XenB*) from *P. putida*^[44] or OYE-1 from *Saccharomyces carlsbergensis*.^[138, 530] To test the reductive activity of *NemR* toward **14** and **16**, the *nemA* single KO *E. coli* K-12 mutant was ordered from the Keio library.^[48] Subsequently, CFEs were prepared from *E. coli* BL21(DE3) and the K-12 mutant as described in G I.5. Biotransformations were performed according to G III.1. Indeed, **14** and **16** were not reduced in the presence of a slight excess of NADPH and the CFE from the *NemR* deficient mutant (Figure E-9). Based on this finding, *NemR* was confirmed as the competing enzyme activity and was knocked-out by N. Oberleitner employing the TargeTron® system (Figure C-16).^[44, 138]

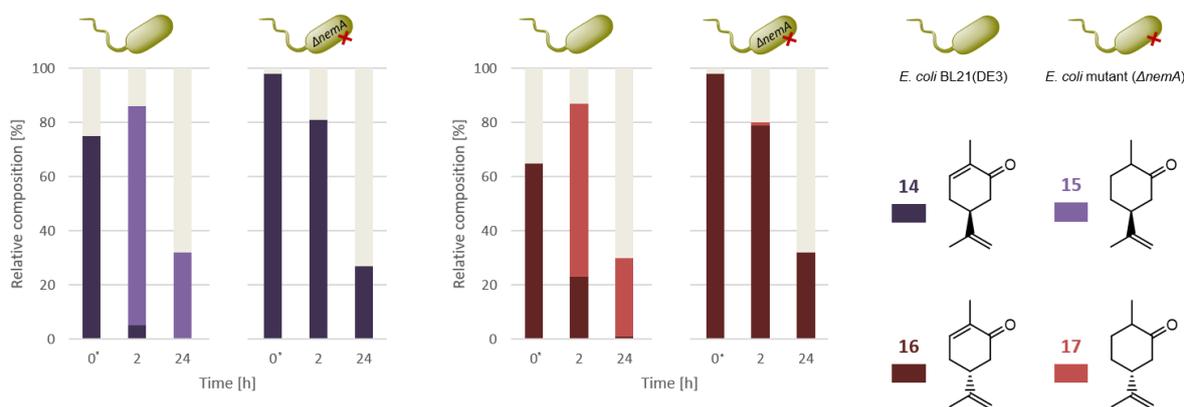
Optimization strategies for synthetic pathways *in vivo*

Figure E-9. The KO of *nemA* reduces the reduction of C=C bonds. The carvones **14** and **16** were not reduced in the *nemA* deficient mutant strain. Screenings were performed with CFEs from *E. coli* BL21(DE3) or *E. coli* K-12 $\Delta nemA$ as indicated; 5% (v/v) dioxane as cosolvent. Results presented as mean values of two independent experiments monitored by GC/FID; t_0^* sample taken immediately after addition of substrate and mixing. Stacked bars add up to 100% and represent the sum of carvone (dark color), dihydrocarvone (light color), and not recovered material (grey). Loss of material due to insufficient mixing at t_0^* and compound volatility.

However, higher substituted cyclohexenones in α position (e.g., 2-benzyl cyclohexanone) were still reduced by the background of the *nemA* KO strain according to N. Oberleitner. The KO of another endogenous ERED, 2,4-dienoyl-CoA reductase (*DCR*), was necessary to optimize the host cell and to construct a strain with decreased enoate reducing activity.^[138]

To impede the endogenous reduction of (aromatic) aldehydes in *E. coli*, more than only two consecutive rounds of KOs were necessary, illustrating that the metabolic background of host cells can greatly interfere with the flux through *de novo* pathways.^[6, 42, 406] Regarding the mini-pathway consisting of *AlkJ* and *Fsa1-A129S*, a high excess of donor molecules (e.g., DHA) were inevitable to shift the equilibrium of the aldolase mediated aldol formation to the product side. At lower DHA concentrations, retro-aldol reaction was pronounced and formed the aldehyde intermediate again, which was subject to the reduction and oxidation to the corresponding alcohols and carboxylic acids, respectively. Therefore, these two competing endogenous transformations pulled on the retro-aldol reaction, ultimately, decreasing aldol yields (**Figure D-29B**).^[348]

Consequently, to study the influence of endogenous enzyme activities on the retro-aldol reaction and the related high excess of DHA to shift the equilibrium, the mini-pathway was transferred into the engineered RARE strain (**Figure E-10**).^[42]

To transform competent RARE cells, desired plasmids must pass through one round of intermediate transformation of *E. coli* DH5 α . The genome of *E. coli* K-12 MG1655, which led to the construction of the RARE strain, contains several genes responsible for the restriction of foreign DNA including *hsdR*, *mcrA*, *mcrBC*, and *mrr*. The restriction enzyme *EcoKI*, encoded by *hsdR*, attacks DNA that is not protected by adenine methylation at the appropriate recognition site 5'-AAC[N₆]GTGC-3'.^[531] *McrA*, *McrBC*, and *Mrr* are methylation-dependent systems that cleave only DNA that is methylated at specific positions.^[532-534] All three of the latter systems restrict DNA modified by CpG methyltransferases and do not restrict DNA modified at *dcm* sites.^[535] Additionally, *Mrr* does not cleave DNA modified at *dam*, *EcoKI*, or *EcoRI* sites.^[533-534] For example, the restriction and methylation systems (*r/m*) were deleted in *E. coli* BL21(DE3) (*r⁻/m⁻*), whereas *E. coli* DH5 α lacks *hsdR* but can still methylate plasmid DNA (*r⁻/m⁺*). Therefore, plasmids methylated post transformation of DH5 α cells will not be degraded after reisolation and transformation of the RARE strain (*r⁺/m⁺*).

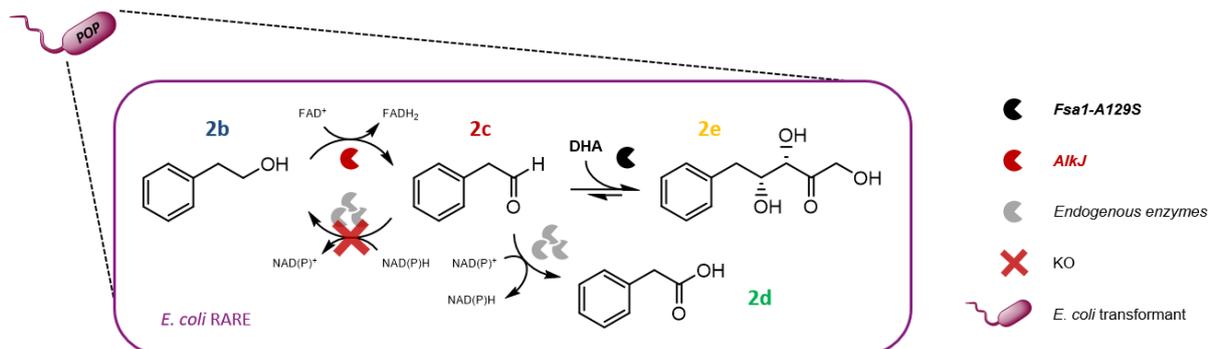


Figure E-10. The RARE strain in application. The metabolic background of the RARE strain exhibits decreased aromatic aldehyde reducing activity. By preventing the rapid reduction of **2c** to the substrate **2b**, thereby, increasing the intracellular concentration of **2c**, the equilibrium is thought to be shifted toward the target aldol **2e** at reduced DHA concentrations. *AlkJ* and *Fsa1-A129S* were coexpressed from the previously constructed POP plasmid.^[348, 435]

After successful transformation of the RARE strain with pPOP, *AlkJ* and *Fsa1-A129S* were coproduced and RCs prepared as before (**Figure E-11**). Biotransformations with the standard substrate **2b** were performed in the presence of different DHA concentration (0–100 mM; **Figure E-12B**).

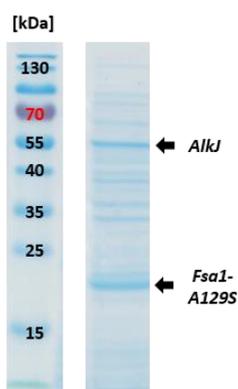


Figure E-11. Mini-pathway expression in the RARE strain. SDS-PAGE analysis of *E. coli* RARE whole cell samples expressing *AlkJ* and *Fsa1-A129S* from pPOP. Protein production in AIM (G VI.3.1.1). Sample loading normalized to $OD_{590} = 7.0$. The figure was adapted from T. Wiesinger (2017).^[435]

Although *E. coli* BL21(DE3) is a well-characterized and broadly applied host system for synthetic pathway expression, the metabolic environment (i.e., host context) can impair and redirect the carbon flux from target cascade compounds. In the first example, the confirmation of *NemR* as endogenous enzymatic activity reducing carvones (and other cyclohexenone derivatives; **Figure E-9**) led to the construction of a single ($\Delta nemA$) and a double ($\Delta nemA \Delta fadH$) KO mutant strain of *E. coli* BL21(DE3), showing minimized endogenous activity toward cyclic α, β -unsaturated C=C bonds.^[138]

Furthermore, the mini-pathway could be further optimized in the host context of the engineered RARE strain. In the host context of the unoptimized *E. coli* BL21(DE3), a high initial DHA loading (20 eq) was crucial to shift the equilibrium to the target aldol **2e** and achieve high yields of 89% according to calibrated HPLC analysis. DHA loading could be reduced 4-fold in the RARE strain under otherwise unchanged experimental conditions, yielding 86% of **2e** in only 2 h reaction time (**Figure E-12B**).

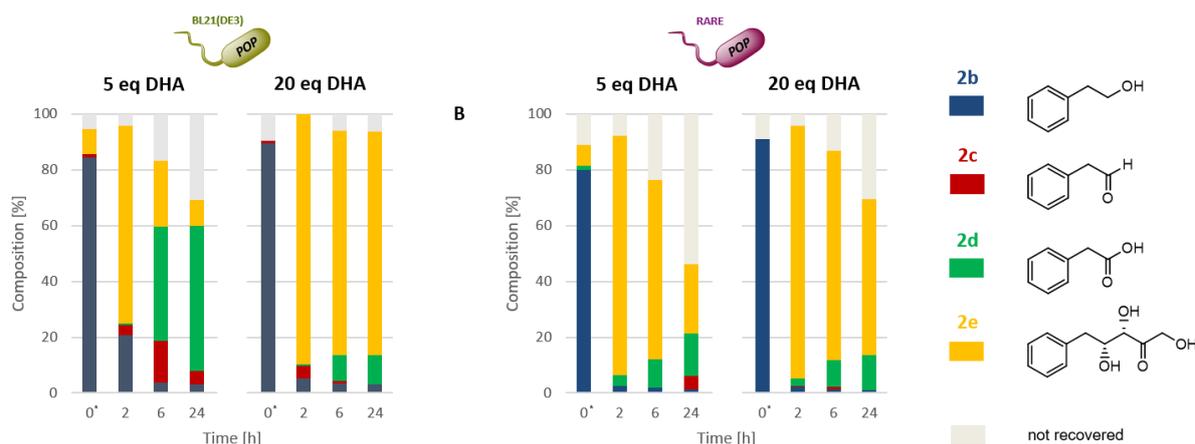


Figure E-12. Mini-pathway validation in the RARE strain. (A) *E. coli* BL21(DE3) RCs expressing *AlkJ* and *Fsa1-A129S* from the POP plasmid efficiently produced the target aldol **2e** in the presence of high DHA concentrations (20 eq = 100 mM) in 2 h reaction time. (B) RCs of *E. coli* RARE also expressing the mini-pathway. The DHA loading could be reduced to only 5 eq (25 mM), yielding comparably high yields of **2e** in 2 h. Screenings performed as usual; t_0^+ sample taken immediately after the addition of 5 mM **2b** and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <5% according to calibrated HPLC. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of **2b** (blue), **2c** (red), **2d** (green), aldol **2e** (yellow), and not recovered material (grey). The figure was adapted from T. Wiesinger *et al.* (2017).^[435]

In conclusion, the often-applied engineering of host cells has been a tool of great value for metabolic and microbial cell factory design. Although the multiple gene KOs in the RARE strain contributed to the decrease of the initial DHA load and to maximize target aldol titers, the many changes in the genome of host cells might affect the expression of neighboring loci and lead to undesired context effects. Reduced growth rates of the RARE strain in comparison to *E. coli* BL21(DE3), for example, might point toward such unpredictable consequences on the overall cellular system.

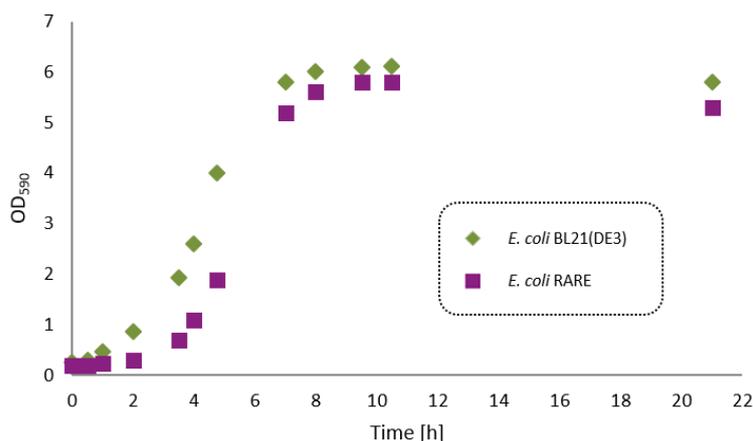


Figure E-13. The RARE strain and cellular growth. Unburdened *E. coli* RARE was characterized by significantly prolonged lag phase compared to *E. coli* BL21(DE3). The engineering of the genomes by directed KOs might lead to context effects that affect growth and other vital processes. Cellular growth studies were performed in baffled flasks in duplicates. Cultivation conditions: Inoculation with 1% (v/v) preculture of LB-Miller medium, 37°C (200 rpm) for 21 h.

E I.4 A 'hidden reservoir' for reactive aldehyde species: Addressing a competing side reaction and the issue of aldehyde toxicity *in vivo*

A major obstacle in whole cell biocatalysis, the accumulation of aldehydes (as products), was overcome with the introduction of the genomically engineered RARE strain by the Prather group.^[42] Seven genes had to be disrupted to minimize the inherent reducing activity of *E. coli* toward (aromatic) aldehydes, which beneficially changed the genetic host context. However, the accumulation of aldehydes triggers complex cellular responses to counter the oxidative and electrophilic stress from the reactive carbonyl group including the upregulation of ALDHs, for example.^[280-281, 400] These enzymes catalyze the irreversible oxidation of aldehydes to the carboxylic acids, an obstacle that has not been addressed by directed knock-out of the genes involved.

In this work, a conceptually distinct but complementary strategy, formerly known as 'substrate funneling', was applied to enhance the pathway flux through the artificial mini-pathway consisting of *AlkJ* and *Fsa1-A129S* (Figure D-29A). *AlkJ* oxidizes primary alcohols (e.g., **2b**) to the corresponding aldehydes (**2c**), which undergoes aldol addition with the donor DHA to form target polyhydroxylated compounds (**2e**). At insufficient concentrations of the extracellularly added donor DHA to shift the equilibrium toward the desired aldol product, the reactive aldehyde intermediate was oxidized to the thermodynamically favored carboxylates (e.g., **2c**; Figure D-29B). To redirect the carbon flux from the undesired byproducts toward the aldehyde cascade intermediates, the previously characterized *CAR_{Ni}* (and *Pptase_{Ec}*) was introduced to access the carboxylic acid sink (shown for **2b** in Figure E-14).^[348]

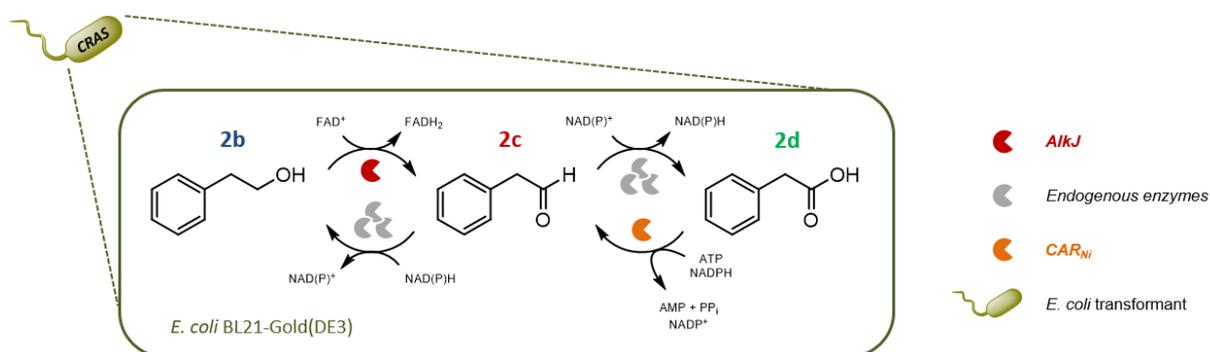


Figure E-14. An engineered strain for the containment of reactive aldehyde species (CRAS). *AlkJ* and *CAR_{Ni}* are supposed to shuttle the toxic aldehyde **2c** between different redox states, the primary alcohol **2b** and the carboxylate **2d**.

For proof of concept studies, competent *E. coli* BL21-Gold(DE3) cells were cotransformed with pKA1_alkJ::fsa1-A129S and pETDuet-1_pptase_{Ec}::car_{Ni}. Successful coexpression of *AlkJ*, *CAR_{Ni}*, *Pptase_{Ec}*, and *Fsa1-A129S* could be confirmed by SDS-PAGE analysis (Figure E-15A). *CAR_{Ni}* was readily expressed, whereas *AlkJ* and *Fsa1-A129S* showed reduced expression levels in comparison to pPOP transformants due to the increased metabolic burden (Figure E-15B). Subsequent screenings were performed in RCs under standard conditions and the conversion of the model substrates **2b** and **2c** monitored over time (0–24 h) in the absence of DHA (Figure E-16).

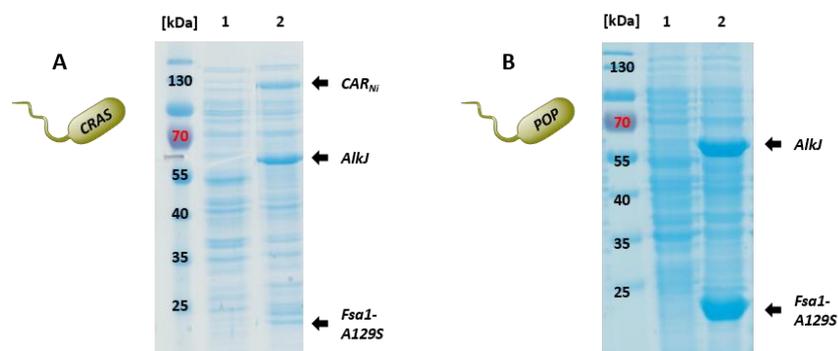


Figure E-15. Enzyme coproduction in the CRAS strain. (A) SDS-PAGE analysis of whole cell samples of untransformed *E. coli* BL21-Gold(DE3) (1) and the CRAS strain coexpressing *AlkJ*, *Fsa1-A129S*, and *CAR_{Ni}* (2). The *Pptase_{Ec}* is below detection limit under expression conditions as expected.^[348] (B) Whole cell samples of untransformed *E. coli* BL21-Gold(DE3) (1) and pPOP transformants coexpressing *AlkJ* and *Fsa1-A129S* (2). Protein coexpression according to G VIII.1. Sample loading normalized to OD₅₉₀ = 7.0. The figure was adapted from T. Bayer *et al.* (2017).^[348]

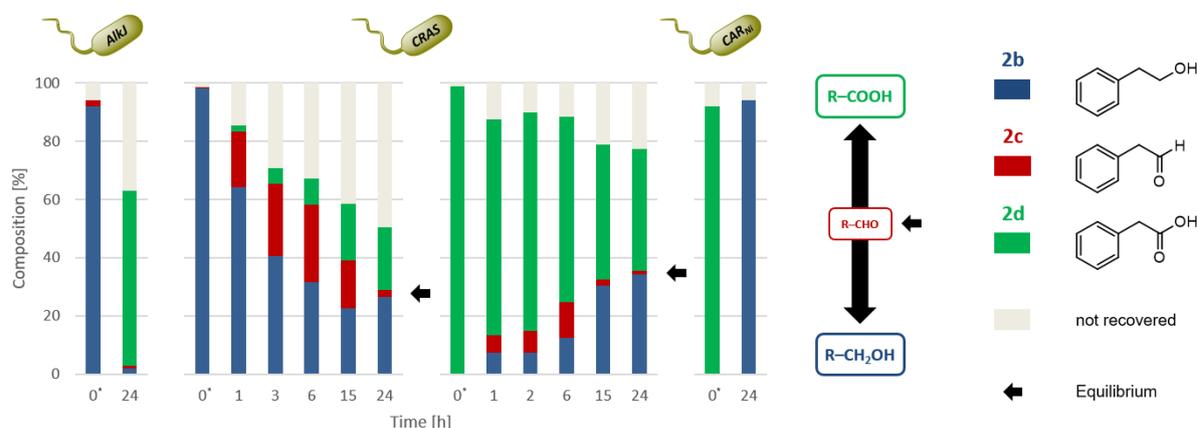


Figure E-16. The ‘hidden aldehyde reservoir’ in the CRAS strain. Single expression of *AlkJ* oxidized the alcohol substrate **2b** to the toxic aldehyde intermediate **2c**, which was further oxidized and accumulated as **2d** (left). *CAR_{Ni}* and *AlkJ* in the CRAS strain equilibrate **2a**, **2c**, and **2d** and formed a distinct redox equilibrium (black arrows; center). Single expression of *CAR_{Ni}* reduced **2d** to **2c**, which was rapidly metabolized to **2a** (right). Screenings performed under standard conditions; t_0^* sample taken immediately after the addition of 5 mM substrate and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <10% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material due to volatility of the aldehyde, for example. Parts of the figure were adapted from T. Bayer *et al.* (2017).^[348]

In previous biotransformations employing RCs expressing *AlkJ*, **2b** was oxidized to **2c**, which was irreversibly oxidized to the carboxylate **2d** mainly by endogenous host enzymes (Figure D-6). Contrary, RCs expressing *CAR_{Ni}* (and the accessory *Pptase_{Ec}*) efficiently reduced **2c** to **2b**, which was rapidly converted to **2b** by the metabolic background in response to the reactive carbonyl group in the aldehyde **2c** (Figure D-12).^[42, 281, 400, 404] In the presence of both enzymes, *AlkJ* and *CAR_{Ni}*, the primary alcohol **2b** was oxidized to corresponding aldehyde **2c** by *AlkJ*. The subsequent overoxidation to the carboxylic acid **2d** was successfully reversed by *CAR_{Ni}*. Endogenous *E. coli* enzyme activities reduced **2c** again to **2b**, which, ultimately, resulted in a redox equilibrium between **2b**, **2c**, and **2d** (Figure E-16). The system was balanced to a minimal but detectable amount of highly reactive aldehyde **2c** even after 24 h reaction time. The same effect was observed starting with the carboxylate **2d**. Again, the redox equilibrium was formed after 15 h (Figure E-16).^[348] To confirm the efficient formation of the ‘aldehyde

reservoir' in the CRAS strain, structurally different primary aromatic alcohols (**1b** and **4–5b**) and carboxylic acids (**1d** and **4–5d**) were tested and the balance between the primary alcohol, aldehyde, and carboxylate monitored as before (Figure E-17). Since *AlkJ* and *CAR_{Ni}* showed overlapping aldehyde production profiles for the tested substrates (Figure D-6 and Figure D-12, respectively), distinct redox equilibria were formed. Equilibria reflected the different activity of *AlkJ* and *CAR_{Ni}* toward the alcohol and carboxylic acid substrates, respectively (Figure E-17A). The compositions of the reaction mixtures after 24 h are also given in Table E-1.

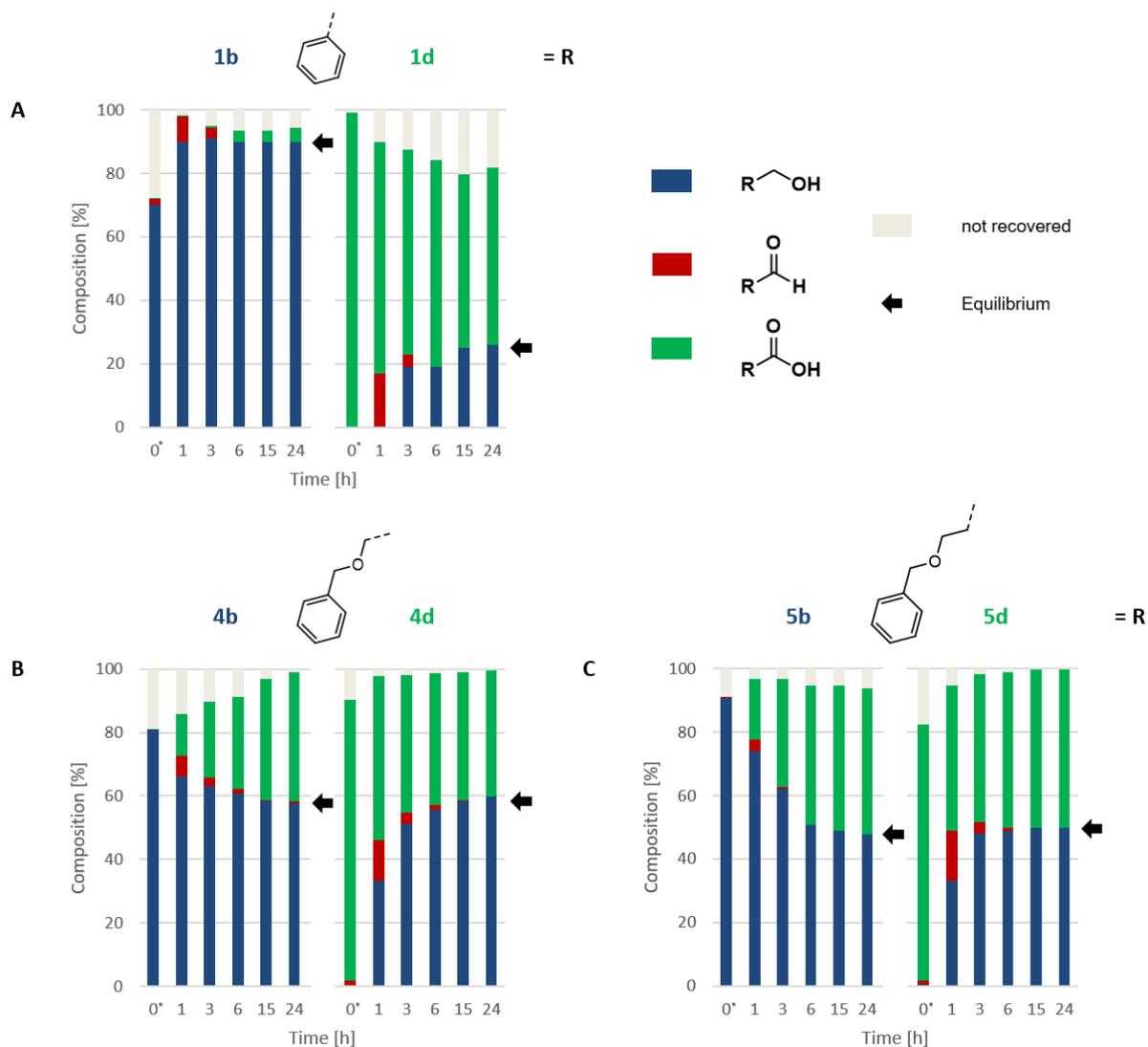


Figure E-17. 'Hidden reservoirs' of structurally different aldehydes. *AlkJ* and *CAR_{Ni}*, an enzyme pair with opposing functional group activity, established and maintained a redox equilibrium between primary alcohols, aldehydes, and carboxylic acids *in vivo*: (A) **1b** and **1d**, (B) **4b** and **4d**, and (C) **5b** and **5d**. Screenings performed in CRAS strain RCs under standard conditions; t_0^* sample taken immediately after the addition of 5 mM substrate and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <5% according to calibrated GC/FID. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of alcohol (blue), aldehyde (red), carboxylic acid (green), and not recovered material (grey). Reduced recovery of material at t_0^* due to insufficient mixing. Parts of (B) and (C) were adapted from T. Bayer *et al.* (2017).^[348]

Table E-1. ‘Hidden aldehyde reservoir’ compositions after 24 h

Substrate	Equilibrium composition [%]		
	Alcohol	Aldehyde	Carboxylic acid
1b	88.4±5.1	n.d.	6.1±1.3
1d	24.5±3.7	n.d.	57.3±4.8
2b	26.4±7.8	2.4±1.6	21.6±6.8
2d	34.4±6.7	1.0±1.4	42.1±7.9
4b	57.4±1.1	0.2±0.1	42.0±0.2
4d	58.7±0.7	n.d.	41.3±0.3
5b	48.0±2.0	n.d.	44.5±0.9
5d	49.8±0.7	n.d.	50.2±0.2

Results presented as mean values of triplicates ($n = 3$) \pm SD measured by calibrated GC/FID. Differences to 100% depict loss in mass that could not be recovered due to aldehyde reactivity and/or volatility. Results for substrates **2** are shown in **Figure E-16** and **1, 3–4** in **Figure E-17**; n.d. = not detected.

To study the influence of CAR_{Ni} on the mini-pathway performance, aldehyde **2c** was removed from the redox equilibrium by the *Fsa1-A129S* mediated cascade reaction (**Figure E-18**).

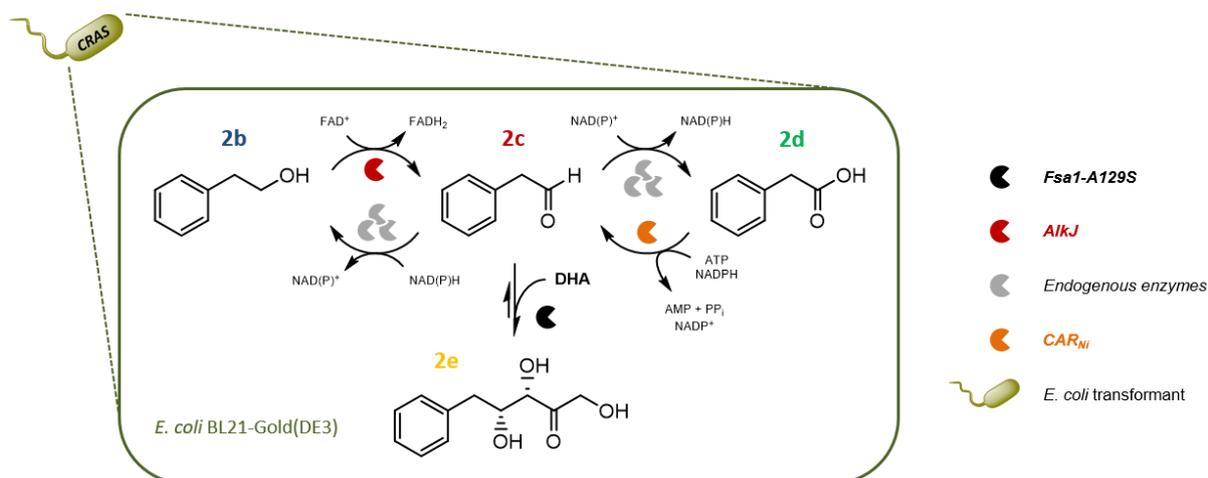


Figure E-18. Draining the ‘hidden aldehyde reservoir’ in a cascade-type aldol reaction. *AlkJ* and CAR_{Ni} contain the reactive aldehyde **2c**, which is withdrawn from the equilibrium between **2a** and **2d** by the cascade-type aldol reaction catalyzed by *Fsa1-A129S*. In the presence of DHA, the nontoxic polyhydroxylated product **2e** is formed.^[348]

To tap the aldehyde reservoir and pull **2c** toward the nontoxic polyhydroxylated aldol product **2e**, DHA was added in different concentrations (**Figure E-19**). First, in the absence of DHA, the enzyme pair formed the distinct redox equilibrium between **2b**, **2c**, and **2d** within 6 h reaction time. Second, in the presence of DHA (5 eq), the target aldol product **2e** was produced but the amount was insufficient to shift the aldol formation equilibrium completely. However, *AlkJ* and CAR_{Ni} established the redox equilibrium at prolonged reaction times. In the presence of a high DHA excess (20 eq), the desired aldol product **2e** was efficiently produced with the highest detectable amount (90.0±0.7%) after 6 h according to calibrated HPLC analysis (**Figure E-19**).^[348]

Although the application of the four-enzyme system, consisting of *AlkJ*, CAR_{Ni} , *Pptase_{Ec}*, and *Fsa1-A129S*, could be successfully demonstrated, the overall cascade performance was slower in the four-enzyme system than in the previously evaluated mini-pathway system, which yielded up to 95% in only 2 h (**Figure D-29**). This was mainly

attributed to the additional metabolic burden arising from the coproduction of CAR_{Ni} . This resulted in decreased amounts of the aldolase and an unbalanced overall enzyme expression. This is also depicted in the maximal aldol production of only 20% after 6 h according to calibrated HPLC when starting the cascade from the carboxylate **2d** (Figure E-21B).^[348]

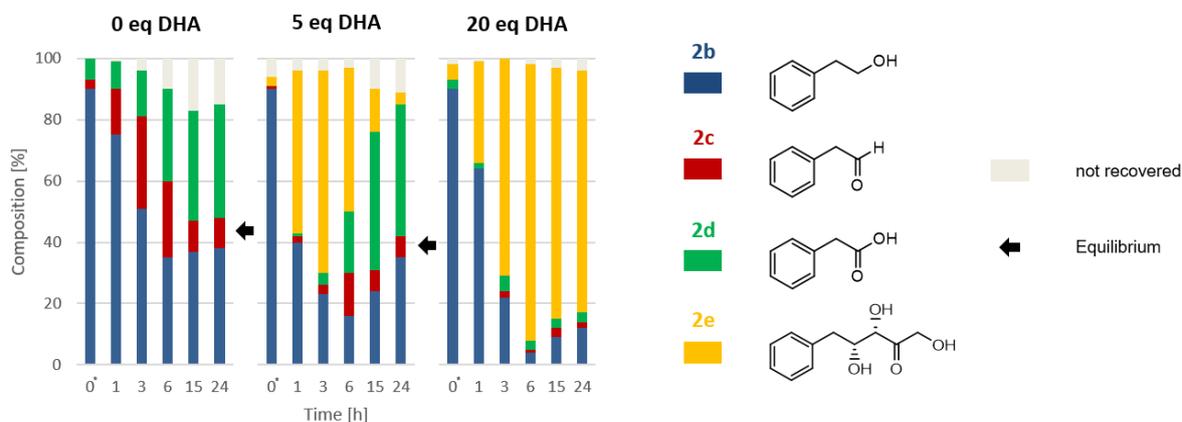


Figure E-19. Aldol production from the intracellular aldehyde reservoir. To RCs of the CRAS strain, **2b** was added. The production of **2e** was monitored at varying DHA concentrations (0–20 eq = 0–100 mM) over time. In the absence of DHA or at low concentrations (5 eq), *AlkJ* and CAR_{Ni} formed the distinct redox equilibrium after prolonged reaction times. An excess of DHA (20 eq) yielded up to 90% **2e** in 6 h. Screenings performed under standard conditions; t_0^* sample taken immediately after the addition of 5 mM **2b** and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <5% according to calibrated HPLC. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of **2b** (blue), **2c** (red), **2d** (green), aldol **2e** (yellow), and not recovered material (grey). The figure was adapted from T. Bayer *et al.* (2017).^[348]

To prove that insufficient yields of *Fsa1-A129S* represented the bottleneck, the same RC experiment was performed but with compensation of the lack of *Fsa1-A129S*. Therefore, *Fsa1-A129S* was purified by HS as before. The lyophilized catalyst was added to the reaction mixture and the formation of aldol was followed by HPLC analysis. Prior to HPLC measurements, extracellular *Fsa1-A129S* was precipitated as described in G III.2. Since all pathway intermediates diffuse through the cell membrane, a significant improvement in the cascade performance was observed. After only 2 h reaction time, up to 90% of desired aldol product **2e** were obtained according to calibrated HPLC (Figure E-20).

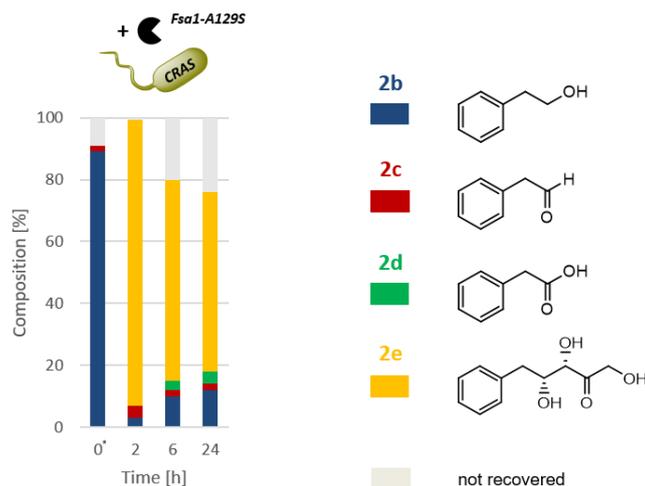


Figure E-20. Enzyme balancing by the extracellular addition of *Fsa1-A129S*. Compensation of the lack of *Fsa1-A129S* in the CRAS strain by the extracellular addition of the purified biocatalyst, yielding up to 90% of the target aldol **2e** in 2 h reaction time. Screenings performed in RCs under standard conditions; t_0^* sample taken immediately after the addition of 5 mM **2b** and mixing. Results presented as mean values of biological triplicates ($n = 3$); SD <5% according to calibrated HPLC. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of **2b** (blue), **2c** (red), **2d** (green), aldol **2e** (yellow), and not recovered material (grey). The figure was adapted from T. Bayer *et al.* (2017).^[348]

The results shown in **Figure E-20** confirmed the hypothesis that the *Fsa1-A129S* expression level was too low in the four-enzyme cascade. Moreover, the reversing enzyme activity of CAR_{Ni} counteracted retro-aldol formation and reduced carboxylate byproduct formation about 2-fold in comparison RCs expressing the mini-pathway and employing 20 eq of DHA after 24 h (**Figure D-29** and **Figure E-19**).

Since the engineered CRAS strain exhibits carboxylic acid reducing activity conveyed by CAR_{Ni} , it provides flexibility regarding the choice of substrate class to synthesize target aldol products. Consequently, **2b** was substituted by **2d** in RC experiments employing standard screening conditions. Unfortunately, **2d** (5 mM) was only slowly converted to the aldehyde intermediate **2c** and the aldol adduct **2e** was not produced under these conditions. Given that the CRAS strain insufficiently produces the aldolase, immediate attempts included the reduction of the overall metabolic burden by truncation of the *alkJ* gene in the POP plasmid and, as described above, the addition of *Fsa1-A129S* lyophilisate (**Figure E-21**).

The construction of $pKA1_alkJ_{trnc}::fsa1-A129S$ is described in G VI.2.4.3. Sanger sequencing (H I.1.10.4), subsequent SDS-PAGE analysis (**Figure E-21A**), and abolished oxidation of **2b** and **4-5b** in RCs (**Figure E-3B**; data not shown for **5b**) confirmed the truncation of *AlkJ*.

In the presence of DHA (20 eq), RCs expressing CAR_{Ni} , *Fsa1-A129S*, and $AlkJ_{trnc}$, only produced 20% **2e** when starting with 5 mM **2d** according to calibrated HPLC (**Figure E-21B**). Although the disruption of the *alkJ* gene produced inactive $AlkJ_{trnc}$, the amount of *Fsa1-A129S* increased *in vivo* according to SDS-PAGE analysis (**Figure E-21A**). However, this simple approach clearly did not suffice to balance overall pathway enzyme stoichiometry. Since **2d** was converted by CAR_{Ni} at slower rates compared to the unburdened system (**Figure D-12**), results suggest that the amount of the holo-CAR enzyme must be tuned, for example, by introducing other genetic regulatory elements or changing the genetic context to improve aldol titers.^[6]

The extracellular addition of aldolase increased **2e** yields to 49% after 2 h under otherwise unchanged screening conditions. In control experiments employing *E. coli* BL21-Gold(DE3) RCs exclusively expressing CAR_{Ni} (and $PPTase_{Ec}$) and extracellular *Fsa1-A129S*, the highest HPLC yield of 70% **2e** was detected after 6 h reaction time (**Figure E-21B**).

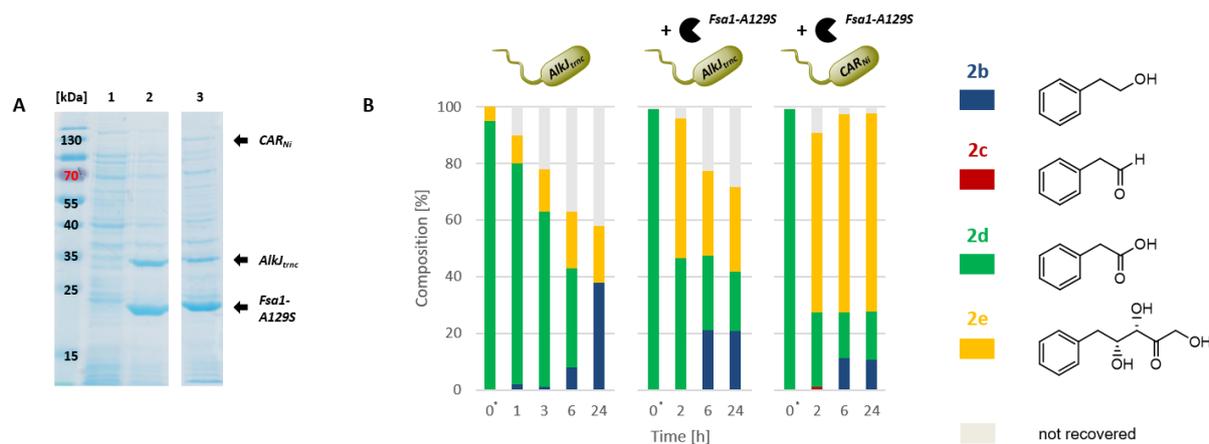


Figure E-21. From carboxylic acids to polyhydroxylated compounds. (A) SDS-PAGE analysis of untransformed *E. coli* BL21-Gold(DE3) whole cell samples (1), cells producing *AlkJ_{trnc}* and *Fsa1-A129S* from pKA1_alk_{trnc}::f_{sa1}-A129S (2), or *AlkJ_{trnc}*, *Fsa1-A129S*, *CAR_{Ni}*, and *Pptase_{Ec}* (3). Protein coproduction in AIM (G VI.3.1.1). Sample loading normalized to OD₅₉₀ = 7.0. (B) Production of target aldol **2e** in RCs harboring pETDuet-1_pptase_{Ec}::car_{Ni} and pKA1_alk_{trnc}::f_{sa1}-A129S (*AlkJ_{trnc}* strain) or only the pETDuet-1 vector with or without the addition of extracellular aldolase. RC screenings performed under standard conditions; t₀* sample taken immediately after the addition of 5 mM **2d** and mixing. Results presented as mean values of biological triplicates (n = 3); SD <5% according to calibrated HPLC. Error bars omitted for clarity. Stacked bars add up to 100% and represent the sum of **2b** (blue), **2c** (red), **2d** (green), aldol **2e** (yellow), and not recovered material (grey). (A) and parts of (B) were adapted from T. Bayer *et al.* (2017).^[348]

In summary, the modular expansion of the mini-pathway by introducing *CAR_{Ni}* and *Pptase_{Ec}*, which posttranslationally modifies apo-CAR to produce the holo-CAR enzyme, not only resulted in the redirection of the pathway flux from the carboxylic acid sink to the target aldehydes; beneficially, the opposing activities of *AlkJ* and *CAR_{Ni}* equilibrated the reactive aldehydes between the nontoxic primary alcohols and the carboxylates. This ‘biocatalytic reservoir’ contained the aldehyde intermediates below nonviable concentrations but freely accessible for subsequent aldol reaction (Figure E-18). ‘Chemical reservoirs’ (e.g., biphasic liquid-liquid systems)^[289, 411] or scavenger strategies (e.g., aldoximes)^[412] have been applied to address aldehyde toxicity *ex vivo* but depended on additional organic solvents and chemicals. The presented approach and the engineered CRAS strain greatly enhanced aldehyde persistence *in vivo* and aldehyde containment increased cell viability, simply by the introduction of an additional enzyme activity. Furthermore, *CAR_{Ni}* reduced the formation of carboxylic acid byproducts post retro-aldol reaction. Despite the unbalanced pathway enzyme stoichiometry, a final preparative scale experiment starting from **2b** (5 mM; 20 eq DHA, V_{total} = 0.2 L) yielded 90% **2e** according to calibrated HPLC in 6 h reaction time (Figure E-19). By applying a SPE purification protocol (see G IX), 70% pure aldol (>99% *de*)^[536] could be obtained after isolated isolation. The longer reaction time of 6 h (see Figure D-29 for comparison with the mini-pathway) was a consequence of imbalanced enzyme production, mainly *Fsa1-A129S*, which could be compensated by the extracellular addition of the purified biocatalyst.^[348]

Concluding, the synthetic enzyme cascade consisting of four enzymes (i.e., *AlkJ*, *CAR_{Ni}*, *Pptase_{Ec}*, and *Fsa1-A129S*) successfully produced the target aldol products. The flux through the *de novo* pathway could be enhanced by the reversion of carboxylate byproduct formation by *CAR_{Ni}*. Together with *AlkJ*, this enzyme pair with opposing functional group activity resulted in the establishment of a reservoir for highly reactive aldehyde cascade intermediates. Since intracellular aldehyde concentrations could be kept below nonviable levels, this approach offers an immediate solution to address aldehyde toxicity in living cells. In general, the containment of toxic pathway intermediates offers a complementary substrate channeling strategy independent from protein modifications such as the tagging of enzymes and the construction of synthetic scaffolds.^[6, 283] Moreover, this

approach can be utilized to reverse other byproduct-forming reactions and given that a reversing enzyme activity exists and recombinant expression is functional and not interfering with the host metabolism, it promises broader application.^[348]

E I.5 Once the microbial cell factory is done: Downstream processing by solid phase extraction (SPE) to maximize aldol product yields

The examples above show the many strategies available to tweak synthetic enzyme cascades and the host *E. coli* to efficiently produce (nonnatural) chemicals and provide the metabolic platform to do so. Genetic tools to modify pathway architecture, engineer the genetic context for enhanced target enzyme production, provide for an increased cofactor demand, or to knock-out competing endogenous have been applied to maximize product titers. These strategies optimize *de novo* pathway performance on different molecular levels and were successfully applied in this thesis. However, combined optimization efforts are rendered insignificant if the engineered microbial cell factory produces high yields of target molecules that cannot be isolated.

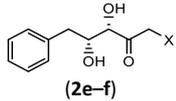
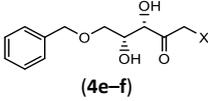
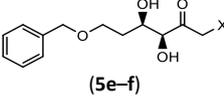
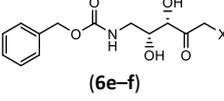
As summarized in D III.1.2, the newly assembled mini-pathway consisting of *AlkJ* and *Fsa1-A129S* produced various polyhydroxylated compounds in up to 95% HPLC yields. However, acyclic aldol compounds are not only water soluble, hence, difficult to isolate from aqueous media; acyclic aldol adducts are temperature sensitive.^[435] Consequently, aldol isolation and purification represented the yield limiting step toward an efficient production (**Table D-5**).^[293, 436, 444] The downstream processing was refined by T. Wiesinger and led to the establishment of an easy to apply solid phase extraction (SPE) protocol.^[435, 487]

Briefly, to identify a suitable method to isolate polyhydroxylated compounds, *E. coli* BL21(DE3) RCs expressing *AlkJ* and *Fsa1-A129S* from the POP plasmid were prepared as before. Primary aromatic alcohols (**2b** and **4-6b**) were converted in the presence of monomerized DHA or HA. Reactions were performed under standard screening conditions and the reaction progress was monitored by complementary GC/FID and HPLC analysis. Biotransformations were terminated after full conversion of starting materials to the target aldol adducts (DHA: **2e** and **4-6e**; HA: **2f** and **5-6f**). RCs were centrifuged and the supernatants were subjected to isolation methods including the extraction with organic solvents such as CH₂Cl₂ or EtOAc, for example. Extraction was suitable for the purification of less hydrophilic HA aldol adducts (**2f** and **4-5f**) while failing for more water-soluble products like **2e** and **4-6e** (**Table D-5**).^[435, 487] Alternatively, preparative reversed phase (RP)-HPLC purification employing octadecyl carbon chain (C18)-bonded silica material resulted in moderate yields for the HA aldol products **2f** and **4-6f** but also gave access to the more hydrophilic DHA adducts **2e** and **4-6e**. Major limitations of the isolation by RP-HPLC rooted in both the concentration by lyophilization and the poor product solubility in the presence of buffer salts.^[435, 487] To separate buffer salts and other water-soluble components, SPE was performed as described in G IX. After transferring supernatants on short C18 columns, eluting buffer salts and solutes of the reaction medium with 5% (v/v) MeOH in water, aldol products were eluted in pure MeOH. The solvent was evaporated at reduced pressure yielding significantly higher amounts of up to 91% of the desired aldol adducts over two steps (**Table E-2**). Compared to the unoptimized downstream processing (**Table D-5**), the final yields in this work exceeded the isolated yields reported in the literature for single-step transformations.^[444, 508]

In summary, the SPE purification protocol is independent from elaborate sample preparations, which, regarding the temperature sensitivity of some aldol adducts, represents an advantage compared to other isolation methods.^[435] SPE can be easily applied and was performed to isolate the aldol adducts **2e**, **4-6e** and **2f**, **4-6f** in good to excellent yields (*de* >99%).^[435] Moreover, the use of MeOH as solvent is beneficial for potential transformations of **6e** toward D-fagomine since it is one of the most prominent solvents applied in catalytic hydrogenations.^[435] The eluate after the SPE purification could be directly used without further treatment.^[293, 537]

The established artificial mini-pathway is a showcase example that the optimization of product isolation/purification is equally important and cannot be entirely solved by tools from metabolic and genetic engineering.

Table E-2. Isolated yields of target aldol compounds with optimized SPE product isolation

Products	Substrate	Donor	Isolated yields after SPE [%] ^[a]	
			This work	Literature ^[b]
 (2e-f)	2b	DHA (e)	78 ^[c]	46 ^[432]
		HA (f)	70	48 ^[432]
 (4e-f)	4b	DHA (e)	60	28 ^[432]
		HA (f)	89	71 ^[432]
 (5e-f)	5b	DHA (e)	64	n.a.
		HA (f)	61	n.a.
 (6e-f)	6b	DHA (e)	91	79 ^[484]
		HA (f)	83	n.a.

^[a] Isolated compounds contain 10% water based on ¹H-NMR experiments in MeOH-*d*₄ and DMSO-*d*₆.^[435, 487]
^[b] Isolated yields from *in vitro* preparations employing *Fsa1-A129S* lyophilisates^[444] or *HL-ADH*, NOX, and *Fsa1-A129S*.^[508] X = OH (e) or H (f); n.a. = not available. ^[c] Production in the CRAS strain. Isolated yields adapted from T. Wiesinger *et al.* (2017).^[435]

E I.6 A difficult case: The optimization of de novo pathways employing DHAP-dependent aldolases

This subchapter summarizes the challenges of implementing DHAP-dependent aldolases (e.g., *FucA*) *in vivo* and addresses obstacles such as the insufficient amounts of intracellular DHAP.

E I.6.1 Consideration of potential bottlenecks

The production of **2e_{FucA}** from the primary alcohol **2b** in the AFucP strain (**Figure D-36**) and directly from the aldehyde **2c** in the FucP strain, respectively, failed under several screening conditions (**Table D-6**). Adaptions to the conditions of Wei *et al.*, who implemented a cascade *in vivo* consisting of the *FucA* from *Thermus thermophilus* HB8 and the *YqaB* phosphatase from *E. coli* (in the following referred to as FucY strain), did not result in the successful production of **2e_{FucA}** either.

The set-up employing the FucP strain was not feasible due to (i) the innately high toxicity levels of aldehydes^[280, 471] when used as substrates (**Figure D-37**) and (ii) indications that overexpressing the nonnative phosphatase *PhoN-Sf* from *S. fleroxneri* might be toxic for *E. coli*. Whereas the expression of *PhoN-Sf* slightly reduced cellular growth of host cells, the expression of *YqaB* did not (**Figure D-40**). Furthermore, downregulation of *PhoN-Sf* expression levels in periplasmic and cytosolic protein fractions were reduced at prolonged expression times (**Figure D-22** and **Figure D-23A**, respectively). Although these observations are not necessarily related to the toxicity of *PhoN-Sf*, Tanaka *et al.* examined the dephosphorylation of G6P catalyzed by *PhoN-Sf* (and *PhoN-*

S_{esp}).^[511] Although this study was not carried out under the physiological conditions of *E. coli* cytosol, the significantly lower $K_M = 0.02$ mM for G6P compared to $K_M = 5.3$ mM for glucose suggest a strong binding of the phosphorylated compound. Since the implementation of *PhoN-Sf* in the AFucP and the FucP strain and, consequently, the different metabolic host context might lead to undesired interferences between the heterologous enzyme and the dephosphorylation of metabolites of central carbon metabolism such as G6P.

Since Wei *et al.* demonstrated the applicability of *YqaB* for the irreversible dephosphorylation and the secretion of resulting aldol adducts out of the cell,^[293] *E. coli* BL21(DE3) was engineered accordingly and is described in the next section.

Another potential bottleneck already discussed in D III.2.2 is the aldehyde compounds to be converted.^[293] Wei *et al.* mainly converted water-soluble (and less toxic) aldehydes (e.g., 3-TFAP) and certainly took advantage of high reactant concentrations of up to 40 mM aldehyde to shift the equilibrium toward the aldol product side. In contrast, the concentration of aldehydes like **2c** must be below toxicity levels to keep host cells viable and the synthetic enzyme cascade running. The CRAS strain offers the possibility to produce **2c** from **2b** *in situ* by *AlkJ* and to reduce the generated **2d** byproducts into **2c** by *CAR_{Ni}*. The opposing functional group transformation activities contain the reactive aldehyde **2c** at viable concentrations (**Figure E-14**),^[348] which is certainly beneficial for the performance of the DHAP-dependent aldolase cascade.

Since increasing the aldehyde acceptor concentration for pushing the equilibrium of the aldol reaction toward the adduct side is clearly not an option, strategies to increase the intracellular DHAP donor concentration were considered. One simple enzymatic route to produce DHAP *in situ* is the ATP-dependent phosphorylation of DHA by the *DhaK* from *C. freundii*.^[293, 363, 447] Reviewing the intracellular concentration of DHAP of approximately 0.2 mM in *E. coli* in the stationary phase,^[512, 538] it can be expected to be a limiting reactant for aldol reactions *in vivo*. Furthermore, the K_M values of known DHAP-dependent aldolases are rather high ranging from 0.5 to 1.0 mM (e.g., 0.6 mM for *RhuA*).^[439] Thus, to increase the intracellular DHAP amounts is indispensable.^[154]

Consequently, the optimization strategy described in the following aims at the construction of an enzymatic ‘push and pull’ module consisting of *DhaK* and *YqaB* to feed the DHAP pool from extracellularly added DHA and to pull the *FucA*-mediated aldolase reaction toward the nonphosphorylated target compound **2e_{FucA}** by irreversible dephosphorylation, respectively.

E I.6.2 *DhaK* and *YqaB*: An enzymatic ‘push and pull’ module for pathway optimization

The plasmid encoding the *dhaK* gene was kindly provided by Prof. Garcíá-Junceda who is with the University of Madrid, Spain. *E. coli* BL21(DE3) was transformed with pRSETa_*dhaK* and *DhaK* successfully expressed as soluble protein according to Iturrate *et al.* (G VI.6.1.2; **Figure E-22**)^[363]

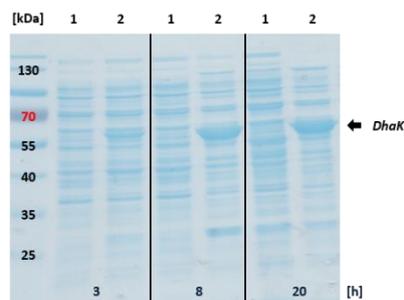


Figure E-22. Soluble expression of *DhaK* from pRSETa. SDS-PAGE analysis of CFEs (soluble fractions) from *E. coli* BL21(DE3) transformed with pET22b(+)_*dhaK*^[539] (1) and pRSETa_*dhaK*^[363] (2). Only pRSETa produced soluble *DhaK* over time (3–20 h). Protein production as in G VI.6.1. Sample loading normalized to OD₅₉₀ = 7.0.

Noteworthy, the preliminary production of *DhaK* from a pET22b(+) as published by Daniel *et al.*^[539] exclusively produced insoluble protein (G VI.6.1.1; **Figure G-25**). The gene was previously synthesized by GenScript™ and delivered in the target pET22b(+) vector (G VI.6.1.1). Sanger sequencing confirmed the integrity of the gene. However, comparison revealed differences in the nucleotide sequence in the pET22b(+)_*dhaK* construct that translate into a protein with 96% amino acid identity to pRSETa_*dhaK*. BLAST (basic local alignment search tool) revealed that the sequence published by Daniel *et al.* probably encodes the *cfaS* gene from *C. freundii*, which would produce a cyclopropane fatty acid synthase and not *DhaK* (H I.1.10.2; **Figure H-22**). Hence, the pET22b(+) construct was not further used in this thesis.

After expression, intracellular DHAP concentrations in untransformed *E. coli* BL21(DE3) and *E. coli* transformants expressing *DhaK* were determined in both exponential and stationary phase. Metabolomics measurements were performed in triplicates as outlined in G VI.6.2. Analysis and evaluation of results was done by S. Milker.^[154] In summary, the addition of 20 mM DHA after 5.8 h and another 10 mM DHA after 24.5 h showed that the *DhaK* producing strain showed an 4-fold increase in the intracellular DHAP concentration compared to the empty vector control in the exponential phase. Elevated DHAP concentrations were stable for 2 h before decreasing again. The untransformed strain also showed slightly increased DHAP concentrations after the addition of DHA, which can be explained by the metabolization of DHA through endogenous metabolic pathways. However, the DHA concentration in the cultivation medium remained constant at later time points, making the initial DHAP increase in *E. coli* BL21(DE3) neglectable. In both strains, DHAP concentrations did not increase during stationary phase.^[154]

Generally, the DHAP concentration during the stationary phase was up to 4-fold smaller compared to the exponential phase,^[154] which is disadvantageous for aldol reactions *in vivo* and supports the use of growing or LEDs to produce **2e_{FucA}** in biotransformations. As the transformation of DHA to DHAP by *DhaK* consumes ATP, the impact on the energy metabolism of host cells was elucidated. Therefore, intracellular ATP, ADP, and AMP concentrations were determined by metabolomic analysis and the energy charge calculated (**Figure E-23**).^[540] The energy charge is an index used to measure the energy status of biological cells and ranges from 0.8 to 0.95 for most cells.^[13]

$$\text{Energy charge} = \frac{[\text{ATP}] + \frac{1}{2}[\text{ADP}]}{[\text{ATP}] + [\text{ADP}] + [\text{AMP}]}$$

Figure E-23. Calculation of the energy charge. The energy charge determines the energy status of cells.

Figure adapted from Nelson *et al.* (2013).^[13]

Upon DHA addition, a difference in concentration was observed for all three energy metabolites in both strains for a short period of time and stabilized to 0.85 afterwards according to S. Milker.^[154] This is a good indication of a functional metabolism and the recycling of cofactors. Careful metabolomic analysis of intracellular ATP concentrations revealed that ATP (and DHA) concentrations were successively decreasing during bacterial growth. At the end of the growing phase, the ATP concentration dropped to $\approx 10 \mu\text{M}$.^[154] Considering the K_M value of *DhaK* for ATP of $70 \mu\text{M}$,^[539] the low ATP concentration might be insufficient to perform the phosphorylation under nongrowing conditions. This result further substantiates the need of growing (e.g., LEPs) cells to efficiently boost aldol formation *in vivo*.

Considering these bottlenecks, RbCl-competent *E. coli* BL21(DE3) was cotransformed with pKA1_ *fucA*,^[509] pCDFDuet-1_ *yqaB*, and pRSETa_ *dhaK*. The resulting strain will be referred to as FucYD strain in the following.

Coexpression of different enzyme combinations was achieved by adapting the fermentation medium of the Li group, who recombinantly coexpress up nine pathway enzymes in a single cell, and the expression conditions by Wei *et al.* as outlined in G VIII.5.^[43, 293] Whereas preliminary expression studies at an earlier stage of this thesis resulted in the successful coexpression of *FucA* and *DhaK*, for example (**Figure E-24A**), *DhaK* production from the pRSETa vector was abolished due to unknown reasons at later stages (**Figure E-24B**). Although Sanger sequencing confirmed the integrity of the *dhaK* sequence, expression could not be restored until the end of this thesis. To rule out mutations in the regulatory sequences (e.g., P_{T7} or RBS), pRSETa should be fully sequenced soon and, ultimately, requested again from our cooperation partner Prof. Garcíá-Junceda if the plasmid backbone is mutated.

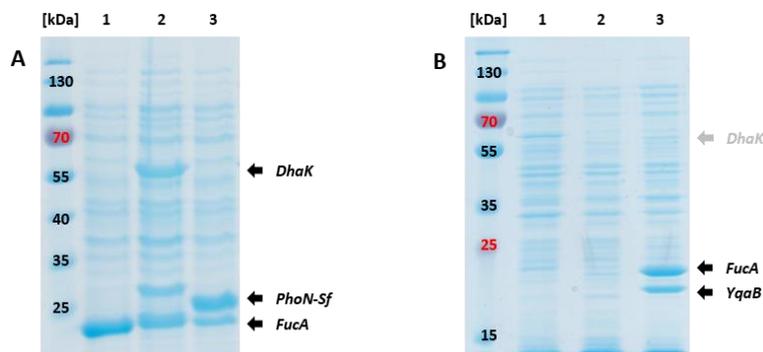


Figure E-24. Expression of *DhaK* and other pathway enzymes. SDS-PAGE analysis of CFEs from *E. coli* BL21(DE3) (A) coexpressing *DhaK* (*FucA* only, **1**; *FucA* + *DhaK*, **2**; *FucA* + *PhoN-Sf*, **3**) and (B) with abolished *DhaK* expression due to unknown reasons (untransformed *E. coli* BL21(DE3), **1**; *DhaK* only, **2**; *FucA* + *DhaK* + *YqaB*, **3**). Protein production for (A) according to G VIII.3 and for (B) G VIII.5.

Nevertheless, pathway validation was pursued without *DhaK* and employed a mixed cell approach of LEPs. The CRAS strain, coexpressing *AlkJ* and *CAR_{Ni}* (pKA1_ *alkJ* and pCDFDuet-1_ *PPTase_{EC::CAR_{Ni}}*), will produce the aldehyde *in situ* and keep intracellular concentrations viable in this strain and, supposingly, the overall system. Aldehyde intermediates need to pass the cell membrane of CRAS cells and enter the FucY strain, which, on the other hand, coexpresses *FucA* and *YqaB* (pKA1_ *fucA* and pCDFDuet-1_ *yqaB*). Experimental set-up and results will be discussed in the last section of this chapter.

E 1.6.3 There is always room for improvement: Unsuccessful production of **2e_{FucA}**

One major requirement in the following reaction set-up was the *in situ* formation of aldehydes by *AlkJ* and sufficient amounts of *FucA*. Since unbalanced coproduction of *AlkJ* and *FucA* from the pKA1_ *fucA::alkJ* assembled in this thesis hardly expressed the DHAP-dependent aldolase, the pKA1_ *fucA* solely expressing the aldolase was

employed. Consequently, *AlkJ* could not be stably expressed in the same cell from two plasmids containing the same ORI and a mixed cell approach was pursued.^[6]

Both the CRAS and the FucY strain were cultivated as outlined in G VIII.5 and protein coproduction checked after expression (**Figure E-25A**) and after biotransformations (24 h) in mixed cell cultures (**Figure E-25B**).

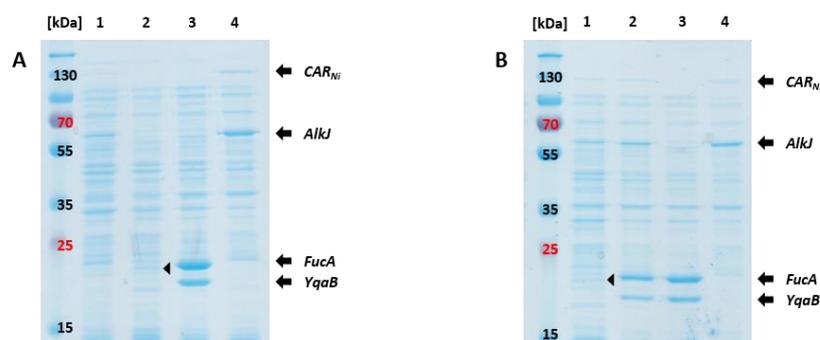


Figure E-25. Expression control of the mixed culture approach. SDS-PAGE analysis of whole cells (**A**) of single cell cultures and (**B**) after biotransformations with mixed cells (24 h). Weak expression of *Fsa1-A129S* indicated by (◄). Sample loading normalized to $OD_{590} = 3.5$. Loading scheme: (**A**) untransformed *E. coli* BL21(DE3) (**1**), *Fsa1-A129S* strain (**2**), FucY strain (**3**), CRAS strain (**4**); (**B**) CRAS + *Fsa1-A129S* strain mix (**1**), CRAS + FucY mix (**2**), FucY strain control (**3**), CRAS strain control (**4**).

For biotransformations, the final OD_{590} values of mixed LEP cultures was ≈ 20.0 . Otherwise, the setting followed the standard screening procedure (see G III.2). Biotransformations were performed with 5 mM of the model substrates **2b** and **2d** to produce the aldehyde **2c** in the CRAS strain. After diffusion of aldehyde intermediates into the FucY strain, *FucA* should catalyze the formation of **P-2e_{FucA}** with intracellular DHAP and *YqaB* finally form the dephosphorylated target aldol **2e_{FucA}** (**Figure E-26A**). As biotransformation controls, *E. coli* BL21(DE3) expressing *Fsa1-A129S* (pET16b_ *fsa1-A129S*) was prepared accordingly and biotransformations with a mix of CRAS cells were performed in the presence of DHA (20 eq; **Figure E-26B**). As negative control, LEPs of the FucY strain were incubated with the model substrates and should not convert primary alcohols and carboxylic acids. Sampling for GC/FID and HPLC was done as before (G III.2).

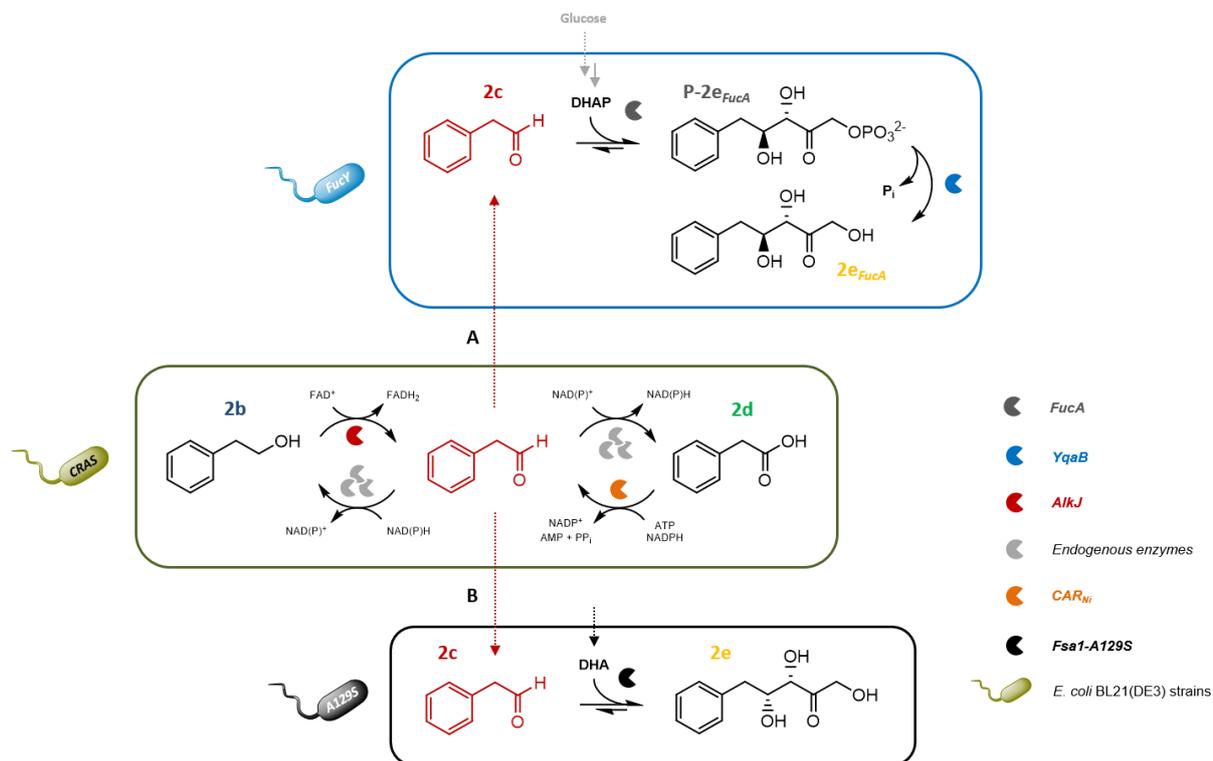


Figure E-26. Mixed culture cascade scheme. *In situ* preparation of **2c** from **2b** or **2d** in the CRAS strain and equilibration of all cascade intermediates. Diffusion of **2c** out of the cell. **(A)** Cascade reaction in the FucY strain to produce **2e_{FucA}** from **2c** and intracellular DHAP. **(B)** Cascade reaction in *E. coli* BL21(DE3) expressing *Fsa1-A129S* to form **2e** in the presence of DHA.

The analysis of biotransformation samples by calibrated GC/FID confirmed the activity of *AlkJ* and *CAR_{Ni}* in the CRAS strain. As expected, both enzymes formed the distinct redox equilibrium between **2b**, **2c**, and **2d**. The FucY strain does neither exhibit oxidation activity toward **2b** nor reduction activity toward **2d** and was employed as negative control. Indeed, the two substrates were not converted and could be well recovered from the biotransformation reactions (**Figure E-27**).

Regarding mixed cell biotransformations containing the CRAS and the FucY strain, the aldehyde intermediate **2c** was again produced from **2b** and **2d**. Whereas **2b** was quickly oxidized (**Figure E-27A**), **2d** was only slowly reduced indicating low activity of *CAR_{Ni}* (**Figure E-27B**). In both cases, the redox equilibrium was formed after 24 h reaction time.

In mixtures containing the CRAS and the *Fsa1-A129S* expressing strain, **2b** was poorly and **2d** hardly converted (**Figure E-27A** and **Figure E-27B**, respectively). Low recoveries according to GC analysis are usually a good indication for the production of aldols. However, control transformations containing 20 eq DHA only yielded low amounts of **2e** (from substrate **2b**: 5.6% after 1 h reaction time; **2d**: <5%). This can be attributed to insufficient amounts of *Fsa1-A129S* in the used batch (**Figure E-25**), which usually converts reactive **2c** into the nontoxic aldol **2e**. Furthermore, retro-aldol reaction and endogenous enzyme activities form the substrates again and outpace the activity in the CRAS strain (**Figure E-27A**).

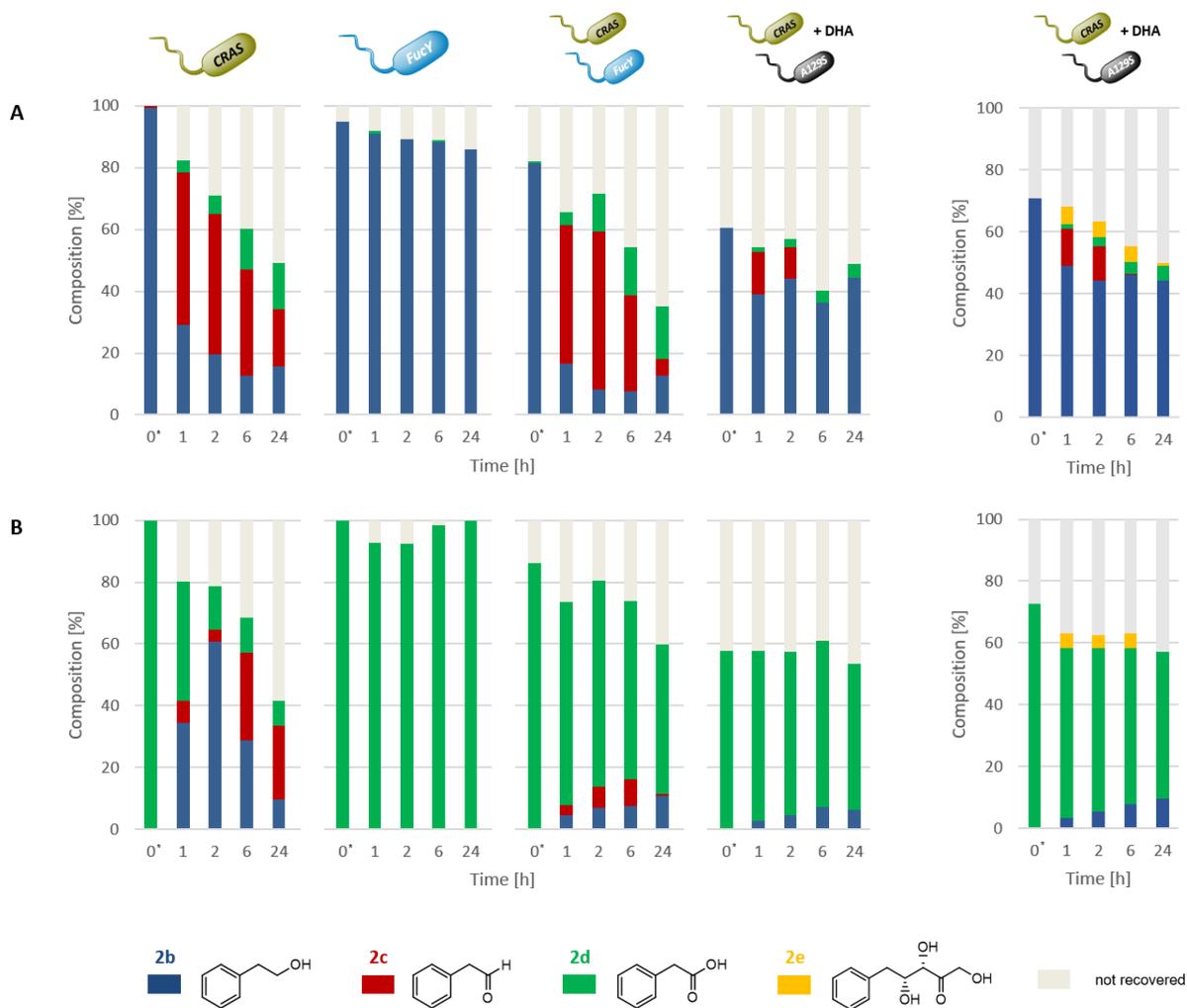


Figure E-27. Aldol production in the mixed culture approach. Cascade starting from (A) the primary alcohol **2b** or (B) the carboxylic acid **2d**. Control experiments employing the CRAS or FucY strain; mixed cultures to produce **2e_{FucA}** (CRAS + FucY strain) from intracellular DHAP or **2e** (CRAS + *Fsa1-A129S* strain + 20 eq DHA). On the left: GC results of all biotransformations. On the right: Confirmation of **2e** production. Results from single biotransformations analyzed by calibrated GC/FID (**2a–d**) and HPLC (**2a–e**). Stacked bars add up to 100% and represent the sum of **2b** (blue), **2c** (red), **2d** (green), and not recovered material (grey). Reduced recovery of material at t_0^* due to insufficient mixing and a loss in mass balance due to the reactivity/volatility of **2c**.

Regarding the DHAP-dependent aldolase cascade, if the desired **2e_{FucA}** was produced in mixed cultures employing the CRAS and the FucY strain, it was below quantifiable amounts.

E 1.6.4 Interim summary II

Unfortunately, the backup strategy separating aldehyde and aldol production in the CRAS and the FucY strain, respectively, failed. Since this two-cell system depends on the passage of the reactive intermediate through two cellular membranes, from the CRAS into the FucY strain, equilibration below toxicity levels is not achieved (**Figure E-27**), which certainly triggers previously discussed stress responses and impair both cell viability and the flux through the two synthetic cascades investigated (**Figure E-26**).

Although the coproduction of *FucA* and *YqaB* could be improved by the adaption of expression conditions (G VIII.5) and the reduced overall metabolic load in the FucY strain, the target aldol **2e_{FucA}** could not be detected. One explanation, other than the unfavorable longevity of the **2c** intermediate in the mixed cell culture, could root in a bottleneck discussed previously. According to SDS-PAGE analysis, *YqaB* is readily expressed to pull the

equilibrium of the aldol reaction from the phosphorylated **P-2e_{FucA}** to the dephosphorylated target **2e_{FucA}** (Figure E-25). However, metabolomic analysis by S. Milker and recalculation of intracellular DHAP concentrations (<0.2 mM) suggest that the mere application of late exponential phase FucY cells (i.e., LEPs) simply does not suffice to efficiently produce **P-2e_{FucA}** and, finally, **2e_{FucA}**.

Consequently, if considering intracellular ATP levels, the immediate solution to this bottleneck remains the implementation of *DhaK* suggested in this thesis (see E 1.6.1 and E 1.6.2).

Table E-3. Unsuccessful production of the aldol **2e_{FucA}** in mixed cell cultures

Engineered strains ^[a]	Expression conditions	Expression levels ^[b]				Screening conditions ^[c]	Aldol production	Associated literature
		<i>AlkJ</i>	<i>CAR_{Ni}</i>	Aldolase	<i>YqaB</i>			
	G VIII.5	++	+	+++ ^[d]	++	LEPs (5 mM 2b or 2c)	2e_{FucA} n.d.	This work
	G VIII.5	++	+	± ^[e]	n.a.	LEPs (5 mM 2b or 2c)	2e 5.6% ^[f]	This work

^[a] Mixed culture of CRAS (pKA1_alkJ and pETDuet-1_PPtase_{Ec}::carNi) + FucY (pKA1_fucA and pCDFDuet-1_yqaB) or Fsa1-A129S expressing *E. coli* BL21(DE3) (pET16b_fsa1-A129S). ^[b] Expression levels according to SDS-PAGE analysis; ^[c] LEPs (OD₅₉₀ ≈ 20.0), 1% (w/v) glucose, 5 mM substrate, 5% (v/v) ACN, 25°C, 250 rpm; after 6 h reaction time, 0.5% (w/v) glucose were spiked. ^[d] DHAP-dependent aldolase *FucA*. ^[e] DHA-dependent aldolase *Fsa1-A129S*. ^[f] Highest detectable amount according to calibrated HPLC; n.a. = not applicable, n.d. = not detected.

F Conclusion and outlook

Past and current research has successfully assembled *de novo* pathways to accumulate natural compounds and to produce valuable nonnatural chemicals with applications in all areas of today's modern life.^[6, 43-44] The repertoire of organic reactions mediated by biocatalysts is increasing fast, driven by improved methods for enzyme discovery, engineering, and characterization.^[17, 68, 277, 541] The growing understanding of efficiently assembling (synthetic) genetic elements and enzyme-coding genes into artificial metabolic pathways not only expands the biocatalytic toolbox; it accelerates developments in the field of synthetic biology to evolve *de novo* cascades, transforming host cells from a mere chassis for heterologous pathway production into a highly optimized microbial cell factory.^[6]



In summary, this thesis successfully designed enzymatic cascades for the production of polyhydroxylated compounds (i.e., aldols) by biocatalytic retrosynthesis (D I), identified and characterized enzymes for subsequent pathway assembly from five major families (D II):

- Esterases (*BS2*, *Pfel*, *PEST*)
- ADHs (*ADH_{Lk}*, *ADH_{Rr}*, *ADH-A*, *ADH-ht*, *AlkJ*)
- CARs (*CAR_{Ni}*, *CAR_{Mm}*)
- (D)HA-dependent aldolases (*Fsa1*, *Fsa1-A129S*)
- DHAP-dependent aldolases (*FruA*, *FucA*, *RhuA*)
- Phosphatases (*PhoN-Se V78L* variants, *PhoN-Sf*, *YqaB*)



Pathway construction (D III) was achieved by state of the art SLIC techniques and conventional restriction enzyme cloning (D III.1 and D III.2). The latter was revisited to refine standard procedures and, for example, isolate linearized vectors for subcloning in high yields and purity independent from expensive commercial kits for DNA purification (G II.1).



Assembled *de novo* pathways were transferred and implemented in the well-characterized host *E. coli*. Pathway validation was performed by complementing analytical methods (e.g., GC, HPLC, LC-MS/MS) to identify potential bottlenecks that impair the flux through the artificial enzymatic cascade (D III.1 and D III.2).



Although *de novo* pathways benefit from the cooperative effect of multiple biocatalysts with their inherently high chemo-, regio-, and stereoselectivity in combination with the omission of intermediate isolation,^[8, 348] they strictly require a balanced carbon flux through the synthetic pathway to meet performance metrics such as high product titers.

The *in vivo* approaches pursued in this thesis offer advantages including the recycling of cofactors, which is essential for demanding (redox) transformations as the FAD-dependent oxidation of primary alcohols to the corresponding aldehydes by *AlkJ* (D II.2.2) or the ATP- and NADPH-dependent reduction of carboxylates by *CAR_{Ni}* (D II.3).^[6, 348] In contrast, the control over the carbon flux *in vivo* is complicated in many aspects.

The mini-pathway established in this work consisted of the two metabolically nonrelated enzymes, the membrane-associated ADH *AlkJ* from *P. putida* and the cytosolic aldolase *Fsa1-A129S* from *E. coli* (D III.1).^[348] *AlkJ* was previously reported to oxidize primary aliphatic alcohols to the corresponding aldehydes.^[289] In this thesis, the ADH was identified as an efficient biocatalyst for the oxidation of primary, structurally different

aromatic alcohols (B I). However, the *in situ* preparation of highly reactive aldehyde species triggered cellular responses in *E. coli*, leading to the irreversible overoxidation to the corresponding carboxylic acids and, ultimately, driving the carbon flux into a dead end. This bottleneck was addressed by (i) coupling the first cascade step with the subsequent aldolase transformation catalyzed by *Fsa1-A129S*, (ii) the utilization of the highly engineered RARE strain to prevent retro-aldol formation by the KO of competing enzymatic activities, and (iii) the introduction of a CAR from *N. iowensis* (and an accessory PPTase from *E. coli*, which posttranslationally modifies apo-CAR_{Ni} to yield holo-CAR_{Ni}).

For (i), vector-based pathway modules were constructed by SLIC for the coexpression of *AlkJ* and *Fsa1-A129S*. This reduced the plasmid burden and increased the viability of host cells harboring the newly assembled plasmids.^[348] To balance enzyme production, the target genes were arranged in different architectures featuring operons, pseudo-operons, and monocistronic configurations and gave rise to the vectors pOPE, pPOP, pMON1, and pMON4. The latter two monocistronic arrangements were achieved by the insertion of the synthetic terminator sequences *B0011* and *B0014*, respectively, in the short intergenic region between the *alkJ* and the *fsa1-A129S* gene. Although transcriptional terminators are known to play a key role in regulating genetic systems, the systematic study of different termination signals is still in its infancy.^[505, 542] Therefore, the influence of the two terminator sequences on protein production, overall pathway performance, and the standardization as regulatory element for cascade design was investigated. An unintended context effect resulting from interactions between the synthetic terminators and adjacent enzyme-coding ORFs resulted in unsteady *Fsa1-A129S* expression levels. Importantly, *AlkJ* activity was strongly impaired *in vivo*, surprisingly, at unchanged expression levels according to SDS-PAGE analysis. The extension of the spacer sequences flanking *B0014* restored *AlkJ* activity in the two additional plasmids with monocistronic configurations, pMON5 and pMON6 (E I.1).

In the context of aldol production, the newly assembled pPOP proved to be beneficial regarding both soluble aldolase expression levels and cell viability of transformants. A high excess of (D)HA (10–20 eq) was necessary to shift the equilibrium of the aldol reaction toward target polyhydroxylated compounds, prevent retro-aldol reaction, and the subsequent formation of carboxylate byproducts *in vivo*. Initial product titers of the model compound (3*S*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e**) achieved up to 95% according to calibrated HPLC in 2 h reaction time (D III.1).

For (ii), the mini-pathway was transferred into the *E. coli* K-12 RARE strain exhibiting reduced aromatic aldehyde reducing activity due to the KO of endogenous ADHs and AKRs.^[42] In the absence of these competing enzymatic activities, retro-aldol reaction was less pronounced and the DHA loading decreased 4-fold (5 eq) by achieving comparably high yields of **2e** in only 2 h.

For (iii), CAR_{Ni} accessed the carboxylate sink but the produced aldehydes were rapidly converted to the corresponding primary aromatic alcohols by the metabolic background of *E. coli*. The activity of the two enzymes with opposing functional group activity, *AlkJ* and CAR_{Ni}, formed a redox equilibrium between alcohols, aldehydes, and carboxylic acids. Beneficially, intracellular aldehyde concentrations were stably equilibrated below cytotoxic levels, yet freely available for subsequent aldolase reaction. The formation of this ‘hidden reservoir’, enzymatically containing the reactive aldehyde species, addressed the biocatalytic challenge of aldehyde toxicity and ephemerality of intermediates in living cells independent from extensive strain engineering by target gene KOs, for example.^[42, 348] Since this unperceived strategy to reroute the pathway flux from undesired byproducts toward target cascade intermediates offers potential to reverse other byproduct forming reactions, given that an reversing enzyme activity has already been added to the biocatalytic toolbox, it certainly complements existing flux enhancement methods.^[348]

Consequently, the ‘biocatalytic reservoir’ was tapped by *Fsa1-A129S*-mediated cascade reaction to yield up to 95% of **2e** in 6 h. The longer reaction time was attributed the unbalanced coproduction of *AlkJ*, *Fsa1-A129S*,

CAR_{Ni}, and *PPTase_{Ec}* in the same cell. Insufficient amounts of the aldolase could be compensated by the extracellular addition of the purified catalyst. This hybrid system yielded equally high amounts of **2e** in only 2 h, importantly, at decreased amounts of the carboxylate byproduct (E 1.4).^[348]

The second pathway constructed in this thesis consisted of *AlkJ* and the model DHAP-dependent aldolase *FucA* from *E. coli* (D III.2). Although four stereocomplementary DHAP-dependent aldolases – *FruA* (3*S*,4*R*), *FucA* (3*R*,4*R*), *RhuA* (3*R*,4*S*), and *TagA* (3*S*,4*S*) – are available from the biocatalytic toolbox that accept a broad range of aldehyde acceptors, their use at larger production scales *in vitro* is strongly limited due to the labile and expensive DHAP donor.^[293, 415] To tackle this issue, the constructed pathway aimed at hijacking DHAP from the intracellular metabolite pool of *E. coli* to feed the *de novo* cascade with the aldol donor solely from glucose.

Therefore, pathway design was extended to modules additionally expressing phosphatases (*PhoN-Sf* from *S. felchneri* and *YqaB* from *E. coli*) to irreversibly dephosphorylate aldol adducts to shift the reaction equilibrium toward the nonphosphorylated target compounds. Furthermore, metabolomics showed that the *DhaK* from *C. freundii* can phosphorylate the extracellularly added donor DHA to DHAP *in vivo*. The combined activities of *DhaK* and *YqaB* were thought to act as a ‘push and pull’ module for pathway optimization (E 1.6.2). Although (3*R*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e_{FucA}**) could not be produced under both various pathway expression and different screening conditions, this work identified crucial bottlenecks to be considered for the future application and optimization of *de novo* pathways employing DHAP-dependent aldolases (D III.2 and E 1.6). Immediate optimization strategies will involve both the construction of pathway modules for the balanced coproduction of *FucA*, *YqaB*, and *DhaK* and further refinement of fermentation conditions. As to cloning, a target insert like *FucA* or *DhaK* can be easily inserted into the MCS-2 of pCDFDuet-1_*yqaB* yielding pCDFDuet-1_*yqaB::insert*.

However, the efficient production of polyhydroxylated compounds (or other value-added chemicals) *via* optimized synthetic cascades in engineered host cells is rendered insignificant if the target molecules cannot be isolated. This essential point has long been neglected for polyhydroxylated molecules and *in vitro* systems suffered from low yields as aldols and sugar derivatives are highly water-soluble and acyclic adducts temperature sensitive. In course of this work, a refined SPE downstream purification protocol was successfully applied to translate the high product titers of the optimized mini-pathway into isolated yields of up to 91% (E 1.5).^[435]

In conclusion, the cooperative effects of multiple flux optimization strategies from genetics, metabolic engineering, and synthetic biology maximized target aldol product titers and isolated yields (chapter E). Optimization strategies were successfully applied on different molecular and process levels and included:

- The construction of vector-based modules coexpressing multiple pathway enzymes to reduce the plasmid burden and to enable the combination of different cascade modules *via* compatible ORIs.
- Enzyme balancing by different genetic pathway architectures featuring additional regulatory elements (e.g., promoters and terminators), the optimization of the synthetic genetic contexts, and the employment of an *in vitro/in vivo* hybrid system.
- Reversion of competing byproduct formation reactions *in vivo* simply by the introduction of a single, nonnative enzymatic activity, simultaneously tackling the unmet challenge of aldehyde toxicity in living cells.
- The utilization of KO strains and to streamline carbon flux through *de novo* pathways.
- The optimization of reaction parameters including recombinant expression of multiple pathway elements, adaption of biotransformation conditions, and, importantly, the efficient purification of aldols by SPE.

The synthesized polyhydroxylated compounds are precursors for sugar derivatives (e.g., D-fagomine) and have potential applications in the pharmaceutical industry and medical research. Furthermore, the modular pathway

design employed in this thesis provides the possibility to enter the *de novo* cascade from different substrates like carboxylic esters (D II.1), couple the *AlkJ*-mediated *in situ* preparation to another cascade-type reaction such as transamination by ω -TAs (G VI.9.1) or chiral carbonylations by *PDC_{Ap}* (G VI.9.2), or implement cofactor recycling modules such as *GDH_{2xBs}* (E I.2) to ease the usage of NAD(P)H from the cellular redox equivalent pool.

This work highlights the fact that with realization of more complex biocatalytic cascade reactions *in vivo*, single optimization strategies may not suffice to move pathway design toward industrial applications. Consequently, the application of cross-disciplinary strategies from biological engineering, biochemistry, genetics, metabolic engineering, synthetic biology, and systems biology will be vital to push biocatalytic processes forward to take today's challenges and produce fine and bulk chemicals of the future.

G Experimental part

G I Materials and methods: Standard molecular biology techniques

Unless noted otherwise, all reagents were purchased from commercial suppliers and used without further purification. All (plastic) consumables and standard glass equipment were either sterile upon purchase or sterilized prior to use by autoclaving (121°C, 15 min, elevated pressure; Tuttnauer 2540EL autoclave). All reagent and media solutions were sterilized prior to use by autoclaving (121°C, 20 min, elevated pressure; Tuttnauer 2540EL autoclave).

G I.1 General stock solutions

Aqueous stock solutions were sterilized by filtration (0.2 µm cellulose acetate syringe filter) and stored at -20°C. Unless noted otherwise, working concentrations were used as in **Table G-1**.

Table G-1. Commonly used stock solutions

Reagent	Concentration in dH ₂ O	Standard working concentration
Amp	50 mg·mL ⁻¹	100 µg·mL ⁻¹
Cam	34 mg·mL ⁻¹ (in abs. EtOH)	34 µg·mL ⁻¹
Kan	50 mg·mL ⁻¹	50 µg·mL ⁻¹
Str	20 mg·mL ⁻¹	20 µg·mL ⁻¹
Tet	10 mg·mL ⁻¹	10 µg·mL ⁻¹
IPTG	0.1 M	varying
PMSF	0.1 mM (in abs. <i>i</i> PrOH)	0.1 µM

G I.2 Standard media preparations

Unless noted otherwise, amounts of reagents refer to the preparation of 1 L medium. All media were stored in the dark at RT after sterilization and, once opened, at 4°C. Visual control was done prior to use. Most media were prepared according to Studier (2005).^[543]

Standard media included variations of lysogeny broth (LB) medium and terrific broth (TB) medium, which were used for bacterial growth studies and expression protocols (**Table G-2**). LB-Luria medium was employed for enzyme production and subsequent isolation of periplasmic proteins (**Table G-2**). The minimal medium M9-N* was employed for the same purposes (**Table G-4**). AIM was conveniently used for IPTG-independent induction of enzyme production (**Table G-3**).

Table G-2. Constituents of bacterial complex media used in this thesis

LB-Miller medium		LB-Luria medium		TB medium		LB-ECAM ^[293]	
10.0 g	bacto-peptone	10.0 g	bacto-peptone	12.0 g	bacto-tryptone	5.0 g	bacto-peptone
5.0 g	yeast extract	5.0 g	yeast extract	24.0 g	yeast extract	2.5.0 g	yeast extract
10.0 g	NaCl	0.5 g	NaCl	16.4 g	K ₂ HPO ₄ · 3 H ₂ O	5.0 g	NaCl
				2.3 g	KH ₂ PO ₄	2.14 g	KCl
						10.8 g	NaH ₂ PO ₄ · H ₂ O
						17.3 g	Na ₂ HPO ₄ · 7 H ₂ O
						1.0 mL	10 M (NH ₄) ₂ SO ₄
						1.0 mL	2.5 M citric acid
						2.5 mL	1 M MgSO ₄
						1.0 mL	1 M CaCl ₂
						0.6 mL	0.1 M FeCl ₃ · 6 H ₂ O
						2.0 mL	1 mM thiamine-HCl
						10.0 mL	100X trace elements solution
						10.0 mL	40% (w/v) glucose

For LB-ECAM, bacto-peptone, yeast extract, and salts were dissolved, filled up to 972 mL with dH₂O, and autoclaved. The 10 M (NH₄)₂SO₄, 2.5 M citric acid, 1 M MgSO₄, 1 M CaCl₂, and 40% (w/v) glucose were prepared and autoclaved separately and added sequentially under sterile conditions. The 0.1 M FeCl₃ · 6 H₂O, 1 mM thiamine-HCl, and trace element solution (see Fehler! Kein gültiges Resultat für Tabelle.) were sterilized by filtration (sterile syringe filter, 0.2 µm cellulose acetate) and added under sterile conditions.

Table G-3. Constituents of bacterial autoinduction medium (AIM)

LB-0.8G		LB-5052		20X NPS		50X 5052	
10.0 g	bacto-peptone	10.0 g	bacto-peptone	66.0 g	(NH ₄) ₂ SO ₄	250 g	glycerol
5.0 g	yeast extract	5.0 g	yeast extract	136 g	KH ₂ PO ₄	25.0 g	glucose
10.0 g	NaCl	10.0 g	NaCl	142 g	Na ₂ HPO ₄	100 g	α-lactose
1.0 mL	1 M MgSO ₄	1.0 mL	1 M MgSO ₄				
20.0 mL	40% (w/v) glucose	20.0 mL	50X 5052				
50.0 mL	20X NPS	50.0 mL	20X NPS				

For LB-0.8G and LB-5052, bacto-peptone, yeast extract and NaCl were dissolved, filled up to 929 mL with dH₂O, and autoclaved. The 1 M MgSO₄, 20X NPS, 50X 5052, and 40% (w/v) glucose were prepared and autoclaved separately and added sequentially under sterile conditions. LB-0.8G was used for preculture preparation and LB-5052 for bacterial growth and enzyme production.

Table G-4. Constituents of bacterial minimal medium used in this thesis

M9-N* medium		10X M9 salts		100X Trace element solution	
100 mL	10X M9 salts	5.0 g	NaCl	0.18 g	ZnSO ₄ · 7 H ₂ O
3.0 mL	1 M MgSO ₄	75.0 g	Na ₂ HPO ₄ · 2 H ₂ O	0.12 g	CuCl ₂ · 2 H ₂ O
1.0 mL	1 M CaCl ₂	30.0 g	KH ₂ PO ₄	0.12 g	MnSO ₄ · 2 H ₂ O
0.6 mL	0.1 M FeCl ₃ · 6 H ₂ O	24.0 g	(NH ₄) ₂ SO ₄	0.18 g	CoCl ₂ · 6 H ₂ O
2.0 mL	1 mM thiamine-HCl			0.03 g	H ₃ BO ₃
10.0 mL	100X trace element solution			0.025 g	Na ₂ MoO ₄ · 2 H ₂ O
100 mL	20% (w/v) glucose			0.084 g	Na ₂ EDTA · 2 H ₂ O

All stock solutions were prepared and sterilized separately. Trace element solution, 1 mM thiamine-HCl, 0.1 M FeCl₃ and 20% (w/v) glucose were sterilized by filtration (sterile syringe filter, 0.2 μm cellulose acetate). For preparation of M9-N* medium, all solutions were mixed under sterile conditions and filled up to the final volume with sterile dH₂O.

Table G-5. Constituents of resting cell medium (RCM)

RCM		10X M9 salts (N-free)	
100 mL	10X M9 salts (N-free)	5.0 g	NaCl
3.0 mL	1 M MgSO ₄	75.0 g	Na ₂ HPO ₄ · 2 H ₂ O
1.0 mL	1 M CaCl ₂	30.0 g	KH ₂ PO ₄
50.0 mL	20% (w/v) glucose		

All stock solutions were prepared and sterilized separately. For resting cell medium preparation, components were mixed under sterile conditions and filled up to the final volume with sterile dH₂O.

G I.3 Strain cultivation on agar plates

E. coli strains were grown in 4 mL LB-Miller medium supplemented with antibiotic(s) if applicable at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. From this culture, cells were streaked on LB agar plates (LB-Miller medium containing 1.5% (w/v) bacteriological agar in standard petri dishes [94x16 mm] supplemented with antibiotic if applicable). Plates were incubated upside down at 37°C for 12–24 h (Heraeus Function line incubator). Agar plates were stored in the dark at 4°C and *E. coli* strains propagated every 4–6 weeks onto freshly prepared agar plates.

G I.4 Permanent culture preparation

E. coli strains were incubated in 4 mL LB-Miller medium (supplemented with antibiotic if applicable) at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. Cell cultures were dispensed into 1.5 mL Eppendorf tubes and an equal volume of 50% (v/v) glycerol was added. The resulting permanent cultures were gently mixed, snap-frozen in liquid nitrogen, and stored at -80°C until further use.

G I.5 Preparation of *E. coli* cell-free extracts (CFEs)

The optimized expression conditions of individual enzymes for CFE preparation are summarized in G IV.

For standard applications, *E. coli* strains were incubated in 4 mL LB-Miller medium supplemented with antibiotic(s) if necessary at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. The TB

medium main culture, supplemented with antibiotic(s) if applicable, was inoculated with 1% (v/v) of the overnight culture and grown at 37°C, 200 rpm until an $OD_{590} = 0.5$ was reached (WPA colourwave, CO7500 Colorimeter). Protein production was performed in the presence of 0.5 mM IPTG at 20–30°C, 150 rpm for 20–24 h.

Cells were harvested by centrifugation (6 000 x g, 4°C, 15 min; Sigma Laboratory Centrifuge 6k15 or 3k30). The cell pellet was resuspended in 1/10 volume of the main culture in 50 mM Tris-HCl (pH 7.5; 6.057 g·L⁻¹ Trizma® base; T6066, Sigma) and centrifuged. The washed pellet was resuspended in 1/40 volume of the main culture in 50 mM Tris-HCl (pH 7.5). To the resulting cell suspension, 0.1 mM PMSF (17.42 mg·mL⁻¹ phenylmethanesulfonyl fluoride in *i*PrOH; 78830, Sigma) was added to the standard working concentration of 0.1 μM prior to cell lysis, which was done on ice by sonication (5 s·min⁻¹ pulse for 9 min, 40% amplitude; KE76 probe, Bandelin Sonoplus HD3500). Cell debris were removed by centrifugation (14 000 x g, 4°C, 25 min; Sigma Laboratory Centrifuge 6k15 or 3k30). Insolubles were resuspended in a sufficient volume of 50 mM Tris-HCl (pH 7.5). The total amount of protein was determined by Bradford assay for both the insoluble fraction and the resulting CFE prior to analysis by SDS-PAGE. CFEs were stored at 4°C for immediate use or kept at -20°C for long time storage if applicable.

G I.6 Preparation of *E. coli* resting cells (RCs)

The optimized expression conditions of individual enzymes for RC preparation are summarized in G IV. The optimized expression conditions of whole synthetic pathways are described in the corresponding subsection of the results and discussion chapter.

Routinely, *E. coli* strains were incubated in 4 mL LB-Miller medium supplemented with antibiotic(s) if applicable at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. The main culture, supplemented with antibiotic(s) if applicable, was inoculated with 1% (v/v) of the preculture culture and grown at 37°C, 200 rpm until an $OD_{590} = 0.5$ was reached (WPA colourwave, CO7500 Colorimeter). Protein production was induced in the presence of IPTG at 0.5 mM final concentration and performed at 20–30°C, 150 rpm for 20–24 h.

Cells were harvested by centrifugation (6 000 x g, 4°C, 15 min; Sigma Laboratory Centrifuge 6k15 or 3k30). Pellets were resuspended in 1/10 volume of the main culture in RCM (1% (w/v) glucose, 8.6 mM NaCl, 42.1 mM Na₂HPO₄, 22 mM KH₂PO₄, 3 mM MgSO₄, 0.1 mM CaCl₂) and centrifuged. The cell pellet was washed and resuspended in a sufficient volume of RCM until an $OD_{590} = 20.0$ was reached. RCs were stored at 4°C up to two days, changing the RCM once a day.

G I.7 Total amount of protein determination by Bradford assay

For routine determination of the total amount of protein, a given protein containing solution was diluted 1:30 with deionized water (dH₂O). In 96-well plates (PS microplate, flat bottom, Greiner Bio-one), 200 μL of 1:5-diluted Bradford reagent (Protein Assay Dye Reagent Concentrate; 500-0006, Bio-Rad) were mixed with 5 μL of the pre-diluted protein solution for 5 s (1 350 rpm; Heidolph Titramax 1000), and incubated at RT for 15 min. The absorbance was determined at 595 nm with a plate reader (Anthos Zenyth 3100) and the amount of protein was determined by bovine serum albumin (BSA) calibration (0–1 mg·mL⁻¹ albumin from bovine serum; A4503, Sigma). An example BSA calibration is shown in **Figure G-1**.

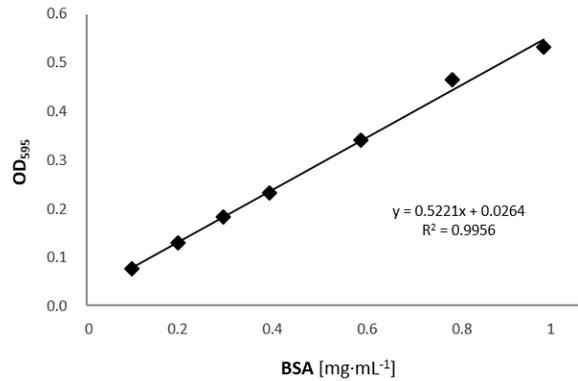


Figure G-1. Exemplary Bradford assay calibration. Calibration was performed with BSA (0–1 mg·mL⁻¹) in 50 mM Tris-HCl (pH 7.5). Calibration was done for fresh Bradford reagents and recalibrations performed quarterly.

G 1.8 Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Routinely, 12% (w/v) polyacrylamide gels with a thickness of 0.75 mm were used and SDS-PAGE was performed under denaturing and reducing conditions in 1X SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 35 mM SDS) at 15–25 mA per gel.

Resolving and stacking gels were prepared as in **Table G-8**. The resolving gel was covered with *i*PrOH, which was removed again prior to pouring the stacking gel on top. Protein samples were denatured with ready-to-use SDS-PAGE sample buffer (65 mM Tris, 4% (w/v) glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue, 5% (v/v) β-mercapto ethanol; **Table G-7**) at 95°C for 4 min before loaded onto the gel. Whole cell samples were denatured at 95°C for 10 min. For comparison between different gels, sample loading was either normalized to 10 μg protein per lane or samples from expressing cultures were collected with their volumes normalized to OD₅₉₀ = 7.0. The latter proved most convenient for monitoring protein expression over time. For protein size determination, 5 μL of prestained protein marker were loaded onto each gel (PageRuler™ Prestained Protein Ladder; 26616, Thermo Scientific). Gels were stained following the microwave staining procedure in G 1.8.2.

G 1.8.1 Buffer and reagent preparations

All buffer and reagent solutions were prepared following the Mini-Protean® 3 Cell instruction manual (Bio-Rad).

Table G-6. Reagent solutions for SDS-PAGE

30% (w/v) Acrylamide (30% T, 2.67% C)	10% (w/v) SDS	0.5% (w/v) Bromophenol blue	10% (w/v) APS
29.2 g acrylamide	1.00 g SDS	50.0 mg bromophenol blue	1.00 g APS
0.80 g N',N'-bis-methylene acrylamide			

The amounts for 30% (w/v) acrylamide account for the preparation of 100 mL final solution. All other reagent amounts account for the preparation of 10 mL final solution. 10% (w/v) SDS and 0.5% (w/v) bromophenol blue were stored at RT; the acrylamide solution in the dark at 4°C. 10% (w/v) APS was dispensed in 1 mL aliquots and stored at -20°C.

Table G-7. Buffer solutions for SDS-PAGE

Resolving gel buffer (1.5 M Tris-HCl, pH 8.8)		Stacking gel buffer (0.5 M Tris-HCl, pH 6.8)		10X SDS-PAGE running buffer		SDS-PAGE sample buffer	
18.15 g	Trizma® base	6.0 g	Trizma® base	30.3 g	Trizma® base	7.1 mL	dH ₂ O
				144 g	glycine	2.5 mL	stacking gel buffer
				10 g	SDS	5.0 mL	glycerol
						4.0 mL	10% (w/v) SDS
						0.4 mL	0.5% (w/v) bromophenol blue

For resolving gel and the stacking gel buffer, the amounts of Trizma® base account for the preparation of 100 mL final solution. Adjustment of pH was done with 5 M HCl. Both buffer solutions were sterilized prior to use. For 10X SDS-PAGE running buffer, reagents were dissolved in dH₂O and filled up to a final volume of 1 000 mL. For ready-to-use SDS-PAGE sample buffer preparation, 50 µL β-mercapto ethanol had to be added to 950 µL SDS-PAGE sample buffer. The resulting sample buffer was stored at 4°C.

Table G-8. Polyacrylamide gel (12%) preparation

Resolving gel		Stacking gel	
2.0 mL	resolving gel buffer	1.0 mL	stacking gel buffer
3.2 mL	30% (w/v) acrylamide	0.52 mL	30% (w/v) acrylamide
2.7 mL	dH ₂ O	2.44 mL	dH ₂ O
80 µL	10% (w/v) SDS	40 µL	10% (w/v) SDS
40 µL	10% (w/v) APS	20 µL	10% (w/v) APS
4.0 µL	TEMED	4.0 µL	TEMED

The volumes account for the preparation of two gels. After TEMED and 10% (w/v) APS were added, gels were poured immediately.

G 1.8.2 Microwave staining

After performing electrophoresis, the stacking gel was carefully removed. The resolving gel was covered in a sufficient volume of dH₂O and incubated at 750 W for 1 min in the microwave. The gel was slightly shaken at RT for 2 min (PSU-10i, Grant-bio). The gel was submerged in fresh dH₂O, incubated at 500 W for 1 min, and shaken at RT for 2 min. The water was discarded. The gel was covered in dyeing solution (SimplyBlue™ SafeStain; LC6065, Novex®) and incubated at 350 W for 45 s. The gel was slightly shaken at RT for 5 min. The dyeing solution was discarded and the gel was washed in dH₂O with shaking for 10 min. Finally, the gel was preserved in 20% (w/v) NaCl, usually overnight, and scanned for documentation.

G 1.9 Transformation of competent *E. coli* cells

Unless noted otherwise, all steps were carried out on ice under sterile conditions. Both freshly prepared competent cells using CaCl₂ (G 1.9.1) and RbCl competent cells from bulk preparation (G 1.9.2) were used for routine transformations. Freshly prepared competent cells using RbCl (G 1.9.3) were used for more delicate transformations such as vector constructions by SLIC methods (G 11.2).

G 1.9.1 Chemical transformation using CaCl₂

A single colony of the *E. coli* strain to be transformed was incubated in 4 mL LB-Miller medium (supplemented with antibiotic if necessary) at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. A fresh LB-Miller culture (supplemented with antibiotic if necessary) was inoculated with 1% (v/v) of the overnight

culture and grown to an $OD_{590} = 0.2\text{--}0.4$. The cell culture was dispensed in 1.5 mL aliquots in Eppendorf tubes and centrifuged (5 000 rpm = 2 292 x g, 4°C, 10 min; Sigma Laboratory Centrifuge 3k15). The supernatant was discarded, cells were resuspended in 0.5 mL ice-cold 0.1 M $CaCl_2$ (11.1 g·L⁻¹; C4901, Sigma) and incubated for 15 min. Cells were centrifuged, resuspended in 0.1 mL ice-cold 0.1 M $CaCl_2$ and 1 µL of plasmid DNA (with a concentration of 50–100 ng·µL⁻¹) was added. Cells were incubated for 1 h. The heat shock was performed at 42°C for 1 min (Eppendorf Thermomixer Comfort). Cells were put on ice immediately for 2 min. For recovery, 0.5 mL prewarmed LB-Miller medium were added and cells were incubated at 37°C with shaking (650 rpm; Eppendorf Thermomixer Comfort) for 1 h. From the resulting solution, 100 µL were plated on prewarmed LB agar plates supplemented with the appropriate antibiotic. Plates were incubated upside down at 37°C for 12–24 h (Heraeus Function line incubator).

For each transformation experiment an empty vector control (positive control) was performed. Additionally, competent cells incubated without plasmid DNA were plated on both a LB agar plate supplemented with the appropriate antibiotic (negative control) and without antibiotic (growth control). Transformation efficiencies were calculated in transformants per µg plasmid DNA using the online tool from <http://www.sciencegateway.org/tools/transform.htm>.

G I.9.2 Bulk preparation of RbCl competent *E. coli* cells

A single colony of the desired *E. coli* strain was incubated in 4 mL LB-Miller medium at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. A fresh LB-Miller culture (100 mL) was inoculated with 1% (v/v) of the overnight culture (1 mL) and grown to an $OD_{590} = 0.35$. Cells were harvested by centrifugation (4 000 x g, 4°C, 10 min; Sigma Laboratory Centrifuge 6k15 or 3k30) and resuspended in 1/5 volume of the main culture in RF1 buffer (20 mL; 100 mM RbCl, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$, 15% (w/v) glycerol). Cells were incubated for 15 min, centrifuged, and resuspended in RF2 buffer (10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2$, 15% (w/v) glycerol) using 1/5 volume of the RF1 suspension (4 mL). Cells were dispensed in 100 µL aliquots in 1.5 mL Eppendorf tubes, snap frozen in liquid nitrogen and stored at -80°C for further use.

G I.9.3 Chemical transformation using RbCl

A single colony of the *E. coli* strain to be transformed was incubated in 4 mL LB-Miller medium (supplemented with antibiotic if necessary) at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) for 12–24 h. A fresh LB-Miller culture (supplemented with antibiotic if necessary) was inoculated with 1% (v/v) of the overnight culture and grown to an $OD_{590} = 0.3\text{--}0.4$. The cell culture was dispensed in cold 1.5 mL aliquots in Eppendorf tubes and centrifuged (5 000 rpm = 2 292 x g, 4°C, 10 min; Sigma Laboratory Centrifuge 3k15). The supernatant was discarded, cells resuspended in 0.3 mL ice-cold RF1 buffer (**Table G-9**; 100 mM RbCl, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$, 15% (w/v) glycerol) and incubated for 15 min. Cells were centrifuged, resuspended in 100 µL ice-cold RF2 buffer (**Table G-9**; 10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2$, 15% (w/v) glycerol), and 1 µL of plasmid DNA (with a concentration of 50–100 ng·µL⁻¹) was added. Cells were incubated for 1 h. The heat shock was performed at 42°C for 1 min (Eppendorf Thermomixer Comfort). Cells were put on ice immediately for 2 min. For recovery, 0.5 mL prewarmed LB-Miller medium were added and cells incubated at 37°C with shaking (650 rpm; Eppendorf Thermomixer Comfort) for 1 h. From the resulting solution, 100 µL were plated on prewarmed LB agar plates supplemented with the appropriate antibiotic. Plates were incubated upside down at 37°C for 12–24 h (Heraeus Function line incubator). Transformation controls were performed as in G I.9.1.

Table G-9. Buffer solutions for RbCl competent cell preparation

RF1 buffer (pH 5.8)		RF2 buffer (pH 6.8)	
1.21 g	RbCl	0.12 g	RbCl
0.99 g	MnCl ₂ · 2 H ₂ O	0.21 g	MOPS
0.294 g	KOAc	0.11 g	CaCl ₂ · 2 H ₂ O
0.148 g	CaCl ₂ · 2 H ₂ O	7.5 g	glycerol
15 g	glycerol		

Combine the reagents for the preparation of 100 mL of RF1 and RF2 buffer, respectively. The pH of RF1 was adjusted to 5.8 with 0.2 M acetic acid; RF2 was adjusted to 6.8 with 1 M NaOH. Both buffers were sterilized and stored at 4°C.

G I.10 Plasmid DNA isolation and quantification

Plasmid DNA isolation was performed according to the GeneJET Plasmid Miniprep Kit (K0503) by Thermo Scientific with the supplied solutions. All purification steps were carried out at RT.

Briefly, a single colony of the *E. coli* strain harboring the desired plasmid was incubated in 5 mL LB-Miller medium supplemented with the appropriate antibiotic at 37°C and 200 rpm (InforsHT Multitron 2 Standard) for 12–24 h. A cell pellet from 4 mL of the overnight culture was harvested in two consecutive centrifugation steps (8 000 x g, 10 min; Sigma Tabletop Centrifuge 1-14) in a 2.0 mL Eppendorf tube. The cell pellet was resuspended in 0.25 mL cold resuspension solution containing RNase A. To the suspension, 0.25 mL lysis solution were added and the suspension was thoroughly mixed by inverting. Neutralization solution (0.35 mL) was added and immediately mixed by inverting. Cell debris and chromosomal DNA were pelleted by centrifugation (>16 000 x g, 5 min; Sigma Tabletop Centrifuge 1-14). The supernatant was transferred to a supplied spin column by pipetting without disturbing the precipitate. It was centrifuged at >16 000 x g for 1 min. Two consecutive washing steps with 0.5 mL washing solution containing EtOH were performed. The flow-throughs were discarded and the column centrifuged for 1 min to remove residual washing solution. The spin column was transferred to a fresh 1.5 mL Eppendorf tube. DNase- and RNase-free water (50 µL) was added, incubated for 2 min and the plasmid DNA eluted by centrifugation at >16 000 x g for 2 min.

The purified DNA was quantified by NanoDrop® (NanoDrop 1000 Spectrophotometer, Thermo Scientific) and stored at -20°C until further use.

G I.11 Restriction enzyme control digestion of plasmid DNA

Restriction enzyme digestions were performed according to the supplier. In the following, the frequently used control digestion with *NcoI* (ER0571, Thermo Scientific) is outlined.

Table G-10. Example *NcoI* digestion mix

<i>NcoI</i> digestion mix		Final concentration
1.0 µL	10X Tango buffer	1X
0.5 µL	DNA (0.5–1.0 µg·µL ⁻¹)	25–50 ng·µL ⁻¹
1.0 µL	<i>NcoI</i> (10 U·µL ⁻¹)	1 U·µL ⁻¹
7.5 µL	nuclease-free water	-

The digestion mix was prepared under sterile conditions. The restriction enzyme was added last, gently mixed, and spun down for a few seconds using a bench top centrifuge. The mixture was incubated at 37°C for 2–3 h (Heraeus Function line incubator). Thermal inactivation of *NcoI* was performed at 65°C for 20 min (Grant BTA thermoblock). The mixture was analyzed on a 1% (w/v) agarose gel. Therefore, a sufficient volume of agarose gel depending on the gel size was prepared. For routine use, agarose (A9539, Sigma) was dissolved in 1X TAE (40 mM Tris, 18 mM CH₃COOH, 1 mM EDTA) in the microwave. Evaporated water was replaced, the agarose was cooled to 45°C, 10 000X gel stain (SYBR® Safe DNA gel stain; S33102, Thermo Scientific) added to 1X final concentration, and the gel was poured immediately. To 5 µL of the digestion mixture, 1 µL of 6X DNA gel loading dye (R0611, Thermo Scientific) was added, spun down, and loaded onto the agarose gel. For determination of DNA fragment sizes, 3–5 µL of DNA marker (GeneRuler™ 1 kb DNA Ladder; SM01313, Thermo Scientific) were loaded. Electrophoresis was performed in 1X TAE (40 mM Tris, 0.35% (v/v) acetic acid, 1 mM EDTA) at 90–120 V. DNA fragments were visualized by exposure to UV light (GelDoc 2000, Bio-Rad).

Table G-11. Preparation of 50X TAE

50X TAE	
242 g	Trizma® base
57.1 mL	glacial acetic acid
100 mL	0.5 M EDTA (pH 8.0)

Reagents were combined for 1 L buffer. No sterilization needed.

G I.12 Ethanol (EtOH) precipitation of DNA

The total volume of the DNA containing solution was determined and 1/10 of this volume of 3 M NaOAc (pH 5.2; 246.1 g·L⁻¹; S2889, Sigma) was added. Afterwards, 2.5 times of the combined volume of ice-cold, absolute EtOH were added. DNA was precipitated at -20°C for at least 2 h, preferably overnight. It was centrifuged at 17 000 x g, 4°C for 10–15 min. The supernatant was carefully removed without disturbing the DNA pellet. It was washed with 1 mL ice-cold 70% (v/v) EtOH, centrifuged, the EtOH carefully discarded and the pellet dried completely, for example, by evaporation at 30°C, <30 mbar for 30 min (Christ Speedvac RVC 2-25 CD plus and Alpha 2-4 LD plus).

The DNA pellet was dissolved in 50 µL nuclease-free water. The quality of DNA gel electrophoresis as described in G I.11. Optionally, DNA concentration was determined by NanoDrop® and quality checked by the A₂₆₀/A₂₈₀ absorbance ratio also measured by NanoDrop® as described in G I.10.

G I.13 Gradient PCR

The gradient PCR procedure was used to determine the optimal annealing temperature for primer pairs for target gene amplification. For the amplification of PCR products ≤2 kb and >2 kb, the *Pfu*⁺ DNA polymerase (E1118, Roboklon) and the *OptiTaq* (E2600, Roboklon) were used, respectively. The PCR mix preparations are given in **Table G-12**. The DNA polymerase was added last, spun down and the PCR mix dispensed in 6 µL aliquots in 0.2 mL PCR tubes (PP PCR tube, 0.2 mL, flat cap; 683201, Greiner Bio-one). The aliquots were spun down and the gradient PCR performed for eight temperatures per run between 40–69°C (MyCycler™ thermal cycler, Bio-Rad) with heated lid (95°C). The thermal cycle conditions for *Pfu*⁺ and *OptiTaq* are given in **Table G-13**.

Table G-12. Preparation of PCR reactions

<i>Pfu</i> ⁺ PCR mix			Final concentration	<i>OptiTaq</i> PCR mix			Final concentration
5.0 µL	10X <i>Pfu</i> buffer		1X	5.0 µL	10X Pol B buffer		1X
2.0 µL	dNTP mix (5 mM each)		0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)		0.2 mM each dNTP
2.5 µL	<i>fwd</i> primer (5 µM)		0.25 µM	2.5 µL	<i>fwd</i> primer (5 µM)		0.25 µM
2.5 µL	<i>rev</i> primer (5 µM)		0.25 µM	2.5 µL	<i>rev</i> primer (5 µM)		0.25 µM
1.0 µL	template DNA (50–100 ng·µL ⁻¹)		1–2 ng·µL ⁻¹	1.0 µL	template DNA (50–100 ng·µL ⁻¹)		1–2 ng·µL ⁻¹
1.0 µL	DMSO		2% (v/v)	1.0 µL	DMSO		2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)		2.5 U	0.5 µL	<i>OptiTaq</i> polymerase (5 U·µL ⁻¹)		2.5 U
35.5 µL	nuclease-free water		-	35.5 µL	nuclease-free water		-

Table G-13. Thermal cycle conditions for *Pfu*⁺ and *OptiTaq*

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (<i>OptiTaq</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	40 s	
Annealing	30–69	30 s	30	Annealing	30–69	30 s	30
Extension	72	1 min/1 kb		Extension	72	1 min/1 kb	
Terminal extension	72	5 min	1	Terminal extension	72	10 min	1
Hold	4	∞	1	Hold	4	∞	1

To 6 µL PCR reaction, 1.2 µL of 6X DNA gel loading dye (R0611, Thermo Scientific) were added, spun down, and loaded onto the agarose gel. Amplification products were analyzed on a 1% (w/v) agarose gel as described in G I.11. For following experiments, the PCR was repeated with the 50 µL PCR reaction mix at optimal annealing temperature.

G I.14 Colony PCR

For colony PCR, the *GoTaq*[®] polymerase (M3001) from Promega was used. Therefore, the following mixture was prepared:

Table G-14. Colony PCR mix

<i>GoTaq</i> [®] DNA polymerase mix		Final concentration
10.0 µL	5X <i>GoTaq</i> [®] reaction buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	<i>fwd</i> primer (5 µM)	0.25 µM
2.5 µL	<i>rev</i> primer (5 µM)	0.25 µM
1.0 µL	DMSO	2% (v/v)
0.5 µL	<i>GoTaq</i> [®] DNA polymerase (5 U·µL ⁻¹)	2.5 U
32.5 µL	nuclease-free water	-

The polymerase was added last, the PCR mix was spun down and dispensed in 6 μL aliquots in 0.2 mL PCR tubes (PP PCR tube, 0.2 mL, flat cap; 683201, Greiner Bio-one). To verify the insertion of target DNA into the target vector, routinely, 8 colonies potentially harboring the target plasmid were picked with a 10 μL pipette tip from an agar plate supplemented with the appropriate antibiotic. Each clone was transferred onto a replica plate. The pipette tips with a colony were dipped into a PCR mix aliquot and let stand for 1 min at RT. It was pipetted up and down twice before removing the tip. The aliquots were spun down and the PCR performed under the following thermal cycle conditions (MyCycler™ thermal cycler, Bio-Rad):

Table G-15. Thermal cycle conditions for GoTaq®

PCR step (GoTaq®)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	10 min	1
Denaturation	95	30 s	
Annealing	40–69	30 s	35
Extension	72	1 min/1 kb	
Terminal extension	72	variable	1
Hold	4	∞	1

To 6 μL PCR reaction, 1.2 μL of 6X DNA gel loading dye (R0611, Thermo Scientific) were added, spun down, and loaded onto the agarose gel. Amplification of the insert DNA was analyzed on a 1% (ω/v) agarose gel as described in G I.11. The plasmid DNA of positive clones was isolated from the corresponding colonies on the replica plate as described in G I.10 and sent for sequencing.

G I.15 Protein purification methods

G I.15.1 Semi-purification of temperature stable proteins by heat shock (HS)

Cells were cultivated and enzyme production performed under optimized conditions as summarized in G IV. Cells were harvested by centrifugation at 6 000 $\times g$, 4°C for 20 min (Sigma Laboratory Centrifuge 6k15 or 3k30), resuspended in 50 mM glycylglycine (GlyGly) buffer (pH 8.0; 6.6 $\text{g}\cdot\text{L}^{-1}$ GlyGly), frozen at -20°C, and thawed on ice. For cell lysis, 72 μL lysozyme (10 $\text{mg}\cdot\text{mL}^{-1}$ dissolved in 10 mM GlyGly buffer), 4 μL 0.25 M EDTA (pH 8.0), 1 μL benzonase (1 $\text{kU}\cdot\text{mL}^{-1}$), and 1 μL PMSF per mL cell suspension were added. Lysis was performed at 37°C with vigorous shaking (350 rpm) at for 1 h. Lysed cells were incubated at 70°C for 30 min and centrifuged (16 000 $\times g$, 4°C for 25 min). The supernatant containing thermostable proteins was transferred into a round bottom flask, snap frozen in liquid N_2 , and lyophilized under high vacuum (Christ Gamma 2-20 lyophilizer).

G I.15.2 Purification of 6xHis-tagged enzymes by IMAC

For purification, HisTrap™ FF prepacked Ni Sepharose™ columns (1 mL) were used (17-5255-01, GE Healthcare) and performed as follows: The prepacked (and precharged) column was washed with 5 mL dH_2O and 5 mL binding buffer (50 mM Tris-HCl, 40 mM imidazole, 0.5 M NaCl; pH 7.5). CFEs were prepared according to G I.5, 1:5-diluted with binding buffer, PMSF added, and filtered through a 0.2 μm cellulose acetate syringe filter. Diluted CFEs were slowly loaded onto the column. The column was washed with 10 mL binding buffer. Elution was done with elution buffer (50 mM Tris-HCl, 0.25 M imidazole, 0.5 M NaCl; pH 7.5). Fractions were collected in 2 mL tubes. The flow-through during sample loading and washing was collected and analyzed as well. Fractions were

analyzed by SDS-PAGE and eluates containing the target protein pooled and concentrated with a centrifugal membrane concentrator (10 kDa MWCO; UFC901024, Millipore).

After five purifications, the column was stripped and recharged. Therefore, the column was purged with 5 mL dH₂O and 10 mL stripping buffer (50 mM Tris-HCl, 50 mM EDTA, 0.5 M NaCl; pH 7.5). The column was washed with 5 mL binding buffer and 10 mL dH₂O. For recharging, 0.5 mL of 0.1 M Ni₂SO₄ solution (nickel sulfate heptahydrate; 203890, Sigma) were loaded. The preparation of all reagent and buffer solutions is given in **Table G-16**.

Table G-16. Reagent and buffer solutions for IMAC

Binding buffer (pH 7.5)		Elution buffer (pH 7.5)		Stripping buffer (pH 7.5)		Recharging solution	
6.06 g	Trizma® base	6.06 g	Trizma® base	6.06 g	Trizma® base	2.80 g	Ni ₂ SO ₄ · 7 H ₂ O
2.72 g	imidazole	17.02 g	imidazole	100 mL	0.5 M EDTA (pH 8.0)		
29.2 g	NaCl	29.2 g	NaCl	29.2 g	NaCl		

For preparation of the recharging solution, Ni₂SO₄ · 7 H₂O was dissolved in 80 mL dH₂O and filled up to 100 mL. All solutions and dH₂O for washing were filtered (0.2 µm cellulose acetate) before loading onto the column.

G II Materials and methods: (Advanced) cloning techniques

All reagents were purchased from commercial suppliers and used without further purification. Standard equipment and consumables were either sterile upon purchase or sterilized prior to use by autoclaving (121°C, 15 min, elevated pressure; Tuttnauer 2540EL autoclave). Unless otherwise noted, all reagent and media solutions were sterilized prior to use by autoclaving (121°C, 20 min, elevated pressure; Tuttnauer 2540EL autoclave). DNase- and RNase-free water was used throughout all protocols concerning DNA manipulations.

HPLC purified PCR primers were ordered from Sigma. Unless noted otherwise, primers were prepared as 100 µM stocks in nuclease-free water from which 5 µM working dilutions were prepared.

Sequencing was performed by LGC Genomics using the T7prom *fwd* and T7term *rev* sequencing primers provided by LGC for routine sequencing. Alternatively, Sanger sequencing was performed by Microsynth using the same standard primers.

G II.1 Optimized 'Florida' cloning

The Florida cloning procedure is an optimized, classical molecular cloning procedure employing restriction enzymes to create matching DNA overhangs that can be ligated by ligases.

G II.1.1 PCR amplification and purification of target insert DNA

PCR was used to amplify the gene of interest (MyCycler™ thermal cycler, Bio-Rad) by employing *Pfu*⁺ DNA polymerase (E1118, Roboklon). If necessary, the optimal annealing temperature of the gene-specific primer pair containing the desired restriction enzyme sites for subsequent directional cloning was determined by gradient PCR as described in G I.13. The success of amplification was determined by DNA gel electrophoresis as described in G I.11 using 5 µL of the reaction mix after PCR.

For purification, the Wizard® SV Gel and PCR Clean-Up System Kit (A9282, Promega) was used. Briefly, an equal volume (usually 45 µL) of membrane binding solution was added to the remaining PCR reaction. A minicolumn was added into the collection tube and the prepared PCR product mixture incubated on the minicolumn at RT for 1 min. It was centrifuged at >16 000 x g for 1 min (Sigma Tabletop Centrifuge 1-14). The flowthrough was discarded and the minicolumn washed with 0.7 mL membrane washing solution containing EtOH. The washing step was repeated with 0.5 mL membrane washing solution and it was centrifuged for 5 min. The flowthrough was discarded again and it was centrifuged for 1 min to remove residual washing solution. The minicolumn was transferred to a sterile 1.5 mL Eppendorf tube, 50 µL of DNase- and RNase-free water were added and incubated at RT for 1 min. The DNA was eluted by centrifugation at >16 000 x g for 1 min and kept at 4°C and -20°C for immediate use and long-time storage, respectively. Before subsequent cloning, 5 µL of purified DNA were analyzed on a 1% (w/v) agarose gel as described in G I.11. Optionally, DNA concentration was determined by NanoDrop®.

G II.1.2 Trimming of target insert DNA

Unless otherwise noted, trimming was achieved by using a restriction enzyme (double) digestion taking advantage of the introduced restriction enzyme sites *via* PCR as described in G II.1.1. (Double) digestions were performed in accordance to the optimal conditions suggested by suppliers. Trimmed DNA was purified as described in G II.1.1.

G II.1.3 Restriction enzyme (double) digestion of target vector and gel purification

The target vector DNA was (double) digested to create overhangs matching the trimmed target DNA insert. The digested target vector was separated on 1% (w/v) agarose as described in G I.11. The linearized vector DNA was visualized by exposure to UV light in preperative mode (UV Transilluminator 2000, Bio-Rad) and the corresponding band cut out with a sterile scalpel. The gel slice was transferred to a sterile 1.5 mL Eppendorf tube and frozen either at -80°C for 30 min or -30°C for 1.5 h. The gel slice was thawed at 37°C and 550 rpm (Eppendorf Thermomixer Comfort) for 15 min and centrifuged at 17 000 x g, 4°C for 10 min (Sigma Laboratory Centrifuge 3k15). The supernatant was transferred to a fresh tube. The remaining gel slice was mixed with 100 µL nuclease-free water, the gel slice disturbed with a pipette tip, frozen and thawed again. It was centrifuged at 17 000 x g, 4°C for 10 min. The volume of the combined supernatants was determined and DNA precipitated by EtOH precipitation as described in G I.12.

G II.1.4 Sticky end ligation of target insert and vector DNA

For sticky end ligation, the T4 DNA ligase (EL0014) from Thermo Scientific was used. Therefore, the following reaction mixture was prepared:

Table G-17. Sticky end ligation mix

	T4 DNA ligation mix	Final concentration
2.0 µL	10X T4 DNA ligase buffer	1X
1.0 µL	linearized vector DNA (20–100 ng·µL ⁻¹)	1–5 ng·µL ⁻¹
1.0 / 4.0 µL	trimmed insert DNA	Varying
0.5 µL	T4 DNA ligase (5 Weiss U·µL ⁻¹)	2.5 Weiss U
12.5 / 15.5 µL	nuclease-free water	-

Routinely, 1:1 and 1:4 volumetric mixtures of vector and insert were prepared. The ligase was added last, the mixtures spun down and incubated at RT (22–26°C) for 20 min. Up to 5 µL of each mixture were used for the transformation of 50 µL chemically competent cells.

G II.1.5 Chemical transformation

Transformation of chemically competent cells was performed as described in G I.9.1.

G II.1.6 Plasmid assembly verification by colony PCR

Routinely, 8 colonies were picked from agar plates supplemented with the appropriate antibiotic. Colony PCR was performed as described in G I.14 and the plasmid DNA of positive clones was sent for sequencing. Plasmid DNA isolation was performed according to the procedure in G I.10.

G II.2 Sequence- and ligation-independent cloning (SLIC) methods

G II.2.1 FastCloning (FC)

FC is an *in vivo* DNA assembly method and was used in this thesis to construct functional plasmids by joining two DNA fragments *via* homologous DNA overhangs.^[87]

G II.2.1.1 Primer design

The desired plasmid harboring the target gene was assembled *in silico* using the software Geneious® 6.1.8. Primer pairs for both the plasmid backbone and the target insert were designed fulfilling the following criteria: Primers had a total length of ≤45 of which ≥10 bp were matching the DNA sequence to be amplified and ≥15 bp contained homologous overhangs for directed assembly of DNA fragments. Primers contained terminal GC pairs and had a GC content of roughly 50%. The estimated T_m can be ≥65°C as calculated by the software since annealing is solely directed by the oligonucleotide portion specific for the target DNA.

Designed primer sequences are given in the corresponding cloning procedures and summarized in **Table H-1**.

G II.2.1.2 PCR amplification and processing of target DNA fragments

The optimal annealing temperatures for each primer pair and the corresponding DNA template were determined by gradient PCR and the amplification products were analyzed on 1% (w/v) agarose as described in G I.13. The PCRs were repeated at optimal annealing temperatures. The amplified backbone and the target insert DNA were mixed in volumetric backbone to insert-ratios of 1:0, 1:1, 1:4, and 1:10. To destroy methylated DNA templates from previous PCRs, the mixtures were applied to *DpnI* (FD1703, Thermo Scientific) digestion. The 1:0-mixture served as *DpnI* digestion control for subsequent transformation.

Table G-18. *DpnI* digestion mix for two unpurified PCR products

	FastDigest® <i>DpnI</i> mix	Final concentration
2.0 µL	10X FastDigest® buffer	0.67X*
5–8.0 µL	amplified target insert DNA	-
2–5.0 µL	amplified backbone DNA	-
1.0 µL	<i>DpnI</i>	-
17.0 µL	nuclease-free water	-

* Only 2 µL of 10X buffer required for unpurified PCR product in a 30 µL reaction volume according to the supplier.

DpnI digestion was performed at 37°C for 10 min and the enzyme was inactivated at 80°C for 5 min. Each DNA mixture was purified using the QIAquick® PCR Purification Kit (28104, QIAGEN). Briefly, five volumes of binding buffer (QIAquick PB buffer) were added to one volume of each DNA mixture. Optionally, 10 µL 3 M NaOAc (pH 5.0; 408.2 g·L⁻¹ NaOAc · 3 H₂O; S8625, Sigma) were added if the color of the mixture was orange or violet. The minicolumn was placed in a 2.0 mL collection tube. It was centrifuged at ≥16 000 x g for 1 min (Sigma Tabletop Centrifuge 1-14) and the flow-through discarded. For washing, 0.75 mL washing buffer (QIAquick PE buffer) were added and centrifuged for 1 min. The flow-through was discarded and the minicolumn centrifuged again for 1 min to remove residual washing buffer. The minicolumn was transferred to a sterile 1.5 mL Eppendorf tube. For DNA elution, 30 µL nuclease-free water were added onto the minicolumn, incubated at RT for 2 min and centrifuged. Optionally, 0.5 µL of 6X DNA loading dye (R0611, Thermo Scientific) were added to 2.5 µL of DNA mixture and purity checked by DNA electrophoresis on 1% (w/v) agarose as described in G I.11.

G II.2.1.3 *In vivo* assembly of DNA fragments

Chemically competent *E. coli* TOP10 were prepared using RbCl. For transformation, 4 µL of DNA mixture per 100 µL competent cells were used. Preparation of chemically competent cells and transformation was performed as described in G I.9.3. After recovery, 0.1 mL and 0.4 mL of the resulting cell culture, respectively, were plated on prewarmed LB agar plates supplemented with the appropriate antibiotic. Plates were incubated upside down at 37°C for 16–24 h (Heraeus Function line incubator). Besides the negative control described above, transformation controls were performed as in G I.9.1.

Colony PCR was performed as described in G I.14 to verify the presence of target insert DNA. Plasmid DNA was then isolated from positive TOP10 clones according to G I.10 and sent for sequencing.

E. coli BL21(DE3) cells were transformed as described in G I.9.1 for subsequent expression studies.

G II.2.2 Seamless and ligation-independent cloning extract (SLiCE) approach

SLiCE is an *ex vivo* DNA assembly method and was used in this thesis to construct ready-to-transform plasmids from up to three DNA fragments.^[97]

G II.2.2.1 SLiCE preparation

A preculture of *E. coli* TOP10 was prepared in 20 mL 2X YT medium (pH 7.0; 16 g·L⁻¹ tryptone, 10 g·L⁻¹ yeast extract, 5 g·L⁻¹ NaCl) in a 100 mL baffled flask at 37°C with shaking (200 rpm; InforsHT Multitron 2 Standard) overnight. For the main culture, 100 mL 2X YT medium were inoculated with 2 mL preculture in a 500 mL baffled flask. Cells were grown at 37°C with shaking (200 rpm) until an OD₅₉₀ = 3.0 was reached (WPA colourwave, CO7500 Colorimeter). The remaining culture was harvested by centrifugation at 5 000 x g, 4°C for 20 min (Sigma Laboratory Centrifuge 6k15 or 3k30). The supernatant was discarded, the pellet washed with 100 mL ice-cold dH₂O and centrifuged at 5 000 x g, 4°C for 20 min. The supernatant was discarded again and the wet weight of the cell pellet determined. The cell pellet was resuspended in 0.3 mL CellLytic™ B Cell Lysis Reagent (B7435, Sigma) per 0.25 g cell pellet. It was briefly vortexed (IKA® Vortex 4 basic), spun down and cells were transferred into a sterile 1.5 mL Eppendorf tube. Lysis was performed at RT for 10 min. Lysates were centrifuged at 20 000 x g, RT for 3 min (Sigma Laboratory Centrifuge 3k15). The supernatant was carefully transferred into a new 1.5 mL tube and mixed with an equal volume of glycerol (G9012, Sigma). The SLiCE was dispensed into 50 µL aliquots in 0.5 mL tubes and stored at -80°C.

G II.2.2.2 Primer design

The desired plasmid harboring the target gene(s) was assembled *in silico* using the software Geneious® 6.1.8. Primer design followed the rules described in G II.2.1.1.

G II.2.2.3 Preparation of DNA fragments for assembly by SLiCE

DNA fragments were PCR amplified with homologous overhangs for directed assembly as in the FC procedure (G II.2.1). After amplification of all DNA fragments at optimal annealing temperatures, the target insert(s) and the backbone DNA were mixed in volumetric ratios of 1:(1):1 and 4:(4):1 [insert #1:(insert #2):backbone] and the mixtures applied to *DpnI* (FD1703, Thermo Scientific) digestion to destroy methylated DNA templates from previous PCR. The *DpnI* digestion mix for two DNA fragments was prepared as described above (**Table G-18**) and is given below for three DNA fragments. For later negative controls, additionally, the backbone DNA was digested both without any insert DNA if assembling two fragments and with one DNA insert of choice if assembling three fragments, respectively.

Table G-19. *DpnI* digestion mix for three unpurified PCR products

FastDigest® <i>DpnI</i> mix		Final concentration
2.0 µL	10X FastDigest® buffer	0.67X*
3–4.0 µL	amplified DNA insert #1	-
3–4.0 µL	amplified DNA insert #2	-
1–3.0 µL	amplified backbone DNA	-
1.0 µL	<i>DpnI</i>	-
8.0 µL	nuclease-free water	-

* Only 2 µL of 10X buffer required for unpurified PCR product in a 30 µL reaction volume according to the supplier.

DpnI digestion was performed at 37°C for 10 min and *DpnI* inactivated by heating at 80°C for 5 min. Each DNA mixture was purified using the QIquick® PCR Purification Kit (28104, QIAGEN).

G II.2.2.4 SLiCE reaction

The 10X SLiCE buffer (0.5 M Tris-HCl, 0.1 M MgCl₂, 10 mM DTT; pH 7.5) was thawed prior to use. The SLiCE reaction mix was prepared as given in the table below:

Table G-20. SLiCE reaction mix

SLiCE assembly mix	Final concentration	10X SLiCE buffer (pH 7.5)
1.0 µL 10X SLiCE buffer	1X	0.605 g Trizma® base
4–7.0 µL DNA fragment mix	-	0.095 g MgCl ₂
1.0 µL ATP (10 mM)	1 mM	0.015 g DTT
1.0 µL SLiCE (from TOP10)	10% (v/v)	
0–3.0 µL nuclease-free water	-	

Amounts for the 10X SLiCE buffer are for the preparation of 10 mL. The buffer was filter-sterilized (0.2 µm cellulose acetate), dispensed into 1 mL aliquots, and stored at -20°C

The extract was added last, the assembly mix spun down, and incubated at 37°C for 1 h. For direct transformation of *E. coli* BL21(DE3) cells, 1–2 µL of the SLiCE reaction mix were used per 50 µL of competent cells. Preparation of competent cells and transformation were performed as described in G I.9.3. For plasmid assembly verification, colony PCR was performed as described in G I.14. Plasmid DNA was isolated from positive *E. coli* BL21(DE3) clones according to G I.10 and sent for Sanger sequencing.

G III Materials and methods: Biotransformations

G III.1 Biotransformations employing CFEs

Enzyme production was performed under optimized expression conditions for single enzymes as summarized in G IV. The preparation of CFEs was conducted as described in G I.5 and the total amount of protein determined as in G I.7. Prior to biotransformations, protein production was checked by SDS-PAGE as in G I.8.

Routinely, biotransformations were performed in 2.0 mL total volume containing 5 mg·mL⁻¹ CFE and 4 mM substrate. For biotransformations employing a redox enzyme, 4.25 mM of the appropriate cofactor were added from 100 mM stock solutions. Biotransformations were performed in 50 mM Tris-HCl (pH 7.5). The biotransformation mix was prepared as follows:

Table G-21. Biotransformation mix using CFEs

	Biotransformation mix	Final concentration
x μL	CFE (usually 7–35 mg·mL ⁻¹)	5 mg·mL ⁻¹
80 μL	substrate (100 mM)	4 mM
85 μL	cofactor (100 mM)	4.25 mM
y μL	Tris-HCl (50 mM; pH 7.5)	-

The substrate was added last. Immediately after adding the substrate, 100 μL of a t_0^* sample were taken and added to 200 μL of EtOAc containing 1 mM of methylbenzoate (115 μL·L⁻¹; 10785, Merck) as internal standard (IS) in a 1.5 mL Eppendorf tube. The sample was extracted by vortexing at maximal speed for 35 s (IKA® Vortex 4 basic) and spun down for 35 s using a bench top centrifuge. The organic layer was transferred into a fresh 1.5 mL tube. The sample was extracted a second time with 190 μL EtOAc containing 1 mM IS and 10 μL 2 M HCl. The combined organic layers were desiccated over Na₂SO₄ and transferred into a 1.5 mL GC vial (548-0003, VWR) with a 0.1 mL micro-insert (548-0006, VWR). The GC vial was crimped with an aluminium cap with rubber septum (548-0010, VWR) and the organic phase was analyzed by achiral GC. GC analysis was carried out on a Thermo Focus GC/FID detector and the method 'STD_Achiral_100-300_6 min' used (initial temperature of 80°C, hold 0.5 min; ramp 1 at 40°C·min⁻¹ to 220°C; ramp 2 at 80°C·min⁻¹ to 300°C, hold 1 min). The injected sample volume was 1 μL.

Samples were then taken immediately after mixing (t_0^*), 2 h, and 24 h as described above.

G III.2 Standard screening conditions: Biotransformations using RCs

Enzyme production was performed under optimized expression conditions for every individual enzyme or coexpressed enzymes as summarized in G IV and the corresponding subsections of results and discussion chapters. The preparation of RCs was done as described in G I.6. Routinely, protein production was checked by SDS-PAGE as described in G I.8 for whole cell samples. Usually, 10 μL of RCs were incubated with 40 μL of SDS sample buffer at 95°C for 10 min and 15 μL of the resulting solution loaded onto the gel.

Routinely, biotransformations with RCs were performed in a total volume of 2.0 mL at OD₅₉₀ = 10.0 starting with 5 mM substrate. Substrates were added from 100 mM stocks in organic solvent, usually ACN. Consequently,

biotransformations were performed in the presence of 5% (v/v) organic solvent. The biotransformation mix was prepared as follows in 8 mL reaction vials:

Table G-22. Biotransformation mix employing RCs

Biotransformation mix		Final concentration
1.0 mL	RCs (OD ₅₉₀ = 20)	OD ₅₉₀ = 10
0.1 mL	substrate (100 mM)	5 mM
0.9 mL	RCM	-

The substrate was added last and screenings were performed at 25°C with shaking (250 rpm). Immediately after adding the substrate, the vial was closed and inverted five times and a t_0^* sample (100 μ L) was taken. The 100 μ L were transferred into a 1.5 mL Eppendorf tube already containing 200 μ L of EtOAc supplemented 1 mM of methylbenzoate (115 μ L·L⁻¹; 10785, Merck) as IS and 10 μ L 2 M HCl. The sample was extracted by vortexing at maximal speed for 35 s (IKA® Vortex 4 basic) and spun down for 1 min. The organic layer was transferred into a fresh 1.5 mL tube. The sample was extracted a second time with 190 μ L EtOAc containing 1 mM IS. The combined organic layers were desiccated over Na₂SO₄ and transferred into a 1.5 mL GC vial (548-0003, VWR) with a 0.1 mL micro-insert (548-0006, VWR). The GC vial was crimped with an aluminium cap with rubber septum (548-0010, VWR) and the organic phase analyzed by achiral GC/FID as before (G III.1).

Samples were usually taken after mixing (t_0^*), 1 h, 2 h or 3 h, 6 h, 15 h, and 24 h as described above. Samples for HPLC analysis were collected at the same time points as follows:

An aliquot (200 μ L) of the reaction mixture was transferred into 1.5 mL Eppendorf tubes and centrifuged (6 000 x g, 4°C, 10 min; Sigma Laboratory Centrifuge 3k15). The supernatant was transferred into a new 1.5 mL tube and if not analyzed immediately, stored at -80°C. The supernatant was filtered through a 0.2 μ m PTFE membrane syringe filter (4552T, Pall Life Sciences) into a 1.5 mL glass vial equipped with a 0.1 mL micro-insert, crimped, and analyzed by HPLC. LC/MS analysis was performed by HPLC (Nexera Shimadzu). The injected volume of the samples was 10 μ L. Supernatants were analyzed with the photodiode array detector (PDA) for quantification of analytes at λ = 190 nm; the refractive index (RI) detector and the electrospray ionization (ESI) ion source with a quadrupole mass analyzer (LC/MS 2020 Shimadzu) for additional confirmation of the substances and overflow metabolites. Separation was performed with an ROA-Organic Acid H⁺ (8%) column (150 x 7.8 mm, Phenomenex) with an isocratic flow of 0.5 mL·min⁻¹ of 0.1% (v/v) formic acid in water (HPLC grade).

For HPLC measurement of samples containing extracellular enzymes, for example, protein precipitation was performed as follows: Supernatants were quenched with two volumes of ice-cold acetonitrile (ACN), snap-frozen in liquid N₂, thawed, centrifuged (18 000 x g, 4°C, 10 min), and filtered through a PTFE (0.2 μ m) syringe membrane before analysis.

G IV Strain and enzyme library

The following table summarizes the optimized expression conditions for single enzymes produced in different hosts in this thesis. If not noted otherwise, precultures were grown in LB-Miller medium and cells cultivated in TB medium (supplemented with antibiotics if applicable). The procedure for expression in AIM is described in G VI.3.1.1. The optimized conditions for the coproduction of pathway enzymes in the same host cell are given in

G VIII. Unless noted otherwise, engineered strains are cultivated in the presence of the appropriate antibiotic(s) at standard working concentrations given in **Table G-1**.

Table G-23. Growth and expression conditions for all enzymes used in this thesis in alphabetical order

#	Enzyme / Strain	Enzyme class	Vector	Marker [*]	Inducer	Inducer concentration	Expression conditions			
							[OD ₅₉₀]	[°C]	[rpm]	[h]
1	<i>3FCR</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
2	<i>3GJU</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
3	<i>3HMU</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
4	<i>3IST</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
5	<i>ADH-A</i> / BL21(DE3)	ADH	pET22b(+)	amp ^R	IPTG	2 mM	≈ 12.0	20	120	24
6	<i>ADH-A</i> / DH5α	ADH	pET22b(+)	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7	<i>ADH-ht</i> / BL21(DE3)	ADH	pET26b(+)	kan ^R	IPTG	0.5 mM	0.5	20	120	22
8	<i>ADH-ht</i> / DH5α	ADH	pMA-T	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
9	<i>ADH_{Lk}</i> / BL21(DE3)	ADH	pET21b(+)	kan ^R	IPTG	1 mM	0.5	30	120	20
10	<i>ADH_{Rr}</i> / BL21(DE3)	ADH	pRR	cam ^R	IPTG	25 μM	0.3	25	120	22
11	<i>AlkJ</i> / BL21(DE3)	ADH	pKA1	cam ^R	IPTG	0.5 mM	0.5	25	120	20–22
12	<i>AlkJ</i> / DH5α	ADH	pGEc47	tet ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
13	<i>AlkJ</i> / DH5α	ADH	pKA1	cam ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
14	<i>AlkJ</i> / RARE	ADH	pKA1	cam ^R	IPTG	0.5 mM	0.5	25	120	20–22
15	<i>AlkJ_{trnc}</i> / BL21(DE3)	ADH	pKA1	cam ^R	α-Lac	0.2% (w/v)	Coexpression from pKA1_alkJ _{trnc} ::fsa1-A129S (G VI.2.4.3)			
16	<i>AlkL</i> / BL21(DE3)	OMP	pCOM	amp ^R	DCPK	0.05% (w/v)	0.5–1.0	n.d.	n.d.	n.d.
17	<i>AlkL</i> / DH5α	OMP	pCOM	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
18	<i>AspFum</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
19	<i>AspTer</i> / BL21(DE3)	ω-TA	pGASTON	amp ^R	L-Rhm	0.2% (w/v)	0.7	20	200	20
20	<i>B0014</i> / DH5α	T _{Syn}	pUC57	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
21	<i>BS2</i> / BL21(DE3)	Esterase	pET28a	kan ^R	IPTG	0.5 mM	0.2–0.4	25	200	20–24
22	<i>CAR_{Ni}</i> / BL21(DE3)-Gold	CAR	pETDuet-1	amp ^R	α-Lac	0.2% (w/v)	AIM	20	150	20
23	<i>CAR_{Ni}</i> / RARE	CAR	pETDuet-1	amp ^R	α-Lac	0.2% (w/v)	AIM	20	150	20
24	<i>CAR_{Mm}</i> / BL21(DE3)-Star	CAR	pETDuet-1	amp ^R	α-Lac	0.2% (w/v)	AIM	20	150	20
25	<i>CAR_{Mm}</i> / RARE	CAR	pETDuet-1	amp ^R	α-Lac	0.2% (w/v)	AIM	20	150	20
26	<i>DhaK</i> / BL21(DE3)	Kinase	pRSETa	amp ^R	IPTG	0.5 mM	0.5–0.8	30	150	18–22
27	<i>DhaK</i> / DH5α	Kinase	pRSETa	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
28	<i>DhaK</i> / RARE	Kinase	pRSETa	amp ^R	IPTG	0.5 mM	0.5–0.8	30	150	18–22
29	<i>DhaK_{trnc}</i> / BL21(DE3)	Kinase	pRSETa	amp ^R	IPTG	0.5 mM	0.5–0.8	30	150	18–22
30	<i>DhaK_{trnc}</i> / DH5α	Kinase	pET22b(+)	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
31	<i>FucA</i> / BL21(DE3)	Aldolase	pKA1	cam ^R	α-Lac	0.2% (w/v)	AIM	25	150	20–22
32	<i>FucA</i> / RARE	Aldolase	pKA1	cam ^R	α-Lac	0.2% (w/v)	AIM	25	150	20–22
33	<i>FucA</i> / BL21(DE3)	Aldolase	pKK	amp ^R	IPTG	0.1 mM	0.3	25	150	20–22
34	<i>FruA</i> / BL21(DE3)	Aldolase	pKK	amp ^R	IPTG	0.1 mM	0.3	25	150	20–22
35	<i>Fsa1</i> / BL21(DE3)	Aldolase	pET16b	amp ^R	IPTG	0.5 mM	0.5	30	200	14–18

36	<i>Fsa1</i> / JM109(DE3)	Aldolase	pET16b	amp ^R	IPTG	0.5 mM	0.5	30	200	14–18
37	<i>Fsa1-A129S</i> / BL21(DE3)	Aldolase	pET16b	amp ^R	IPTG	0.5 mM	0.5	30	200	14–18
38	<i>Fsa1-A129S</i> / JM109(DE3)	Aldolase	pET16b	amp ^R	IPTG	0.5 mM	0.5	30	200	14–18
39	<i>GDH_{2xBS}</i> / BL21(DE3)	GDH	pET26b(+)	kan ^R	IPTG	0.1 mM	0.3–0.5	37	150	16–20
40	<i>GDH_{2xBS}</i> / DH5α	GDH	pET26b(+)	kan ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
41	<i>GDH_{7xBS}</i> / BL21(DE3)	GDH	pET26b(+)	kan ^R	IPTG	0.1 mM	0.3–0.5	37	150	16–20
42	<i>GDH_{7xBS}</i> / DH5α	GDH	pET26b(+)	kan ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
43	<i>LK-ADH_{N-int}</i> / BL21(DE3)	ADH	pTYB21	amp ^R	IPTG	1 mM	0.5	15	150	40
44	<i>LK-ADH_{N-int}</i> / DH5α	ADH	pTYB21	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
45	<i>LK-ADH_{N-int}</i> / ER2566	ADH	pTYB21	amp ^R	IPTG	1 mM	0.5	15	150	40
46	<i>MycVan</i> / BL21(DE3)	ω-TA	pGASTON	amp ^R	L-Rhm	0.2% (ω/v)	0.7	20	200	20
47	<i>NeoFis</i> / BL21(DE3)	ω-TA	pET22b(+)	amp ^R	IPTG	0.1 mM	0.7	20	200	20
48	<i>OhyA2</i> / BL21(DE3)	Hydratase	pET28a	kan ^R	IPTG	0.1 mM	0.6	16	150	16
49	<i>OhyA2</i> / DH5α	Hydratase	pET28a	kan ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
50	<i>PDC_{Ap}</i> / BL21(DE3)	PDC	pET22b(+)	amp ^R	α-Lac	0.2% (ω/v)	AIM	20	150	18–20
51	<i>PDC_{Ap}</i> / DH5α	PDC	pET22b(+)	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
52	<i>PEST</i> / BL21(DE3)	Esterase	pET21a	amp ^R	IPTG	0.25 mM	0.2–0.4	25	200	20–24
53	<i>Pfel</i> / BL21(DE3)	Esterase	pGASTON	amp ^R	L-Rhm	0.2% (ω/v)	0.2–0.4	37	200	3
54	<i>PhoN-Se V78L</i> / BL21(DE3)	<i>PhoN</i>	pET26b(+)	kan ^R	IPTG	0.5 mM	0.5	37	200	5–6
55	<i>PhoN-Se V78L</i> / DH5α	<i>PhoN</i>	pMA-T	amp ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
56	<i>PhoN-Se V78L_{SP}</i> / BL21(DE3)	<i>PhoN</i>	pET26b(+)	kan ^R	IPTG	0.5 mM	0.5	37	200	5–6
57	<i>PhoN-Se V78L_{3xFLAG}</i> / BL21(DE3)	<i>PhoN</i>	pET26b(+)	kan ^R	IPTG	0.5 mM	0.5	37	200	5–6
58	<i>PhoN-Sf</i> / BL21(DE3)	<i>PhoN</i>	pET26b(+)	kan ^R	IPTG	0.5 mM	3.0	37	200	5
59	<i>PhoN-Sf</i> / DH5α	<i>PhoN</i>	pET26b(+)	kan ^R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
60	<i>PhoN-Sf</i> / RARE	<i>PhoN</i>	pET26b(+)	kan ^R	IPTG	0.5 mM	0.5	37	200	5–6
61	<i>Pptase_{Ec}</i> / BL21(DE3)-Gold or -Star	<i>Pptase</i>	pETDuet-1	amp ^R	α-Lac	0.2% (ω/v)	Coexpression with CARs from pETDuet-1 (entry #22–25)			
62	<i>Pptase_{Ec}</i> / RARE	<i>Pptase</i>	pETDuet-1	amp ^R	α-Lac	0.2% (ω/v)	Coexpression with CARs from pETDuet-1 (entry #22–25)			
63	<i>RhuA</i> / BL21(DE3)	Aldolase	pKK	amp ^R	IPTG	0.1 mM	0.3	25	150	20–22
64	<i>VfjH6</i> / BL21(DE3)	ω-TA	pET24b	kan ^R	IPTG	0.1 mM	0.7	20	200	20
65	<i>YqaB</i> / BL21(DE3)	<i>PhoN</i>	pCDF Duet-1	str ^R	IPTG	0.1 mM	0.5	25	200	14–18
66	<i>YqjM</i> / BL21(DE3)	ERED	pET21b	amp ^R	IPTG	0.1 mM	(G VI.8.1)	20	120	16
67	<i>YqjM</i> / BL21(DE3)	ERED	pET22b(+)	amp ^R	IPTG	0.1 mM	(G VI.8.1)	20	120	16
69	<i>YqjM</i> / BL21(DE3)	ERED	pET28a	kan ^R	IPTG	0.1 mM	(G VI.8.1)	20	120	16
70	<i>YqjM</i> / BL21(DE3)	ERED	pHT	amp ^R	IPTG	0.1 mM	(G VI.8.1)	20	120	16
71	<i>YqjM</i> / BL21(DE3)	ERED	pSF1	amp ^R	IPTG	0.1 mM	(G VI.8.1)	20	120	16
72	BL21(DE3)	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
73	BL21(DE3) Δ <i>nema</i>	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
74	BL21(DE3) Δ <i>nema</i> Δ <i>fadH</i>	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24

75	BL21(DE3)-Gold	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
76	BL21(DE3)-Star	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
77	DH5 α	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
78	ER2566	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
79	JM109	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
80	JM109(DE3)	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
81	TOP10	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24
82	RARE	-	n.a.	n.a.	n.a.	n.a.	n.a.	37	200	12–24

[*] Antibiotic concentrations (Amp: 100 $\mu\text{g}\cdot\text{mL}^{-1}$; Cam: 37 $\mu\text{g}\cdot\text{mL}^{-1}$; Cl-Tet: 15 $\mu\text{g}\cdot\text{mL}^{-1}$; Kan: 50 $\mu\text{g}\cdot\text{mL}^{-1}$; Str: 25 $\mu\text{g}\cdot\text{mL}^{-1}$); n.a. = not applicable

G V ‘Florida’ cloning trouble shooting: The establishment of a reliable molecular cloning procedure

Classical molecular cloning employs iterative rounds of restriction enzyme digestion of target DNA fragments and their ligation by ligases. Both the amount and the purity of linear fragments are crucial for the ligation efficiency. Due to poor ligation efficiencies in our lab, a methodical investigation of all cloning steps was executed. For test cloning, the *adh-ht* gene was sub-cloned into pET26b(+) utilizing the two restriction sites *NdeI* and *XhoI*.

All reagents were purchased from commercial suppliers and used without further purification and were either sterile upon purchase or sterilized by autoclaving. DNase- and RNase-free water was used throughout all protocols concerning DNA manipulations. HPLC purified PCR primers were ordered from Sigma. Sequencing was performed by LGC Genomics using, if not noted otherwise, the T7prom *fwd* and T7term *rev* sequencing primers provided by LGC.

G V.1 The purification of PCR amplified target inserts by commercial kits

The *adh-ht* gene was PCR amplified as described below (G VI.2.2.1). After PCR, the mixture was split into three 15 μL aliquots and the PCR product purified with the following commercially available kits:

1. Wizard® SV Gel and PCR Clean-Up System kit from Promega (A9282)
2. innuPREP PCRpure kit from Analytik Jena (845-KS-5010010)
3. GeneJET PCR Purification kit from Thermo Scientific (K0699)

The purified PCR products were analyzed on 1% (ω/v) agarose as described in G I.11. All kits performed equally well according to DNA electrophoresis (**Figure G-2**).

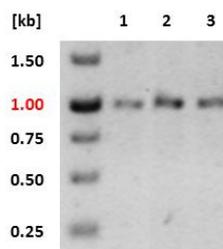


Figure G-2. PCR product purification by different commercial kits. The purification of the amplified *adh-ht* insert (calculated size: 1.03 kb; experimental: 1.00 kb) performed with PCR purification kits from Promega (1), Analytik Jena (2), and Thermo Scientific (3).

G V.2 Gel purification of target vector DNA as a crucial step in molecular cloning

For exemplary subcloning of the *adh-ht* gene into pET26b(+), the target vector was isolated from *E. coli* DH5 α transformants by the PureYield™ Plasmid Midiprep System (A2492, Promega). Briefly, 50 mL LB-Miller medium supplemented with Kan (50 $\mu\text{g}\cdot\text{mL}^{-1}$) were inoculated with a single colony of *E. coli* DH5 α /pET26b(+) and grown at 37°C, 200 rpm overnight. Plasmid DNA was isolated as follows: Cells were pelleted at 5 000 x g, 4°C (Sigma Laboratory Centrifuge 6k15) for 10 min. Cells were resuspended in resuspension solution (3 mL). Cells were mixed with cell lysis buffer (3 mL) and lysis performed at RT for 3 min. Neutralization solution was added, the resulting solution mixed by inversion, and the lysate centrifugated at 16 000 x g, 24°C for 20 min. DNA purification from the lysate was performed by vacuum filtration according to the supplier. Plasmid DNA was eluted with 600 μL nuclease-free water by centrifugation (2 000 x g, RT for 5 min; Sigma Laboratory Centrifuge 6k15 with a swinging bucket rotor). DNA concentration (202 $\text{ng}\cdot\mu\text{L}^{-1}$) was determined by NanoDrop®.

G V.2.1 Commercially available kits

For gel purification experiments using commercial kits, 1 μg pET26b(+) was linearized by the restriction enzymes *NdeI* and *XhoI* as described below (G VI.2.2.1). After digestion, vector DNA was separated on 1% (ω/v) agarose by electrophoresis as described in G I.11. The resulting agarose gel was kept in the gasket, the target DNA band visualized by exposure to UV light in preperative mode (GelDoc 2000, Bio-Rad), and excised. For subsequent gel purification, the following kits were used:

1. Wizard® SV Gel and PCR Clean-Up System kit from Promega (A9282)
2. innuPREP Gel Extraction kit from Analytik Jena (845-KS-5010010)
3. GeneJET Gel Extraction kit from Thermo Scientific (K0699)

Gel purification was performed in triplicates following the corresponding instruction manual, DNA concentration determined by NanoDrop® after elution, and the presence of linearized plasmid DNA checked on 1% (ω/v) agarose. In contradiction to experimental data provided by the suppliers, gel purification was not successful and linearized vector DNA could not be recovered by the commercial kits tested (**Figure G-3A**).

G V.2.2 Agarose gel disruption and EtOH precipitation of vector DNA

For agarose gel purification, 3 μg pET26b(+) were linearized by *NdeI* and *XhoI* as described below (G VI.2.2.1). After digestion, the linearized vector was separated on 1% (ω/v) agarose by electrophoresis (G I.11) and excised. The gel slice was transferred into a sterile 1.5 mL tube and frozen at -80°C for 30 min. Subsequently, the gel slice was thawed at 37°C for 10 min with shaking (550 rpm; Eppendorf Thermomixer Comfort), centrifuged (16 000 x g, 4°C for 10 min), and the supernatant transferred into a fresh 1.5 mL tube. To the remaining gel

debris, nuclease-free water (100 μ L) was added, the gel disturbed with a pipette tip, and frozen/ thawed as before. After centrifugation (16 000 x g, 4°C for 5–10 min), the volume of the combined supernatants was determined. DNA precipitation was performed as outlined in G I.12. The dried DNA pellet was dissolved in 30 μ L nuclease-free and purity checked on 1% (ω/v) agarose (**Figure G-3B**).

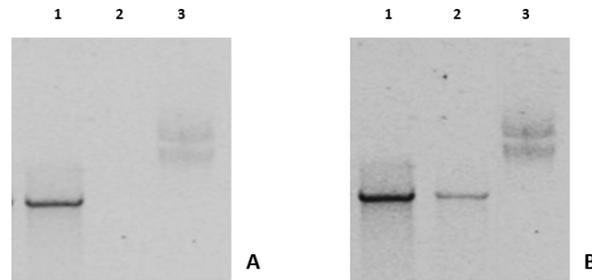


Figure G-3. DNA purification and recovery from agarose gels. (A) Commercial kits. (B) Agarose gel disruption and EtOH precipitation. The vector pET26b(+) was digested (A: 1 μ g; B: 3 μ g). The presence of the linearized pET26b(+) was checked before (1) and after purification (2). The undigested vector was loaded as a control (3); 1% (ω/v) agarose.

G V.3 Testing of ligases from different suppliers

For subcloning into pET26b(+), the *adh-ht* insert was amplified by PCR and digested with *NdeI* and *XhoI* as outlined below (G VI.2.2.1). The ligation mixture was prepared according to **Table G-34**. Two T4 DNA ligases were tested:

1. T4 DNA ligase from Promega (M180B)
2. T4 DNA ligase from Thermo Scientific (E0011)

For sticky end ligation, mixtures were incubated at RT for 3 h and 10 min in the presence of the ligase (2.5 Weiss U) from Promega and Thermo Scientific, respectively. As a control, linearized pET26b(+) was incubated accordingly with either of the two ligases in the absence of insert DNA. Transformation of chemically competent *E. coli* BL21(DE3) was performed in triplicates as described in G I.9.1. As expected, transformations with the control ligation mixture did not grow colonies. The number of colonies after transformation for both ligation mixtures were comparably high:

1. 42 ± 4
2. 50 ± 6

In conclusion, the crucial step of Florida cloning was the purification of the linearized pET26b(+) vector DNA and its recovery by agarose gel disruption and subsequent EtOH precipitation (**Figure G-3B**). The tested kits for insert purification after PCR performed comparably well; the short ligation time at RT of only 10 min for the T4 DNA ligase from Thermo Scientific and the slightly higher transformation efficiency made it superior to the ligase from Promega under experimental conditions.

G VI Single enzyme characterizations

G VI.1 Characterization of esterases to prepare primary alcohols *in situ*

G VI.1.1 Cloning and expression of *BS2*, *Pfel*, and *PEST*

- BS2:** Esterase from *Bacillus subtilis* (GenBank: AQZ92317.1)
Gene size: 1 494 bp AA: 497 SDS-PAGE: 55 kDa (experimental)
- Pfel:** Esterase from *Pseudomonas fluorescens* (GenBank: AAB60168.1)
Gene size: 825 bp AA: 275 SDS-PAGE: 26 kDa (experimental)
- PEST:** Esterase from *Pyrobaculum calidifontis* (GenBank: AB078331.1)
Gene size: 942 bp AA: 313 SDS-PAGE: 33 kDa (experimental)

G VI.1.1.1 Cloning and single enzyme expression

The three esterases from *B. subtilis* (*BS2*),^[463] *P. fluorescens* (*Pfel*),^[464-465] and *Pyrobaculum calidifontis* VA1 (*PEST*)^[466] were cloned and kindly donated by the group of Bornscheuer. According to the plasmid maps provided, the *bs2* gene was subcloned into pET28a(+) by *NdeI/PstI* restriction, *pfel* into pGASTON utilizing the same restriction enzymes, and *pest* into pET28a(+) by *NdeI/BamHI*. *BS2* and *Pfel* bear C-terminal 6xHis tags for purification. All three esterases were successfully expressed in *E. coli* BL21(DE3) transformants harboring pET28a_ *bs2*, pGASTON_ *pfel*, and pET21a_ *pest*, respectively (**Figure D-2A–B**). Cultivation was performed in TB medium supplemented with the appropriate antibiotics following the single enzyme expression conditions in G IV. SDS-PAGE analysis was performed according to G I.8.

G VI.1.1.2 Functional screening by pNPA assay and substrate acceptance screening

CFEs were prepared as described in G I.5 and their activities toward pNPA tested (**Figure D-2C**). The assay was performed in 50 mM Tris-HCl (pH 7.5) with 0.05 mg·mL⁻¹ CFE containing esterase, 10 mM pNPA in $V_{\text{total}} = 1.0$ mL. The increase in absorption was followed at $\lambda = 405$ nm, 37°C for 30 min (Anthos Zenyth 3100 plate reader).

Substrate acceptance screenings were performed in RC duplicates with the appropriate *E. coli* BL21(DE3) transformants following the procedure in G III.2. The esters **2–4a** were tested starting at 5 mM substrate loading. Results are summarized in **Table G-24** and shown in detail in **Table D-2**.

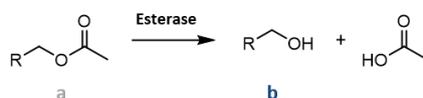
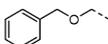


Table G-24. Substrate acceptance screening of *BS2* and *Pfel* after 24 h

	Substrate	Product	Product formation [%]	
			<i>BS2</i>	<i>Pfel</i>
	2a	2b	≥99	92
	3a	3b	≥99	≥99
	4a	4b	≥99	91

G VI.2 Characterization of ADHs for the oxidation of primary alcohols to aldehydes

G VI.2.1 Old friends letting one down: *ADH_{Lk}* and *ADH_{Rr}*

***ADH_{Lk}*:** NADPH-dependent *R*-specific ADH from *Lactobacillus kefir* (GenBank: AAP94029.1)
Gene size: 759 bp AA: 252 SDS-PAGE: 27 kDa (experimental)

***ADH_{Rr}*:** NADH-dependent *S*-specific ADH from *Rhodococcus ruber* (GenBank: CAD36475.1)
Gene size: 1 041 bp AA: 346 SDS-PAGE: 22 kDa (experimental)

G VI.2.1.1 Cloning and single enzyme expression

Both ADHs were cloned by our cooperation partners from the Bornscheuer group at the Ernst-Moritz-Arndt-Universität Greifswald, Germany.

The *lk-adh* gene from *L. kefir* was cloned into pET21b(+) utilizing the *NdeI* and the *BamHI* restriction sites. Protein production was performed according to the expression protocol from the Bornscheuer group without modifications.

Briefly, 4 mL of LB-Miller medium supplemented with Amp were inoculated with a single colony of *E. coli* BL21(DE3)/pET21b(+)-*lk-adh* and incubated at 37°C, 200 rpm overnight. TB medium supplemented with Amp was inoculated with 1% (v/v) preculture. Cells were grown at 37°C, 120 rpm until an OD₅₉₀ = 0.5 was reached. Protein production was induced with 1 mM IPTG final concentration from a 0.1 M IPTG stock (1.19 g isopropyl β-D-1-thiogalactopyranoside dissolved in 50 mL dH₂O) and expression performed at 30°C, 120 rpm for 20 h.

The *rr-adh* gene from *R. ruber* including a P_{T7} and a T7 terminator (T_{T7}) and flanking sequences was cloned into a pACYC-derived vector (in the following called pRR) utilizing unique *NdeI* and *BamHI* restriction sites. Protein production was performed according to the expression protocol from Greifswald without modifications.

Briefly, 4 mL of LB-Miller medium supplemented with Cam were inoculated with a single colony of *E. coli* BL21(DE3)/pRR-*rr-adh* and incubated at 37°C, 200 rpm overnight. TB medium supplemented with Cam was inoculated with 1/100 culture volume of the preculture. Cells were grown at 37°C, 120 rpm until an OD₅₉₀ = 0.3 was reached. ZnCl₂ was added from a 0.1 M stock (13.63 g·L⁻¹; 793523, Sigma) to 1 mM final concentration. Cells were incubated with shaking at 25°C for 30 min. Protein production was induced in the presence of 25 μM IPTG and expression performed at 25°C, 120 rpm for 22 h.

From both cultures, CFEs were prepared (G I.5) to confirm enzyme production and to determine the distribution of proteins in soluble and insoluble fractions. The total amount of protein of all samples was determined by Bradford assay according to G I.7. Gels were loaded with 10 μg of total protein per lane and SDS-PAGE performed following G I.8. Gels were stained with Coomassie Blue G250 (Bio-Rad) overnight and destained with a solution of dH₂O, EtOH, and acetic acid in a ratio of 50:40:10 (Figure G-4).

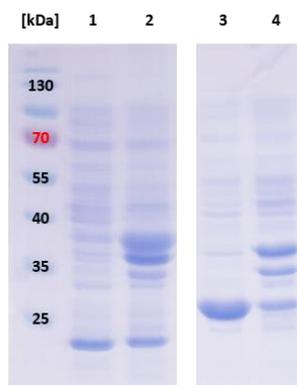


Figure G-4. Expression of *ADH_{Rr}* and *ADH_{Lk}*. *ADH_{Rr}* expressed as soluble (**1**) and insoluble (**2**) protein with the expected size of 22 kDa. *ADH_{Lk}* expressed as soluble (**3**) and insoluble (**4**) protein with the expected size of 27 kDa. Sample loading normalized to 10 µg total protein per lane.

G VI.2.1.2 Substrate acceptance screening

Biotransformations were carried out in duplicates with CFEs (5 mg·mL⁻¹) containing *ADH_{Lk}* or *ADH_{Rr}*. Substrates were added to 4 mM final concentration and a slight excess (4.25 mM) of NADP⁺ and NAD⁺ cofactor, respectively. The substrate acceptance screening was carried out according to G III.1 with compounds **1b**, **2b**, **3b**, **4b**, and **5b** (B I). The results are summarized in **Table G-25** and discussed in D II.2.1 (**Figure D-4**).

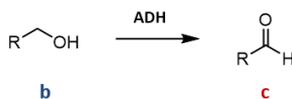


Table G-25. Substrate acceptance screening of *ADH_{Lk}* and *ADH_{Rr}* after 24 h

R-CH ₂ -OH	Substrate	Product	Product formation [%]	
			LK-ADH	RR-ADH
	1b	1c	[a]	8
	2b	2c	n.c.	[a]
	3b	3c	[b]	[b]
	4b	4c	n.c.	n.c.
	5b	5c	[b]	[a]

[a] Traces of the product detected by GC/FID. [b] <10% of the carboxylic acid overoxidation product (**d**); n.c. = no conversion

G VI.2.1.3 Cloning of the *adh_{Lk}* gene into pTYB21 and expression of N-terminally intein-tagged enzyme
The untagged *adh_{Lk}* gene was amplified from the pET21b(+) template with the following primers:

LKSapI fwd: 5'-GGTGGTTGCTCTTCCAACATGACTGACCGTTTG-3'

LKNcoI rev: 5'-GGTCCATGGCTATTGAGCAGTGTAG-3'

The preparation of the PCR mixture and the thermal cycle conditions are summarized in **Table G-26** and **Table G-27**, respectively.

Table G-26. PCR mix for *lk-adh* amplification

	<i>Pfu</i> ⁺ PCR mix	Final concentration
5.0 µL	10X <i>Pfu</i> ⁺ buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	LKSapI <i>fwd</i> (5 µM)	0.25 µM
2.5 µL	LKNcoI <i>rev</i> (5 µM)	0.25 µM
1.0 µL	pET21b(+)- <i>lk-adh</i> (50 ng·µL ⁻¹)	1 ng·µL ⁻¹
1.0 µL	DMSO	2.0% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-

Table G-27. Thermal cycle conditions for *lk-adh* amplification

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	2 min	1
Denaturation	95	30 s	30
Annealing	63.4	20 s	
Extension	72	47 s	
Terminal extension	72	2 min	1
Hold	12	∞	1

PCR yielded a single amplification product of expected size (calculated: 0.76 kb; experimental: 0.8 kb) analyzed on 1% (w/v) agarose. The PCR product was purified with the Wizard® SV Gel and PCR Clean-Up Kit (Promega). Subsequently, the *lk-adh* coding fragment was double digested in CutSmart® buffer with *SapI* (R0569, NEB) and *NcoI* (R0193, NEB) at 37°C for 1 h. The target vector pTYB21 (IMPACT™ Kit, E6901S, NEB) was double digested accordingly (**Table G-28**). The restriction enzymes were inactivated at 65°C for 20 min. Both the trimmed *lk-adh* insert and the linearized vector were purified with the Wizard® SV Gel and PCR Clean-Up Kit as before. Insert and vector were ligated by T4 ligase (Fermentas) at 22°C for 1 h (**Table G-29**). The construct was transformed into RbCl-competent *E. coli* BL21 (DE3) cells as described in G I.9.3. The plasmid DNA from five putative positive clones was isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and submitted to *PvuII* (ER0631, Thermo Scientific) control digestion at 37°C for 2.5 h in accordance to G I.11. *PvuII* digestion produced the expected DNA fragments from the putative pTYB21-*intein-CBD:adh_{Lk}* construct (4.05 kb, 3.37 kb, 0.75 kb, and 0.09 kb; **Figure G-5**). Finally, the sequence of plasmid DNA isolated from clone #3 (referred to as Int-LK3) was confirmed by Sanger sequencing (**Figure G-6**; H I.1.3.4).

Table G-28. *SapI/NcoI* double digestion

	Double digestion mix	Final concentration
2.0 µL	CutSmart® buffer	1X
2.0 / 11.0 µL	pTYB21 (0.5 µg·µL ⁻¹) / insert	50 ng·µL ⁻¹ / -
1.0 µL	<i>SapI</i> (10 U·µL ⁻¹)	0.5 U
1.0 µL	<i>NcoI</i> (10 U·µL ⁻¹)	0.5 U
14.0 / 5.0 µL	nuclease-free water	-

Table G-29. Ligation of pTYB21_intein-CBD:adh_{LK}

T4 DNA ligation mixtures		Final concentration
1.0 µL	10X T4 DNA ligase buffer	1X
1.0 µL	linearized vector DNA (30 ng·µL ⁻¹)	3.0 ng·µL ⁻¹
1.0 / 4.0 µL	trimmed insert DNA	-
0.5 µL	T4 DNA ligase (5 Weiss U·µL ⁻¹)	2.5 Weiss U
6.5 / 3.5 µL	nuclease-free water	-

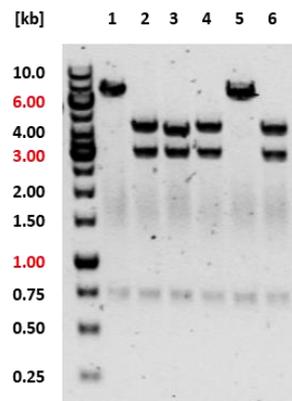


Figure G-5. PvuII control digestion of pTYB21_intein-CBD:adh_{LK} plasmids. Digestion of the empty pTYB21 vector (**1**) and putative positive clones #1–5 (**2–5**). The empty vector produced the expected DNA fragments of 6.67 kb and 0.75 kb; the 0.09 kb fragment could not be detected due to the small size (**1**). Clones #1–3 (**2–4**) and #5 (**6**) produced the desired DNA fragment pattern of 4.05, 3.37, and 0.75 kb. Again, the 0.09 kb fragment was not detected.

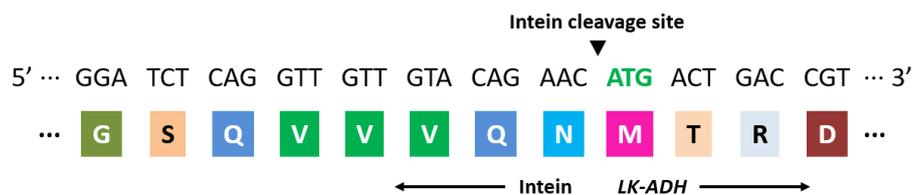


Figure G-6. Sanger sequencing of the Int-LK3 plasmid. Sanger sequencing was performed with P_{T7} and T_{T7} standard primers, T7prom *fwd* and T7term *rev* (LGC Genomics), respectively, and confirmed the in-frame fusion of the intein-CBD tag coding region to the N-terminus of the *lk-adh* gene. Cleavage of the tag (black triangle) would release untagged LK-ADH. Start codon (ATG) in green bold letters, amino acids indicated by single letter code abbreviation.

For preliminary expression studies, 4 mL LB-Miller supplemented with Amp were inoculated with a single colony of *E. coli* BL21(DE3)/pTYB21_intein-CBD:*lk-adh* (clone: Int-LK3) and incubated at 37°C, 200 rpm overnight. LB-Miller medium supplemented with Amp was inoculated with 1% (v/v) of the preculture. Cells were cultivated in the presence of 1 mM MgCl₂ at 37°C, 200 rpm until OD₅₉₀ = 0.5 was reached. Enzyme production was induced at 1 mM IPTG final concentration. Expression was performed at 20°C, 150 rpm and enzyme production followed over time (0–20 h). Whole cell samples were analyzed by SDS-PAGE as described G 1.8 (**Figure G-7**).

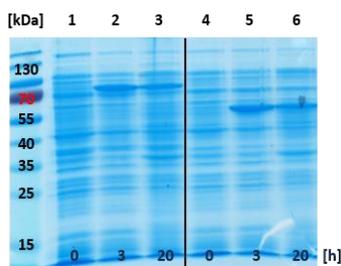


Figure G-7. Preliminary expression of *LK-ADH_{N-int}*. Production of N-terminally intein-tagged *LK-ADH* fusion protein in *E. coli* BL21(DE3) was monitored before induction (0 h; **1**) and followed over time (3–20 h; **2–3**). Production of the intein tag from pTYB21 in *E. coli* BL21(DE3) before (0 h; **4**) and after induction (3–20 h; **5–6**). The fusion protein had a size of 80 kDa (**3**), whereas the intein tag occurred at 60 kDa (calculated: 60.8 kDa; **6**). Sample loading normalized to OD₅₉₀ = 7.0.

As alternative enzyme production host, *E. coli* ER2566 (IMPACT™ Kit, E6901S, NEB) was propagated on LB agar plates from a 50% (v/v) glycerol stock (G I.3). Chemically competent cells were transformed with the Int-LK3 plasmid (G I.9.1) and preliminary expression was successfully performed (not shown) as described above.

The intein mediated purification with affinity chitin binding tag (IMPACT) kit enables the purification of *LK-ADH_{N-int}* (and other fusion proteins) *via* the chitin binding domain (CBD) from *Bacillus circularans* on a chitin resin.^[479-480, 544] In the pTYB21 vector, the *adh_{Lk}* gene is N-terminally fused to the vacuolar membrane ATPase 1 (*VMA1*) intein from *S. cerevisiae*.^[545-546] Both *LK-ADH_{N-int}* expression and purification of *ADH_{Lk}* by dithiothreitol (DTT) induced cleavage of the intein-CBD tag^[478] were optimized and performed by S. Milker.^[154]

G VI.2.2 Being thermostable is not enough: *ADH-ht*

***ADH-ht*:** Zn²⁺ and NADH-dependent ADH from *Geobacillus stearothermophilus*
(GenBank: KFL15473.1)
Gene size: 1 020 bp AA: 339 SDS-PAGE: 36 kDa (experimental)

G VI.2.2.1 Cloning, single enzyme expression, and heat shock purification

The *adh-ht* gene from *G. stearothermophilus*^[547] was synthesized by GeneArt™ Gene Synthesis with flanking *NdeI*, conveniently including the ATG start codon of the target gene, and *XhoI* restriction sites, which were utilized for subcloning into pET26b(+). The gene was delivered in a pMA-T vector conferring Amp resistance. The lyophilized plasmid DNA (5 µg) was reconstituted in 50 µL nuclease-free water and the resulting solution (100 ng·µL⁻¹) stored at -20°C for further use.

For strain and enzyme library expansion, competent *E. coli* DH5α cells were transformed with 1 ng of pMA-T/*adh-ht* as described in G I.9.1. From the resulting agar plate supplemented with Kan, single colonies were picked to prepare permanent cultures and to isolate plasmid DNA as described in G I.4 and G I.10, respectively.

Subcloning was done by Florida cloning as described in G II.1 except that the insert was amplified by KOD polymerase (71086, Merck Millipore; **Table G-30**). The PCR was performed at 65°C annealing temperature with the pMA-T_*adh-ht* template and the following primer pair (**Table G-31**):

HTNdeI *fwd*: 5'-CTCCATATGAAAGCAGCAGTTGTG-3'
HTXhoI *rev*: 5'-CACCTCGAGTTATTTATCTTCCAGGG-3'

Table G-30. Thermal cycle conditions for KOD amplifying the *adh-ht* insert

PCR step (KOD)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	2 min	1
Denaturation	95	20 s	
Annealing	65	10 s	30
Extension	70	65 s	
Terminal extension	70	3 min	1
Hold	12	∞	1

Table G-31. PCR mix for *adh-ht* amplification

	KOD DNA polymerase mix	Final concentration
10.0 µL	5X KOD reaction buffer	1X
1.0 µL	pMA-T_ <i>adh-ht</i> (100 ng·µL ⁻¹)	2 ng·µL ⁻¹
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.0 µL	HTNdel <i> fwd</i> (7.5 µM)	0.3 µM
2.0 µL	HTXhol <i> rev</i> (7.5 µM)	0.3 µM
0.5 µL	KOD DNA polymerase (5 U·µL ⁻¹)	2.5 U
32.5 µL	nuclease-free water	-

PCR yielded a single amplification product of the expected size (**Figure G-8A**). Insert trimming was performed by restriction enzyme double digestion with *NdeI* (ER058, Thermo Scientific) and *XhoI* (ER0691, Thermo Scientific) at 37°C for 2.5 h (**Table G-32**). Restriction enzymes were inactivated at 65°C for 20 min. To linearize the vector, 3 µg pET26b(+) were also digested with *NdeI* and *XhoI* (**Table G-32**). Vector purification was done by agarose gel disruption and subsequent EtOH precipitation of DNA overnight. The dried DNA pellet was dissolved in 30 µL nuclease-free water and the resulting solution (95 ng·µL⁻¹ according to NanoDrop®) used for ligation (**Table G-34**).

Table G-32. *NdeI/XhoI* double digestion of pET26b(+)

	Double digestion mix	Final concentration
4.0 µL	O buffer	1X
21 µL	pET26b(+) [144 ng·µL ⁻¹]	3 µg
1.0 µL	<i>NdeI</i> (10 U·µL ⁻¹)	5.0 U
2.0 µL	<i>XhoI</i> (10 U·µL ⁻¹)	10.0 U
12 µL	nuclease-free water	-

Table G-33. *NdeI/XhoI* double digestion of the PCR fragment containing the *adh-ht* gene

	Double digestion mix	Final concentration
2.0 µL	O buffer	1X
15 µL	insert DNA	-
0.5 µL	<i>NdeI</i> (10 U·µL ⁻¹)	5.0 U
1.0 µL	<i>XhoI</i> (10 U·µL ⁻¹)	10.0 U
1.5 µL	nuclease-free water	-

Table G-34. Ligation of pET26b(+)*_adh-ht*

T4 DNA ligation mixtures		Final concentration
2.0 μ L	10X T4 DNA ligase buffer	1X
1.0 μ L	linearized vector DNA	-
1.0 / 4.0 μ L	trimmed insert DNA	-
0.5 μ L	T4 DNA ligase (5 Weiss U· μ L ⁻¹)	2.5 Weiss U
15.5 / 12.5 μ L	nuclease-free water	-

Ligation was performed at RT for 30 min. For transformation of competent *E. coli* BL21(DE3), 5 μ L of ligation mixture were used for the transformation of 100 μ L CaCl₂ competent cells. After recovery, cells were plated on LB agar plates supplemented with Kan. The plasmid DNA was isolated from three positive clones and applied to *Pvu*II restriction enzyme control digest (**Figure G-8B**). The plasmid DNA from the HTB2 clone was successfully sequenced and propagated for enzyme expression studies.

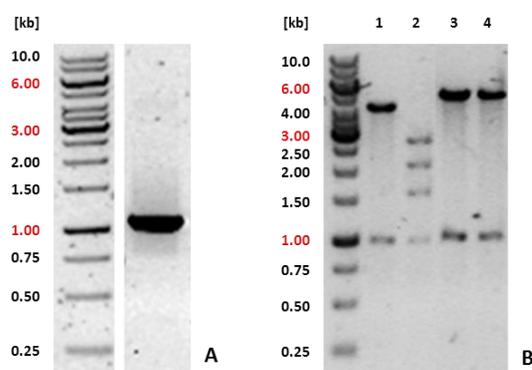


Figure G-8. Subcloning of pET26b(+)*_adh-ht*. (A) PCR amplification of the *adh-ht* insert with a calculated size of 1.03 kb. (B) *Pvu*II control digest of the empty pET26b(+) vector (1) and putative target vectors from three clones HTB1 (2), HTB2 (3), and HTB3 (4) after transformation and plasmid re-isolation; 0.85% (w/v) agarose.

For preliminary expression studies, 4 mL LB-Miller supplemented with Kan were inoculated with a single colony of *E. coli* BL21(DE3)/pET26b(+)*_adh-ht* (clone: HTB2) and incubated at 37°C, 200 rpm overnight. LB-Miller and TB medium supplemented with Kan were inoculated with 1/100 culture volume of the preculture. Cells were cultivated in the presence of 1 mM ZnCl₂ at 37°C, 200 rpm until OD₅₉₀ = 0.5 was reached and enzyme production induced by adding IPTG to 1 mM final concentration. Expression was performed at 20°C, 120 rpm and enzyme production determined over time (0–22 h) by preparation of CFEs and SDS-PAGE analysis as described in G 1.5 and G 1.8, respectively (**Figure G-9**).

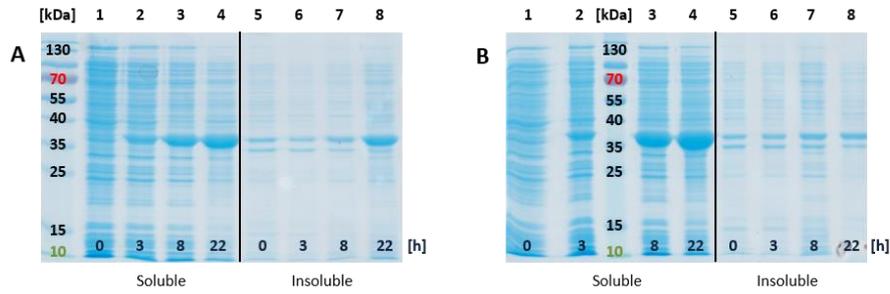


Figure G-9. Preliminary expression of *ADH-ht* in *E. coli* BL21(DE3). Production of *ADH-ht* in cells cultivated in (A) LB-Miller or (B) TB medium. Soluble and insoluble enzyme production (1–4 and 5–8, respectively) was monitored before induction (0 h) and followed over time (3–22 h). Sample loading normalized to $OD_{590} = 7.0$.

Expression was optimized following the literature protocol by Guagliardi and coworkers.^[547] Briefly, 10 mL of LB-Miller supplemented with Kan were inoculated with a single colony of HTB2 and incubated at 30°C, 120 rpm for 16 h. LB-Miller and TB medium supplemented with Kan and 1 mM $ZnCl_2$ were inoculated with 1/335 culture volume of the preculture and incubated at 37°C, 120 rpm for 2 h. Expression was performed in the presence of 1 mM IPTG at 20°C, 120 rpm for 22 h. CFEs were prepared as before and *ADH-ht* purified by heat shock (HS) as described in G I.15.1 with the following modifications: The CFEs were incubated at 60°C in a water bath (M3 Lauda with MT Lauda thermostat) for 20 min and rested on ice for 2 h. The heat shock was repeated and insolubles separated by centrifugation at 14 000 x g, 4°C for 45 min (**Figure G-10**).

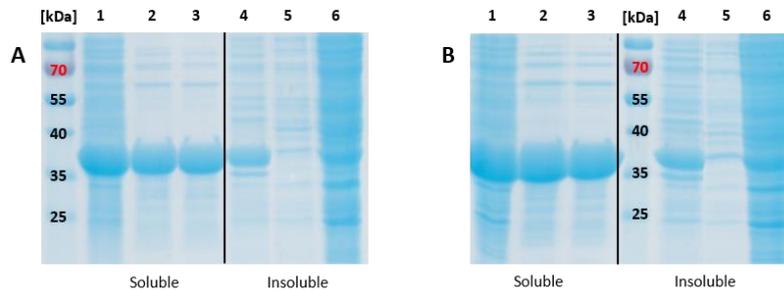


Figure G-10. Optimized expression of *ADH-ht* in *E. coli* BL21(DE3) and HS purification. Production of *ADH-ht* in cells cultivated in (A) LB-Miller or (B) TB medium. Soluble and insoluble fractions (1–3 and 4–6, respectively) before HS (1 and 4), first HS (2 and 5), and second HS (3 and 6). Sample loading normalized to $OD_{590} = 7.0$.

The purified *ADH-ht* was concentrated by a centrifugal membrane concentrator (10 kDa MWCO; UFC901024, Millipore) according to the supplier instructions. The concentration was performed once with approximately 16 mL of pooled soluble fractions ($\beta = 2 \text{ mg}\cdot\text{mL}^{-1}$). Bradford assay was performed in triplicates with 1:50 dilutions of all fractions. Centrifugal concentration yielded 2.0 mL of purified *ADH-ht* which was analyzed by SDS-PAGE (**Figure G-11**). According to Bradford assay, concentration of *ADH-ht* could be increased 27.5-fold to $55 \text{ mg}\cdot\text{mL}^{-1}$.

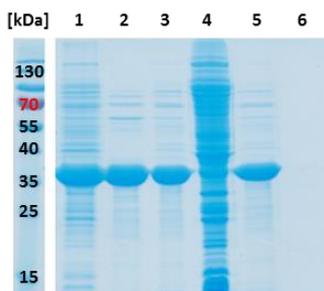


Figure G-11. HS purification and concentration of *ADH-ht*. Production of *ADH-ht* in cells cultivated in TB-Kan medium and analysis of soluble fractions before HS (**1**), after the first and the second HS (**2** and **3**, respectively), the insoluble fraction after HS purification (**4**), the concentrate containing proteins >10 kDa (**5**), and the <10 kDa flow-through (**6**). Sample loading normalized to 10 µg protein per lane.

To monitor the denaturation of *ADH-ht* at higher temperatures, the enzyme was produced in TB medium supplemented with Kan under optimized expression conditions. The resulting CFE was dispensed in 2.0 mL aliquots, which were incubated at temperatures between 50 and 95°C for 20 min. Analysis of soluble and insoluble fractions was performed by SDS-PAGE (**Figure G-12**).

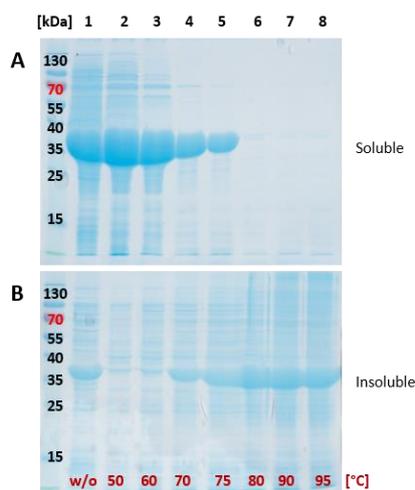


Figure G-12. Thermostability of *ADH-ht*. Determination of *ADH-ht* in (A) soluble and (B) insoluble fractions after incubation without HS (**1**) and at temperatures between 50°C and 95°C (**2–8**). Sample loading normalized to $OD_{590} = 20.0$.

G VI.2.2.2 Substrate acceptance screening

Biotransformations were carried out in duplicates under standard screening conditions employing RCs (G III.2). Substrates **1–6b** (**Table G-35**) were tested. The results are discussed in D II.2.1 (**Table D-3**). However, under standard screening conditions, only **6b** was accepted and yielding 27% of **6c**. The alcohols **1–5b** were not accepted and results are summarized in **Table D-3**.^[435]

**Table G-35.** Substrate acceptance screening of *ADH-ht* after 2 h

R-CH ₂ -OH	Substrate	Product	Product formation
			[%]
			<i>ADH-ht</i>
	1b	1c	n.c.
	2b	2c	n.c.
	3b	3c	n.c.
	4b	4c	n.c.
	5b	5c	n.c.
	6b	6c	27

n.c. = no conversion

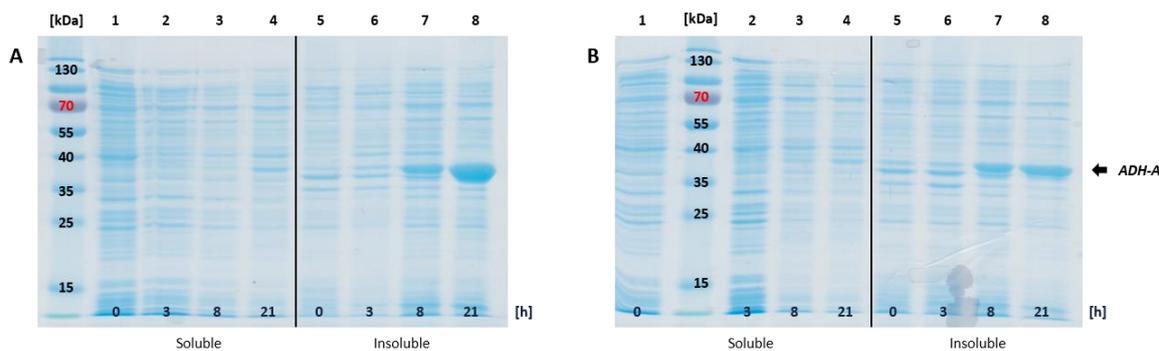
G VI.2.3 A robust biocatalyst for the oxidation of secondary alcohols: *ADH-A***ADH-A:** Zn²⁺ and NADH-dependent ADH from *Rhodococcus ruber* DSM 44541

Gene size: 1 047 bp AA: 345 SDS-PAGE: 36 kDa (experimental)

G VI.2.3.1 Cloning and optimization of *ADH-A* production

The *adh-A* gene from *R. ruber*^[548] was synthesized by GeneArt™ Gene Synthesis with flanking *NdeI*, conveniently including the ATG start codon of the target gene, and *NotI* restriction sites, which were utilized for subcloning into pET22b(+). The gene was delivered as lyophilized plasmid DNA (5 µg) and reconstituted in 50 µL nuclease-free water. The resulting solution (100 ng·µL⁻¹) was stored at -20°C for further use.

Preliminary expression studies were performed in *E. coli* BL21(DE3) cultivated in LB-Miller and TB medium at 30°C (data not shown) and 20°C (**Figure G-13**). Cells were grown at 37°C, 200 rpm until and OD₅₉₀ = 0.5 was reached. Protein production was induced with 0.5 mM IPTG and performed for 20 h. Cell lysis was done by sonication (2 s pulse/28 s break at 4°C for 5 min, 30% amplitude; KE76 probe). SDS-PAGE analysis was performed as described in G I.8.

**Figure G-13.** Preliminary expression of *ADH-A*. Soluble and insoluble production of *ADH-A* in (A) TB medium and (B) LB-Miller medium over time (0–21 h) at 20°C. Sample loading normalized to OD₅₉₀ = 7.0.

The optimized expression protocol was adapted from Edegger *et al.* as follows: TB medium supplemented with Amp and 1 mM ZnCl₂ was inoculated with a single colony of *E. coli* BL21(DE3) harboring pET22b(+)*_adh-A*. Cells were grown at 30°C, 120 rpm for 20 h. The OD₅₉₀ ≈ 12.0 was checked before Amp was added to a final concentration of 50 µg·µL⁻¹. Protein production was induced upon addition of IPTG (2 mM) and performed at 20°C, 120 rpm for 24 h.^[548] Cells were harvested and lysed as before and analyzed by SDS-PAGE. Additionally, cultivation and protein production was performed in M9-N* and LB-Miller medium with lower or similar *ADH-A* yields (**Figure G-14**).

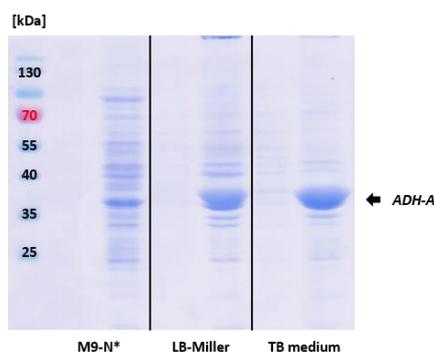


Figure G-14. Optimized expression of *ADH-A*. Soluble and insoluble production of *ADH-A* in M9-N* medium (left), LB-Miller (center), and TB medium (right). Sample loading normalized to OD₅₉₀ = 7.0.

G VI.2.3.2 Preparation of whole cell lyophilisates

As shown above, *ADH-A* exclusively accumulated in insoluble fractions after cell lysis. To prepare whole cell lyophilisates of *E. coli* BL21(DE3) expressing the target ADH, protein expression was performed as before. Cells were harvested by centrifugation (8 000 x g, 4°C for 10 min). The cell pellet was resuspended in ice-cold dH₂O, transferred to a cooled round bottom flask, snap-frozen in liquid N₂, and lyophilized at high vacuum overnight (Christ Gamma 2-20 lyophilizer). Lyophilisates were transferred into a 50 mL Greiner tube (GN227261, Sigma) and stored at 4°C until further use. The addition of 20% (w/v) sucrose did not improve the activity of the whole cell biocatalyst in subsequent biotransformations converting acetophenone to (*R*)-1-phenyl ethanol (data not shown).^[549]

G VI.2.3.3 Substrate acceptance screening

Biotransformations were carried out in duplicates under standard screening conditions employing RCs (G III.2). The same substrates as in G VI.2.2.2 (1–6b) were tested (see also **Table G-35**). Alcohols were not converted even after 24 h reaction time.^[435] The results are summarized in D II.2.1 (**Table D-3**).

The reduction of substituted acetophenones was performed by P. Schaaf in triplicates with whole cell lyophilisates (10 mg·mL⁻¹) reconstituted in 320 mM Tris-HCl (pH 8.0) in V_{total} = 0.5 mL at 30°C, 150 rpm for 1 h. Substrates were added to a final concentration of 100 mM and 2-propanol was added as cosubstrate for cofactor recycling (**Figure G-15**). Biotransformations were performed at 30°C, 200 rpm for 24 h. The reaction was stopped by adding 1 mL EtOAc containing 1 mM IS. It was extracted, the organic layer separated by centrifugation, dried over Na₂SO₄, and analyzed by GC/FID.^[549]

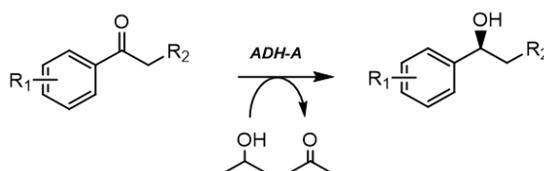


Figure G-15. (S)-selective reduction of acetophenones by ADH-A. Reaction scheme employing whole cell lyophilisates and 2-propanol as cosubstrate; $R_1 = \text{H, Br, CH}_3, \text{CF}_3, \text{Cl, F, OCH}_3$, $R_2 = \text{H, CH}_3$. Screening conditions and substrate scope according to P. Schaaf.^[549]

G VI.2.4 The winner takes it (almost) all: *AlkJ*

***AlkJ*:** FADH₂ dependent ADH from *Pseudomonas putida* (GenBank: CAB54054.1)

Gene size: 1 674 bp AA: 558 SDS-PAGE: 57 kDa (experimental)

G VI.2.4.1 FC of the *alkJ* gene and single enzyme expression

The pGEc47 cosmid containing the *alkBGHJKL* cluster from *P. putida*^[289] was kindly donated by Dr. Bruno Bühler from the TU Dortmund, Germany.

E. coli DH5 α was transformed with pGEc47 (54 597 bp) according to G I.9.1 except that 2 μL of pGEc47 (56 ng $\cdot\mu\text{L}^{-1}$) were used for transformation. After recovery, 400 μL of the cell mixture were added to 4 mL LB-Miller medium supplemented with 7.5 $\mu\text{g}\cdot\text{mL}^{-1}$ chlortetracycline (Cl-Tet) added from a Cl-Tet stock (15 mg $\cdot\text{mL}^{-1}$ in 70% (ω/v) EtOH; 26430, Sigma). Cells were grown at 37°C, 200 rpm overnight, streaked on LB agar plates supplemented with 15 $\mu\text{g}\cdot\text{mL}^{-1}$ Cl-Tet and incubated at 37°C for 24 h. Plasmid DNA isolation was performed with a single colony from the resulting plate as described in G I.10. For the elution of cosmid DNA, nuclease-free water had to be heated to 70°C prior to adding to the spin column. Cosmid DNA was quantified by NanoDrop® (139 ng $\cdot\mu\text{L}^{-1}$).^[348]

For FC, primer pairs were designed as described in G II.2.1.1 for the amplification of the *alkJ* insert and the pKA1 backbone:

ALKJ *fwd*: 5'-GAAGGAGATATACATATGTACGACTATATAATCGTTGG-3'

ALKJ *rev*: 5'-GTTAGCAGCCGGATCCTTACATGCAGACAGCTATCATG-3'

pKA1 *fwd*: 5'-GATTATATAGTCGTACATATGTATATCTCCTTCTTAAAGTTAAAC-3'

pKA1 *rev*: 5'-GATAGCTGTCTGCATGTAAGGATCCGGCTGCTAAC-3'

FC was performed following the procedure in G II.2.1. The preparation of both PCR reactions and the thermal cycle conditions are summarized in **Table G-36** and **Table G-37**, respectively.

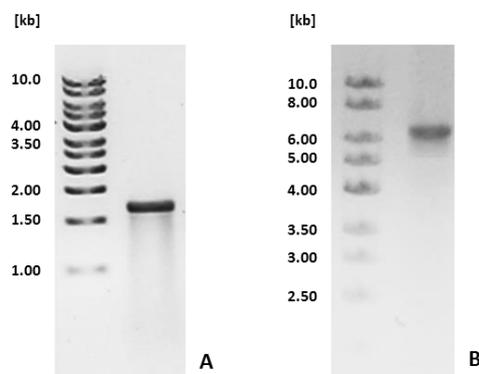
Table G-36. PCR reactions for FC of pKA1_alkJ

<i>Pfu</i> ⁺ PCR mix			Final concentration	<i>OptiTaq</i> PCR mix			Final concentration
5.0 µL	10X <i>Pfu</i> buffer		1X	5.0 µL	10X Pol B buffer		1X
2.0 µL	dNTP mix (5 mM each)		0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)		0.2 mM each dNTP
2.5 µL	ALKJ <i>fwd</i> (5 µM)		0.25 µM	2.5 µL	pKA1 <i>fwd</i> (5 µM)		0.25 µM
2.5 µL	ALKJ <i>rev</i> (5 µM)		0.25 µM	2.5 µL	pKA1 <i>rev</i> (5 µM)		0.25 µM
0.5 µL	pGEc47 (139 ng·µL ⁻¹)		1.39 ng·µL ⁻¹	1.0 µL	pRR (35 ng·µL ⁻¹)		0.7 ng·µL ⁻¹
1.0 µL	DMSO		2% (v/v)	1.0 µL	DMSO		2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)		2.5 U	0.5 µL	<i>OptiTaq</i> polymerase (5 U·µL ⁻¹)		2.5 U
36.0 µL	nuclease-free water		-	35.5 µL	nuclease-free water		-

Table G-37. *Pfu*⁺ and *OptiTaq* thermal cycle conditions for FC of pKA1_alkJ

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (<i>OptiTaq</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	45 s		Denaturation	95	40 s	
Annealing	48	30 s	30	Annealing	52	30 s	30
Extension	72	110 s		Extension	72	6 min	
Terminal extension	72	5 min	1	Terminal extension	72	10 min	1
Hold	4	∞	1	Hold	4	∞	1

After *DpnI* digestion and purification (Figure G-16), the 1:1 and 4:1 mixtures were prepared and *E. coli* TOP10 cells transformed with the linear DNA fragment mixtures as outlined in G II.2.1.3. Clones were picked and colony PCR performed with *Taq* polymerase according to G I.14 with the thermal cycle conditions for the *alkJ* gene from **Table G-37** to identify positives that were finally confirmed by Sanger sequencing.^[348]


Figure G-16. Linear DNA fragments for pKA1_alkJ assembly. Purified PCR products of (A) the *alkJ* fragment and (B) the pKA1 backbone analyzed on 1.0% and 0.8% (w/v) agarose, respectively.

Competent *E. coli* BL21(DE3) cells were transformed with the plasmid DNA from clone AJ3, which contained the complete *alkJ* gene according to sequencing. The optimized production of *AlkJ* was performed as follows: A single colony was cultivated in 10 mL LB-Miller medium supplemented with Cam ($34 \mu\text{g}\cdot\text{mL}^{-1}$) with shaking (120 rpm) at 30°C overnight. For enzyme production, TB medium supplemented with Cam ($34 \mu\text{g}\cdot\text{mL}^{-1}$) was inoculated with 1% (v/v) of the pre-culture and shaken at 37°C , 120 rpm for 6 h. Enzyme production was performed in the presence of 0.5 mM IPTG at 25°C , 120 rpm for 20–24 h. Since *AlkJ* is a membrane-associated enzyme, it was exclusively found in insoluble fractions of protein preparations according to SDS-PAGE analysis (Figure D-5B).^[348]

G VI.2.4.2 Substrate scope and substrate profile expansion toward nonnative compound classes

The ADH from *P. putida* was previously reported to oxidize primary aliphatic alcohols selectively to the corresponding aldehydes.^[391] In this work, *AlkJ* was identified to efficiently oxidize primary aromatic alcohols as well. Substrate acceptance screenings were performed in RCs under standard conditions (G III.2). From the substrate palette (Table G-38), especially the aromatic substrates **2b** and **4–6b** bearing aliphatic C2 and longer side chains were fully converted. Contradictory, **3b** was poorly oxidized to **3c**. Alcohols with short side chains like **1b** or sterically demanding substrates such as **7b** and **9b** were only moderately converted to the corresponding aldehydes. Furthermore, overoxidation of **2c** and **4–6c** to the corresponding carboxylic acids was observed in *E. coli* BL21(DE3) RCs expressing *AlkJ*.^[348] Results are shown in detail in Figure D-6.

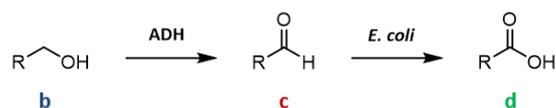


Table G-38. Substrate acceptance screening of *AlkJ* after 24 h

R-CH ₂ -OH	Substrate	Main product	Product formation [%]
			<i>AlkJ</i>
	1b	1c	46
	2b	2d	≥99
	3b	3c	[a]
	4b	4d	≥99
	5b	5d	≥99
	6b	6c	36 [b]
	7b	7c	43
	9b	9c	15

[a] Traces of the aldehyde product. [a] Recoveries <50% after 24 h.

G VI.2.4.3 Truncation of the *alkJ* gene by utilization of encoded *PstI* restriction sites

AlkJ_{trnc}: Truncated variant of *AlkJ* from *Pseudomonas putida*

Gene size: 903 bp AA: 300 SDS-PAGE: 32 kDa (experimental)

For the truncation of the *alkJ* gene, pKA1_alkJ was reisolated from *E. coli* DH5α transformants according to G I.10 and the vector (3 μg) administered to digestion by *PstI* (R0140S, NEB) at 37°C for 2 h (Table G-39). *PstI* was heat-

inactivated at 80°C for 20 min. The digestion mixture was separated on 1% (ω/v) agarose and the major band (6.43 kb) excised and further treated according to Florida cloning as described in G II.1.3.

Table G-39. *PstI* digestion mix

<i>PstI</i> digestion mix		Final concentration
4.0 μ L	NEBuffer 3.1	1X
32.0 μ L	pKA1_alkJ (88 ng· μ L ⁻¹)	\approx 3 μ g
4.0 μ L	<i>PstI</i> (10 U· μ L ⁻¹)	1 U· μ L ⁻¹

The linear fragment was ligated as in **Table G-56** at 25°C for 20 min. After transformation of RbCl-competent *E. coli* BL21(DE3) (G I.9.3), cells were plated on LB agar plates supplemented with Cam. Reisolated plasmid DNA of single colonies was sent for sequencing and the successful in-frame truncation confirmed (H I.1.10.4). Subsequent expression of *AlkJ_{trnc}* was performed in AIM as described in G VI.3.1.1. SDS-PAGE analysis confirmed the production of *AlkJ* and *Fsa1-A129S* (**Figure E-21A**).

RCs of *E. coli* BL21(DE3)/pKA1_alkJ_{trnc} were prepared and the disruption of *AlkJ* activity confirmed with **2b** and **4b** under standard screening conditions (G III.2; **Figure E-3B**).

G VI.3 Characterization of CARs for the reduction of carboxylic acids to aldehydes

G VI.3.1 *CAR_{Ni}* and *CAR_{Mm}*

CAR_{Ni}: ATP/NADPH dependent carboxylic acid reductase from *Nocardia iowensis*
(EMBL: AAR91681.1)
Gene size: 3 525 bp AA: 1 174 SDS-PAGE: 120 kDa (experimental)

CAR_{Mm}: ATP/NADPH dependent carboxylic acid reductase from *Mycobacterium marinum*
(EMBL: ACC40567.1)
Gene size: 3 525 bp AA: 1 174 SDS-PAGE: 120 kDa (experimental)

G VI.3.1.1 Cloning and enzyme expression by autoinduction

Both the pETDuet-1_*PPTase_{Ec}::car_{Ni}* and the pETDuet-1_*PPTase_{Ec}::car_{Mm}* plasmid were constructed by Dr. Margit Winkler from ACIB Graz, Austria. In the following, the cloning of pETDuet-1_*PPTase_{Ec}::car_{Ni}* will be briefly described.^[348]

The fragment corresponding to the *PPTase_{Ec}* gene was amplified from pJexpress401_*PPTase_{Ec}* with following primers using Phusion polymerase (Finnzymes):

EcPPTase_NcoI fwd: 5'-AATCACCATGGTCGATATGAAAACACTACGCATACCTC-3'
EcPPTase_HindIII rev: 5'-AATCAAAGCTTAATCGTGTGGCACAGCGTTATG-3'

The fragment corresponding to the *car_{Ni}* gene was amplified from pEHISTTEV_*car_{Ni}* (originating from pJexpress404_*car_{Ni}*) with following primers:

pMS470d8_HIS-TEVADH2 fwd: 5'-ATACATATGTCGTACTACCATCACCATCACC-3'
CarNoc_XhoI rev: 5'-AATCACTCGAGTTACAGCAGTTGCAGCAG-3'

The optimized thermal cycle conditions for both PCRs are summarized in **Table G-40**. The *PPTase_{Ec}* PCR product as well as pETDuet-1 (Novagen) were digested with fast digest *NcoI* and *HindIII* (Fermentas), gel-separated and

the excised DNA fragments purified with the Wizard® SV Gel and PCR Clean-Up Kit (Promega). Subsequently, *Pptase_{Ec}* was ligated with pETDuet-1, catalyzed by T4 ligase (Fermentas) at 22°C for 1 h. The construct was transformed into electrocompetent *E. coli* TOP10 cells, colonies selected on LB-Amp plates and the sequence confirmed by Sanger sequencing (Microsynth). Plasmid preparation with the correct *Pptase_{Ec}* insert as well as the PCR product of the *car_{Ni}* were subsequently digested with *NdeI* and *XhoI* fast digest restriction enzymes (Fermentas) and ligated as described above. The construct was transformed into electrocompetent *E. coli* TOP10 cells, colonies selected on LB-Amp agar plates. Again, Sanger sequencing confirmed the sequence. Finally, the plasmid preparation with both inserts was transformed into *E. coli* BL21-Gold(DE3).^[348]

Table G-40. Optimized thermal cycle conditions for the cloning of pETDuet-1_*Pptase_{Ec}::car_{Ni}*

PCR step (<i>Pptase_{Ec}</i>)	Temperature [°C]	Time	No. of cycles	PCR step (<i>car_{Ni}</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	98	30 s	1	Initial denaturation	98	30 s	1
Denaturation	98	10 s		Denaturation	98	7 s	
Annealing	57	10 s	25	Annealing	62	20 s	30
Extension	72	115 s		Extension	72	70 s	
Terminal extension	72	7 min	1	Terminal extension	72	7 min	1
Hold	4	∞	1	Hold	4	∞	1

The pETDuet-1_*Pptase_{Ec}::car_{Mm}* construct was cloned accordingly.^[550] Competent *E. coli* BL21-Star(DE3) cells were transformed with the target plasmid containing both inserts.

Production of both CARs was performed according to the expression protocol provided by Dr. Margit Winkler with minor modifications.^[543] Briefly, 12 mL LB-0.8G medium supplemented with Amp (100 µg·mL⁻¹) were inoculated with a single colony of the desired transformant and shaken at 37°C, 275 rpm (InforsHT Multitron 2 Standard) overnight. AIM was used for protein production. Therefore, LB-5052 was inoculated with 0.2% (v/v) preculture and incubated at 37°C, 150 rpm for 4 h. The temperature was switched to 20°C and enzyme expression performed for 20 h. Cells were harvested by centrifugation (8 000 x g, 4°C for 10 min).

For subsequent enzyme characterization, *CAR_{Ni}* and *CAR_{Mm}* were tested with the substrates **1–2d** and **7d** under standard screening conditions employing RCs (**Figure D-11**). Whereas *CAR_{Ni}* readily converted all three carboxylates to the target aldehydes, *CAR_{Mm}* only accepted **1d** to produce **1c**. Additionally, *CAR_{Ni}* was purified by IMAC as described below and the carboxylates **1–2d** and **8d** attempted to be reduced *in vitro*.

G VI.3.1.2 Purification of *CAR_{Ni}* by IMAC

After expression, cells were harvested by centrifugation (8 000 x g, 4°C for 10 min), cells resuspended in 25 mM Tris-HCl (pH 8.0) containing 0.3 M NaCl, and lysed by sonication as described in G I.5. The insoluble fraction was removed by centrifugation (16 000 x g, 4°C for 25 min). Purification was performed as described in G VI.7.2 with modifications according to Finnigan and coworkers.^[64] The binding buffer constituted of 25 mM Tris-HCl (pH 8.0), 40 mM imidazole, and 0.5 M NaCl. The elution buffer contained 25 mM Tris-HCl (pH 7.5), 0.25 M imidazole, and 0.1 M NaCl.

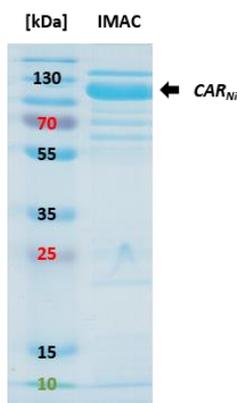


Figure G-17. Purification of CAR_{Ni} . Concentrate of purified enzyme after IMAC. Other protein bands might correspond to apo- CAR_{Ni} . Protein production as in G VI.3.1.1. Sample loading normalized to 10 μ g total amount of protein per lane.

G VI.3.1.3 Substrate acceptance screening of CAR_{Ni}

RCs coexpressing CAR_{Ni} and $Pptase_{Ec}$ were provided with the carboxylic acids **1d–7d** and **9d**. All tested carboxylates were efficiently reduced to the corresponding aldehydes under standard screening conditions outlined in G III.2 (**Table G-41**). Aldehydes were subsequently reduced to the primary alcohols **1b–7b** and **9b** by the enzymatic host background in response to the oxidative and electrophilic stress induced by the reactive carbonyl group in aldehydes.^[42, 280-281, 348, 400, 404] Results are summarized in **Figure D-12**.

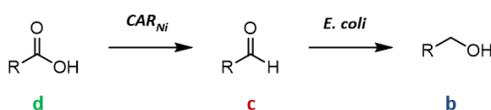


Table G-41. Substrate acceptance screening of CAR_{Ni} after 24 h

R-CH ₂ -OH	Substrate	Main product	Product formation
			[%]
			CAR_{Ni}
	1d	1b	71 ^[a]
	2d	2b	92 ^[a]
	3d	3b	95
	4d	4b	95
	5d	5b	89
	6d	n.a.	n.c.
	7d	7b	88
	8d	n.a.	n.c. ^[b]
	9d	9b	46 ^[c]

^[a] *In vitro* reduction to the aldehyde by purified CAR_{Ni} and GDH_{2xBS} for cofactor recycling (see D II.3). ^[b] Tested *in vitro* only. ^[c] 41% **9c** after 24 h; n.a. = not applicable, n.c. = no conversion. In the reaction scheme $Pptase_{Ec}$ to produce holo- CAR_{Ni} is omitted for clarity.

In vitro reductions of the carboxylic acids **1-2d** and **8d** were performed in 100 mM Tris-HCl (pH 7.5) with purified *CAR_{Ni}* (50 µg), 2.5 mM substrate loading, stoichiometric amounts of ATP, and 25 mM MgCl₂. For NADPH cofactor regeneration, *GDH_{2xBs}* (0.05 mg·mL⁻¹) and 100 mM D-glucose were used (**Figure D-13**).^[332] *GDH_{2xBs}* expression and purification are described in G VI.7.2.2. The reduction was started by the addition of 0.1 mM NADP⁺ and followed by GC/FID as outlined in G III.1. Screenings were performed at 30°C.

G VI.4 Characterization of different aldolases for C–C bond formation

G VI.4.1 DHAP-dependent aldolases: *FruA*, *RhuA*, and *FucA*

FruA: Fructose-1,6-bisphosphate aldolase from *Staphylococcus carnosus* (GenBank: X71729.1)
Gene size: 891 bp AA: 296 SDS-PAGE: 37 kDa (experimental)

RhuA: Rhamnulose 1-phosphate aldolase from *Escherichia coli*
Gene size: 825 bp AA: 274 SDS-PAGE: 35 kDa (experimental)

FucA: Fuculose 1-phosphate aldolase from *Escherichia coli*
Gene size: 648 bp AA: 215 SDS-PAGE: 24 kDa (experimental)

G VI.4.1.1 Cloning and expression of *FruA*, *RhuA*, and *FucA*

The pKK223-3 (pKK) vectors containing the *fruA*, the *rhuA*, and the *fucA* gene were kindly donated by Prof. Wolf-Dieter Fessner and Dr. Michael Kickstein from the TU Darmstadt, Germany.

According to Sanger sequencing, all genes were cloned by *EcoRI/PstI* or *BamHI* restriction sites and cloned into the target vector downstream of a *P_{tac}*. Subsequently, competent *E. coli* BL21(DE3) were transformed with the desired vector. Expression was performed as suggested by M. Kickstein. Briefly, 4 mL LB-Miller supplemented with Amp were inoculated with a single colony of target strain and grown at 37°C, 200 rpm overnight. TB-Amp was inoculated with 1% (v/v) preculture and cultivated at 37°C, 200 rpm until OD₅₉₀ = 0.3 was reached. ZnCl₂ to a final concentration of 1 mM was added and the temperature shifted to 25°C. Protein production was induced in the presence of 0.1 mM IPTG and performed at 25°C, 150 rpm for 20–22 h. An example SDS-PAGE gel is shown in **Figure D-17**.

G VI.4.1.2 Construction of pKA1_*fucA* by FC

The pKK_*fucA* was kindly donated by Prof. Wolf-Dieter Fessner and Dr. Michael Kickstein from the TU Darmstadt, Germany. In course of her B.Sc. thesis, B. Walder performed FC of pKA1_*fucA*.^[509]

Briefly, a single colony of *E. coli* DH5α transformed with pKK_*fucA* according to G I.9.1 was cultivated in 5 mL LB-Miller medium supplemented with Amp at 37°C, 200 rpm overnight. Plasmid DNA isolation was performed as described in G I.10 and quantified by NanoDrop® (80 ng·µL⁻¹).

For FC, primer pairs were designed as described in G II.2.1.1 for the amplification of the *fucA* insert and the pKA1 backbone:

FucA fwd: 5'-GGAGATATACATATGATGGAACGAAATAAACTTGCTC-3'
FucA rev: 5'-GCAGCCGGATCCTTACTTCAATTCGTAACCC-3'

pKA1FucA fwd: 5'-CAAGTTTATTCGTTCCATCATATGTATATCTCCTTC-3'
pKA1FucA rev: 5'-GTTACGAATTGAAGAGTAAGTAAGGATCCGGCTGCTAAC-3'

Gradient PCRs resulted in optimal annealing temperatures of 61.0°C and 48.5°C for the *fucA* coding fragment and the pKA1 backbone, respectively (data not shown).^[509] FC was performed following the procedure in G II.2.1.

The preparation of both PCR reactions and the thermal cycle conditions are summarized in **Table G-42** and **Table G-43**, respectively.

Table G-42. PCR reactions for FC of pKA1_ *fucA*

<i>Pfu</i> ⁺ PCR mix		Final concentration	Opti <i>Taq</i> PCR mix		Final concentration
5.0 µL	10X <i>Pfu</i> buffer	1X	5.0 µL	10X Pol B buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	<i>FucA fwd</i> (5 µM)	0.25 µM	2.5 µL	pKA1 <i>FucA fwd</i> (5 µM)	0.25 µM
2.5 µL	<i>FucA rev</i> (5 µM)	0.25 µM	2.5 µL	pKA1 <i>FucA rev</i> (5 µM)	0.25 µM
1.0 µL	pKK_ <i>fucA</i> (80 ng·µL ⁻¹)	0.8 ng·µL ⁻¹	1.0 µL	pKA1_ <i>alkJ</i> (109 ng·µL ⁻¹)	1.09 ng·µL ⁻¹
1.0 µL	DMSO	2% (v/v)	1.0 µL	DMSO	2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U	0.5 µL	Opti <i>Taq</i> polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-	35.5 µL	nuclease-free water	-

Table G-43. *Pfu*⁺ and Opti*Taq* thermal cycle conditions for FC of pKA1_ *fucA*

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (Opti <i>Taq</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	61	30 s	30	Annealing	48.5	30 s	30
Extension	72	40 s		Extension	72	6 min	
Terminal extension	72	3 min	1	Terminal extension	72	10 min	1
Hold	12	∞	1	Hold	12	∞	1

After *DpnI* digestion and purification (**Figure G-18A**), the 1:1, 4:1, and an additional 8:1 mixture was prepared and *E. coli* TOP10 cells transformed with the linear DNA fragment mixtures as outlined in G II.2.1.3. Four clones (BWA1–4) were picked, plasmid DNA isolated, and administered to *NcoI* control digestion according to G I.11. All four plasmids showed the expected DNA fragment pattern post restriction enzyme digestion (**Figure G-18B**). The *fucA* sequence in BWA2 and BWA4 plasmids was finally confirmed by Sanger sequencing.

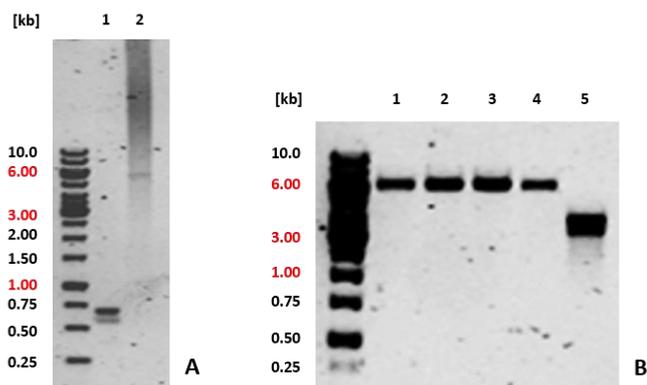


Figure G-18. Linear DNA fragments for pKA1_ *fucA* assembly. (A) Purified PCR products of the *fucA* fragment (1) the pKA1 backbone (2). (B) *NcoI* control digestion of newly assembled pKA1_ *fucA* producing a single 6 kb fragment (BWA1-4: 1-4) and the pKA1_ *alkJ* parent plasmid producing a 3.82 kb and a 3.39 kb fragment as expected (5); 1.0% (w/v) agarose. The figure was adapted from B. Walder (2016).^[509]

Competent *E. coli* BL21(DE3) cells were transformed with the plasmid DNA from clone BWA4. The production of *FucA* was performed as before (G VI.4.1.1) except that pKA1_ *fucA* transformants were cultivated in the presence of Cam (Figure G-19). Protein production was performed as described above (G VI.4.1.1).



Figure G-19. Expression of *FucA* from pKA1_ *fucA*. SDS-PAGE analysis of whole cell samples producing *FucA* before (1) and after induction with IPTG (2). Sample loading normalized to OD₅₉₀ = 7.0. The figure was adapted from B. Walder (2016).^[509]

G VI.4.1.3 Assaying *FucA* activity *in vitro*

Initial studies *in vitro* were performed by T. Wiesinger to access (3*R*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e_{FucA}**) and the relative configuration confirmed by ¹³C- and ³¹P-NMR by T. Wiesinger as well (Figure D-16).^[487]

Biotransformations were performed in duplicates either in a one pot fashion by simultaneously adding CFEs containing *FucA* and the phosphatase from *Shigella flexneri* (*PhoN-Sf*) or by adding the aldolase and *PhoN-Sf* sequentially (Figure D-18). Single enzyme expressions are described in this chapter and summarized in Table G-23. The preparation of CFEs is outlined in G I.5. *In vitro* reactions were carried out at 25°C, 250 rpm overnight with CFEs containing *FucA* and *PhoN-Sf* (5 mg·mL⁻¹ each), 1.7 eq **2c**, and 5% (v/v) ACN as cosolvent. The addition of 1.0 eq DHAP started the reaction. The consumption of **2c** was followed by calibrated GC/FID, whereas the formation of the target aldol **2e_{FucA}** was monitored by calibrated HPLC (Figure D-18).^[487] Proteins were precipitated prior to HPLC measurement as described in G III.2.

G VI.4.2 The dihydroxyacetone (DHA)-utilizing aldolases *Fsa1* and *Fsa1-A129S*

***Fsa1*:** Fructose-6-phosphate aldolase from *Escherichia coli* (GenBank: AMC98045.1)
 Gene size: 663 bp AA: 220 SDS-PAGE: 23 kDa (experimental)

G VI.4.2.1 Cloning and enzyme expression

The pET16b vectors containing the wild type *fsa1* and the mutant *fsa1-A129S* gene^[432] were kindly donated by Prof. Wolf-Dieter Fessner from the TU Darmstadt, Germany.

According to Sanger sequencing, the desired genes were subcloned into pET16b by *NcoI/XhoI*. For protein production, competent *E. coli* BL21(DE3) cells were transformed with the aldolase-coding vectors as described in G I.9.1. Expression was performed as follows and monitored over time: TB medium supplemented with Amp (100 µg·mL⁻¹) was inoculated with 1% (v/v) preculture prepared as described in G I.5 or G I.6. Cells were grown at 37°C, 200 rpm until OD₅₉₀ = 0.5. Protein production was induced upon the addition of 0.5 mM IPTG and performed at 30°C, 200 rpm for 20–22 h.^[348, 435] Cells were harvested by centrifugation (8 000 x g, 4°C for 10 min) and lysis was performed either in 50 mM Tris-HCl (pH 7.5) by sonication as described in G I.5 or in 50 mM GlyGly buffer (pH 8.0) in the presence of lysozyme at 37°C (350 rpm) for 1 h. The detailed lysozyme protocol is outlined in G I.15.1 (**Figure G-20**). Since *Fsa1* and *Fsa1-A129S* are thermostable, HS purification could be applied (G VI.4.2.2). The commonly applied HS protocol is given in G I.15.1. The protein amount of all samples was determined by Bradford assay (G I.7) before SDS-PAGE analysis (G I.8).

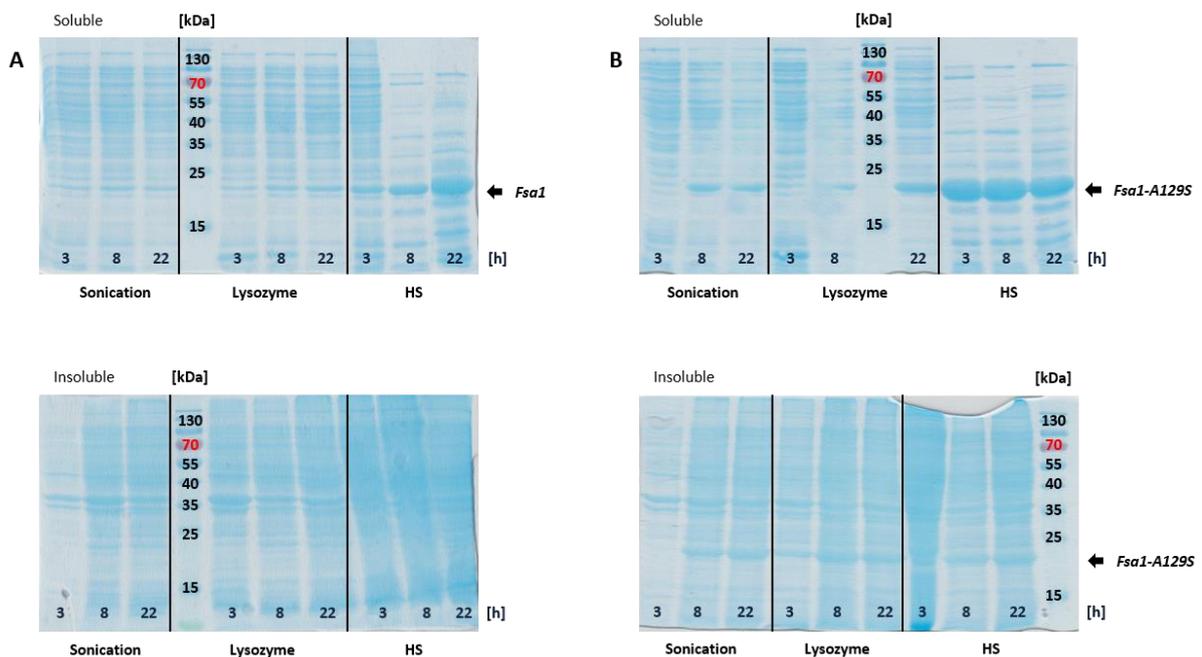


Figure G-20. Expression of *Fsa1* and *Fsa1-A129S*. SDS-PAGE analysis of (A) wild type *Fsa1* and (B) engineered *Fsa1-A129S* in *E. coli* BL21(DE3) from pET16b. Protein production was monitored over time (0–22 h) and cell lysis performed by sonication or lysozyme as indicated. The aldolases were semi-purified by HS after lysis by lysozyme. Soluble and insoluble fractions on top and bottom, respectively. Sample loading normalized to 10 µg total amount of protein per lane.

Transformation of *E. coli* JM109(DE3), followed by enzyme expression as described above, did not yield higher amounts of soluble target *Fsa1* and *Fsa1-A129S* (data not shown).^[486]

G VI.4.2.2 Purification of *Fsa1-A129S* by HS and preparation of lyophilisates

Semi-purification of *Fsa1* and *Fsa1-A129S* was performed after cell lysis in the presence of lysozyme by HS at 70°C for 0.5 h (**Figure D-15B** and **Figure G-20**). Lyophilisates were prepared according to G I.15.1.

G VI.4.2.3 Assaying *Fsa1-A129S* activity *in vitro*

Aldolase activity was tested *in vitro* by the aldol addition of the model aldehyde **2c** and DHA (**Figure G-21** and **Figure D-16**).

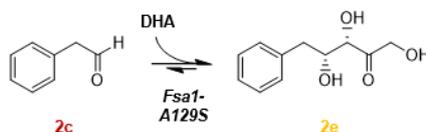


Figure G-21. *In vitro* functionality of *Fsa1-A129S*. Purified *Fsa1-A129S* (or *Fsa1*) to produce target aldol **2e** from **2c** and DHA. Full conversion to the product after 24 h (see D II.4.1).

Screening mixtures contained lyophilisates ($10 \text{ mg} \cdot \text{mL}^{-1}$), dissolved in 50 mM Tris-HCl (pH 7.5), 1 eq **2c**, and 20 eq DHA. The DHA dimer was monomerized beforehand as described in G III.2. The aldehyde substrate was added last and samples taken immediately after mixing (to^{*}), 2 h, and 24 h, also described in G III.2. The consumption of **2c** was followed by calibrated GC/FID, whereas the formation of the target aldol (3*S*,4*R*)-1,3,4-trihydroxy-5-phenylpentan-2-one (**2e**) was monitored by calibrated HPLC (**Figure D-16**). Before HPLC analysis, proteins were precipitated as described in G III.2.^[348, 435]

G VI.5 Characterization of acidic phosphatases and engineered variants

PhoN-Se: Acidic phosphatase from *Salmonella enterica typhimurium* (GenBank: X59036)

Gene size: 699 bp AA: 232 SDS-PAGE: 25 kDa (experimental)

PhoN-Sf: Acidic phosphatase from *Shigella flexneri* (GenBank: BAA11655.1)

Gene size: 744 bp AA: 247 SDS-PAGE: 25 kDa (experimental)

YqaB: Acidic phosphatase from *Escherichia coli* (EcoGene: EG13530)

Gene size: 570 bp AA: 189 SDS-PAGE: 20 kDa (experimental)

G VI.5.1 The mutant acidic phosphatase *PhoN-Se V78L*

PhoN-Se V78L_{fl}: Phosphatase mutant from *S. typhimurium* with SP and 3xFLAG tag

Gene size: 768 bp AA: 255 SDS-PAGE: 26 kDa (experimental)

PhoN-Se V78L_{SP}: Engineered phosphatase from *S. typhimurium* with N-terminal SP

Gene size: 699 bp AA: 232 SDS-PAGE: 25 kDa (experimental)

PhoN-Se V78L_{3xFLAG}: Phosphatase mutant from *S. typhimurium* with C-terminal 3xFLAG tag

Gene size: 723 bp AA: 240 SDS-PAGE: 20 kDa (experimental)

G VI.5.1.1 Cloning and expression of *PhoN-Se V78L* and variants

The gene coding for the engineered phosphatase gene from *S. typhimurium*^[484] including the N-terminal SP and the C-terminal 3xFLAG tag^[494-495] was ordered from GeneArt™ in a pMA-T vector. The lyophilized plasmid DNA (5 μg) was reconstituted in 50 μL nuclease-free dH₂O and the resulting solution ($100 \text{ ng} \cdot \mu\text{L}^{-1}$) stored at -20°C for further use. Subcloning into pET26b(+) utilized the *NdeI/XhoI* restriction sites, which were introduced by the designed primers. Cloning was performed according to the improved Florida cloning procedure except that KOD polymerase (71086, Merck Millipore) was used for target insert amplification (G II.1).

PhoN-Se V78L_{fl} was amplified with the following primer pair:

PhoN-Se_fl_NdeI fwd: 5'-CGGCATATGAAAAGTCGTTATTTAGTATTTTTTC-3'

PhoN-Se_fl_XhoI rev: 5'-GTTTCTCGAGTCACTTATCATCGTCATCCTTG-3'

PhoN-Se V78L_{SP} was amplified with the PhoN-Se_fl_NdeI fwd primer and the following rev primer:

PhoN-Se_w/o tag_XhoI rev: 5'-GGTTCTCGAGTCAGCTCCTCACGG-3'

PhoN-Se V78L_{3xFLAG} was amplified with the PhoN-Se_fl_XhoI rev primer and the following fwd primer:

PhoN-Se_w/o SP_NdeI fwd: 5'-CTGCATATGAAATATACATCAGCAGAAACAGTGC-3'

The following PCR mixture was prepared for each insert amplification:

Table G-44. PCR mixtures for *phoN-Se V78L* insert variants

KOD DNA polymerase mix		Final concentration
10.0 µL	5X KOD reaction buffer	1X
1.0 µL	pMA-T_ <i>phoN-Se_{fl}</i> (100 ng·µL ⁻¹)	2 ng·µL ⁻¹
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.0 µL	primer fwd (7.5 µM)	0.3 µM
2.0 µL	primer rev (7.5 µM)	0.3 µM
0.5 µL	KOD DNA polymerase (5 U·µL ⁻¹)	2.5 U
32.5 µL	nuclease-free water	-

The thermal cycle conditions were according to the amplification of the *adh-ht* insert (**Table G-30**) except varying extension times for the different insert variants:

- 47 s for *phoN-Se V78L_{fl}*
- 42 s for *phoN-Se V78L_{SP}*
- 44 s for *phoN-Se V78L_{3xFLAG}*

PCRs yielded single amplification products of the expected sizes for all inserts (**Figure G-22**). Insert trimming was performed by restriction enzyme double digestion with *NdeI* (ER058, Thermo Scientific) and *XhoI* (ER0691, Thermo Scientific) at 37°C for 2.5 h according to the *adh-ht* gene (**Table G-33**). Restriction enzymes were inactivated at 65°C for 20 min. The pET26b(+) vector was digested with the same enzymes as in **Table G-32**. Vector purification was done by agarose gel disruption and subsequent EtOH precipitation of DNA (G II.1.3 and G I.12, respectively). The dried DNA pellet was dissolved in 30 µL nuclease-free dH₂O and the resulting solution used for ligation as in **Table G-34**. Ligation was performed at RT for 20 min. Transformation of competent *E. coli* BL21(DE3) was performed with 5 µL of the corresponding ligation mixture (G I.9.1). After recovery, cells were plated on LB agar plates supplemented with Kan. Plasmid DNA was isolated from putative positive clones and sent for sequencing (H I.1.6).

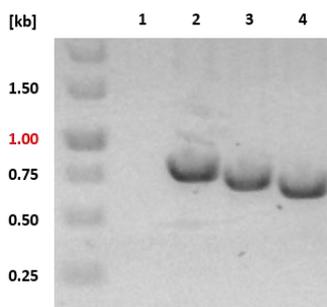


Figure G-22. Insert amplification of *phoN-Se V78L* insert variants. Purified PCR products containing *phoN-Se V78L_{fl}* (**2**), *phoN-Se V78L_{3xFLAG}* (**3**), and *phoN-Se V78L_{SP}* (**4**). The calculated sizes of each fragment are 768 bp, 723 bp, and 699 bp, respectively. The negative control (no DNA template) yielded no PCR product (**1**); 1.0% (w/v) agarose.

The plasmids encoding *PhoN-Se V78L_{fl}* (clone: SeFL1), *PhoN-Se V78L_{SP}* (clone: SeSP3), and *PhoN-Se V78L_{3xFLAG}* (clone: SeTAG3), respectively, were successfully sequenced and propagated for enzyme expression studies and cellular localization of phosphatase variants.

Enzyme production was adapted from Tanaka and coworkers.^[497] Briefly, *E. coli* BL21(DE3) carrying the recombinant plasmid was grown at 37°C, 200 rpm in LB-Luria medium until the absorbance of the culture suspension reached an OD₅₉₀ of 0.4–0.6. The expression of recombinant PhoN-Se variants was induced by adding 0.5 mM IPTG. Growth was continued at 37°C for 4–6 h. Cells were harvested by centrifugation (6 000 x g, 4°C for 15 min) for further analysis.

G VI.5.1.2 Release of periplasmic *PhoN-Se V78L* variants by osmotic shock

For *PhoN-Se* variants bearing a SP for secretion into the periplasmic space, osmotic shock was performed. Therefore, the cell pellet was resuspended in osmotic shock solution 1 (pH 8; 20 mM Tris-HCl, 2.5 mM EDTA, 20% (w/v) sucrose) to OD₅₉₀ = 6.0, and incubated on ice for 10 min. After centrifugation at 10 000 x g, 4°C for 2 min, the cell pellet was resuspended in osmotic shock solution 2 (pH 8; 20 mM Tris-HCl, 2.5 mM EDTA) to OD₅₉₀ = 6.0 and incubated on ice for 10 min. The secreted *PhoN-Se* variants were obtained in the supernatant after centrifuging (8 000 x g, 4°C for 10 min). The supernatant was dialysed against 20 mM sodium acetate buffer (pH 6.0) at 4°C with slight shaking overnight. The solution was passed through a 0.45 µm filter (Millipore) and concentrated by centrifugation with a centrifugal membrane concentrator (10 kDa MWCO; UFC901024, Millipore).

For *PhoN-Se* variants lacking the SP, cell lysis was performed by sonication as described in G I.5.

Protein content of all fractions was determined by Bradford assay (G I.7) and SDS-PAGE analysis (G I.8) performed to determine the cellular localization of *PhoN-Se* variants. Results are shown in **Figure D-20**.

G VI.5.2 The wild type acidic phosphatase *PhoN-Sf*

G VI.5.2.1 Cloning and expression of *PhoN-Sf*

The *phoN-Sf* gene^[497] was subcloned into pET26b(+) by GenScript™ utilizing *NdeI/HindIII* restriction sites (H I.1.6.5). The lyophilized plasmid DNA was reconstituted in nuclease-free water and competent *E. coli* BL21(DE3) transformed with pET26b(+)-*phoN-Sf*.

Enzyme expression was in accordance to Tanaka *et al.* with minor modifications.^[497] As for *PhoN-Se* variants, preculture and main culture were grown in LB-Luria medium supplemented with Kan (50 µg·mL⁻¹). Protein production was performed with 0.5 mM IPTG at 37°C, 200 rpm for 6–8 h. Soluble *PhoN-Sf* was released from the

periplasmic space by osmotic shock as described above, except that it was dialysed against 30 mM Tris-HCl (pH 7.5) containing 30 mM NaCl and 10% (v/v) glycerol. In parallel, CFE were prepared as before (G I.5).

SDS-PAGE analysis revealed that *PhoN-Sf* showed increased and stable expression levels even under atypical phosphatase cultivation conditions such as decreased expression temperatures $\leq 30^{\circ}\text{C}$ (**Figure D-23A**). This is in contrast to *PhoN-Se* variants (e.g., *PhoN-Se V78L_{SP}*; **Figure D-23B**) and advantageous since the simultaneous production of multiple pathway enzymes in the same host cell is routinely performed at lower temperatures.^[6, 43, 348] Furthermore, *PhoN-Sf* yields could be greatly increased when inducing protein production at $\text{OD}_{590} \geq 3.0$ (**Table G-23**, entry #61).

G VI.5.3 The *E. coli* phosphatase *YqaB*

G VI.5.3.1 Cloning and expression of *YqaB*

The *E. coli* phosphatase *YqaB*^[501] was already implemented in an aldolase-coupled cascade *in vivo*.^[293] The gene was ordered from GenScript™ in a pCDFDuet-1 vector utilizing *NcoI/BamHI* restriction sites. The sequence was checked by Sanger sequencing after reisolation of plasmid DNA from *E. coli* DH5 α transformants (H I.1.6.6).

Preliminary expression was performed in LB-Miller medium supplemented with Str, protein production induced at $\text{OD}_{590} = 0.5$ with 0.5 mM IPTG, and cultivation of *E. coli* BL21(DE3) transformants continued at 37°C , 200 rpm for 3–21 h. Subsequently, *YqaB* was successfully produced at different expression temperatures with the highest soluble amounts produced at 20°C after 21 h (**Figure D-24A**).

Expression at low temperatures was also beneficial in terms of greatly reduced yields of target protein in insoluble fractions. Contrary to the *PhoN-Se* variants and *PhoN-Sf* from *S. enterica* and *S. flexneri*, respectively, *YqaB* is a cytosolic phosphatase.^[502] Therefore, the osmotic shock protocol was not performed. released by osmotic shock. The hydrolysis of pNPP confirmed the activity of CFEs containing *YqaB*. Preparation of CFEs and SDS-PAGE analysis were performed as before.

G VI.5.4 Phosphatase activity assays

G VI.5.4.1 Photometric pNPP assay

CFEs (or concentrates) containing the target phosphatase were prepared according to the optimized conditions in **Table G-23**. The assay is based on the hydrolysis of p-nitrophenyl phosphate (pNPP; **Figure G-23**) and performed in 12-well plates (Greiner) in 50 mM Tris-HCl (pH 7.5) containing $0.05 \text{ mg}\cdot\text{mL}^{-1}$ CFE. The pNPP substrate was added last to a final concentration of 20 mM ($V_{\text{total}} = 1.0 \text{ mL}$). The increase in absorbance at $\lambda = 405 \text{ nm}$ was followed at 37°C for 10 min (plate reader: Anthos Zenyth 3100).^[497]

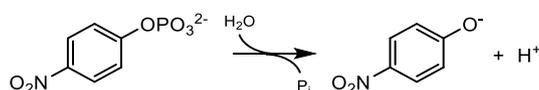


Figure G-23. Reaction scheme pNPP assay. In the presence of phosphatases, the phosphate group from pNPP (left) is cleaved, yielding yellow p-nitrophenol (right) and inorganic phosphate P_i .

G VI.5.4.2 Colorimetric BCIP agar plate assay

The assay is based on the hydrolysis of 5-bromo-4-chloro-3-indolyl phosphate (BCIP; **Figure G-24**).^[499] Solid Sperber medium (pH 7.2; $16 \text{ g}\cdot\text{L}^{-1}$ agar, $10 \text{ g}\cdot\text{L}^{-1}$ glucose, $2.5 \text{ g}\cdot\text{L}^{-1}$ sodium phytate, $0.5 \text{ g}\cdot\text{L}^{-1}$ yeast extract, $0.1 \text{ g}\cdot\text{L}^{-1}$ CaCl_2 , $0.25 \text{ g}\cdot\text{L}^{-1}$ MgSO_4) was autoclaved and prior to pouring $0.025 \text{ g}\cdot\text{L}^{-1}$ BCIP and, if applicable, the appropriate antibiotic to the standard working concentration (**Table G-1**) were added.

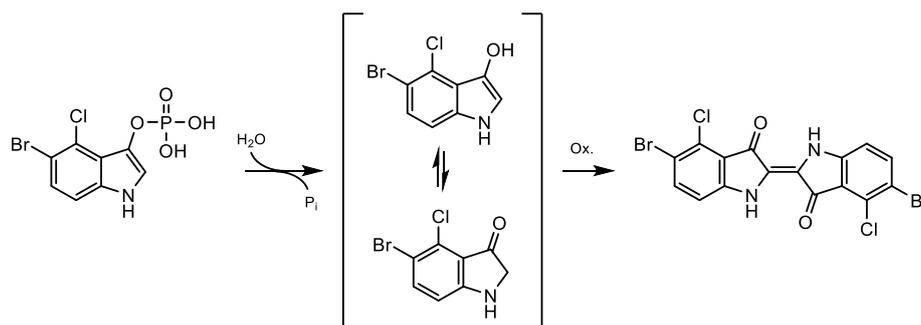


Figure G-24. Reaction scheme BCIP assay. In the presence of phosphatases, the phosphate group from BCIP (left) is cleaved, yielding the intermediate 5-bromo-4-chloro-indolyl (center) and inorganic phosphate P_i . Subsequent oxidation by air forms the blue dye 5,5'-dibromo-4,4'-chloro indigo (right).

The target strains to be tested for phosphatase activity were grown in liquid Sperper medium supplemented with the appropriate antibiotic, if applicable, at 37°C, 200 rpm overnight and streaked out on solid Sperper medium. Plates were incubated upside down at 37°C overnight. Blue colonies accumulated the dye 5,5'-dibromo-4,4'-chloro indigo (**Figure G-24**), which indicates enhanced phosphatase activity. Example plates are shown in **Figure D-21**.

G VI.6 Characterization of the DHA kinase (*DhaK*) from *C. freundii*

DhaK: Dihydroxyacetone kinase from *Citrobacter freundii* CECT 4626 (GenBank: DQ473522.1)
Gene size: 1 659 bp AA: 552 SDS-PAGE: 58 kDa (experimental)

G VI.6.1 Cloning and enzyme expression

G VI.6.1.1 Insoluble enzyme production with pET22b(+)_{dhaK} according to Gottschalk

The *dhaK* gene^[539] was ordered from GenScript™ subcloned into pET22b(+) utilizing *NdeI/XhoI* restriction sites. The lyophilized plasmid DNA was reconstituted in nuclease-free water and competent *E. coli* BL21(DE3) transformed with pET22b(+)_{dhaK} according to G I.9.1.

For preliminary expression, the protocol was adapted from Daniel *et al.* as follows: 4 mL of LB-Miller supplemented with Amp were inoculated with a single colony of the *E. coli* BL21(DE3) transformant and incubated at 37°C with shaking (200 rpm) overnight. TB medium supplemented with Amp was inoculated with 1% (v/v) of the preculture and cultivated at 37°C, 200 rpm until an $OD_{590} = 0.5$ was reached. Protein production was induced with 0.5 mM IPTG and performed at 37°C, 150 rpm for 20 h. CFEs were prepared (G I.5) and soluble and insoluble fractions analyzed by SDS-PAGE (G I.8). Since *DhaK* was exclusively found in insoluble fractions, expression was performed at lower temperatures (**Figure G-25**), which did not yield soluble *DhaK*.

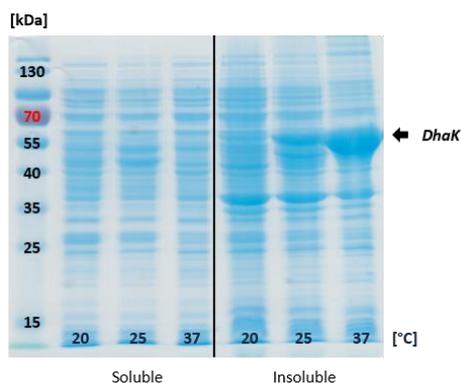


Figure G-25. Preliminary expression of *DhaK* from pET22b(+). Expression with 0.5 mM IPTG at different temperatures did not yield soluble *DhaK*.

Protein BLAST (basic local alignment search tool) from <https://blast.ncbi.nlm.nih.gov/Blast.cgi> of the *DhaK* gave 100% identity to the sequence P45510.3, which is identical to the sequence published by Daniel and coworkers.^[539] However, the sequence is different from the protein sequence from *Citrobacter freundii* CECT 4626 (ID: ABF06666.1), which was successfully expressed and functionally tested by Iturrate *et al.*, for example (H I.1.10.2; **Figure H-22**).^[363]

G VI.6.1.2 Soluble enzyme production with pRSETa_*dhaK* according to García-Junceda

The pRSETa_*dhaK* vector was kindly provided by Prof. Eduardo García-Junceda from the University of Madrid, Spain. Competent *E. coli* DH5 α was transformed with the target plasmid, reisolated, and sent for sequencing. Sanger sequencing confirmed the presence and integrity of the *dhaK* gene from *C. freundii* CECT 4626 (ID: ABF06666.1).

For expression, *E. coli* BL21(DE3) was transformed with pRSETa_*dhaK* (G I.9.1), a single colony picked from LB-Amp agar plates, 4 mL LB-Miller medium supplemented with Amp inoculated, and grown at 37°C, 200 rpm overnight. Cultivation was performed in TB-Amp medium inoculated with 1% (v/v) of the preculture. Cells were grown at 37°C, 200 rpm until OD₅₉₀ = 0.5–0.8, IPTG added to a final concentration of 0.5 mM, and protein production performed at 30°C, 150 rpm for 20 h.^[363] CFEs were prepared as before. SDS-PAGE analysis showed the successful production of soluble *DhaK* from the pRSETa backbone (**Figure E-22**).

G VI.6.2 Functional testing by metabolomic analysis

The experiment was conducted in cooperation with S. Milker, who performed subsequent metabolomic analysis.

For metabolomics, untransformed *E. coli* BL21(DE3) and pRSETa_*dhaK* transformants were cultivated in M9-N* medium without and in the presence of Amp, respectively. Expression was performed in shaking flasks suitable for metabolomics filled with 1/25 of the flask volume and vigorous shaking (350 rpm), but otherwise followed the conditions described above. Expression was successful according to SDS-PAGE analysis (see **Figure E-22** for comparison).^[154]

Sampling at different time points was performed according to the group of Sauer as follow: Immediately after wetting the filter with M9-N* medium, 2 mL cell suspension were dispersed onto a nitrocellulose filter and washed with 4 mL medium. As a sidenote, the filter must not run dry while sampling. The filter was removed with tweezers, placed into a 15 mL Greiner tube, and snap-frozen in liquid N₂. The samples were stored at -80°C until extraction. For extraction, 2 mL of the extraction solution (60% (v/v) EtOH in dH₂O; both HPLC grade) were preheated to 78°C in a water bath. U-¹³C IS was added to the Greiner tube with the filtered sample and the extraction solution (2 mL) were added. The Greiner tube was transferred to the water bath and extracted at 78°C

for 3 min, vortexing once every minute. The samples were stored at -40°C in precooled ethanol until centrifugation ($12000 \times g$, -10°C for 10 min). The samples were dried at 30°C and 0.12 mbar (Christ Speedvac RVC 2-25 CD plus and Alpha 2-4 LD plus). Samples were stored at -80°C until LC-MS/MS measurement.^[154, 551]

For LC-MS/MS measurement, dried extracts were resuspended in 100 μL water, centrifuged ($12\ 000 \times g$, 4°C for 3 min), and the supernatant used for analysis. Separation was achieved with a Luna 3u NH2 100A column (150×2.00 mm, Phenomenex) with a binary gradient method (Solvent A: acetonitrile / Solvent B: 10 mM ammonium acetate; pH 9.9). Gradient parameters were as follows: 20–100% (v/v) B in 0–24 min; 100% (v/v) B from 25–34 min; 20% (v/v) B for 1 min; total time of gradient: 35 min. Detection was performed with a tandem mass spectrometry detector with an ESI ion source (LCMS-8040, Shimadzu) in Multi Reaction Monitoring (MRM) mode. Peak areas were normalized to fully ^{13}C -labeled IS and absolute quantification of metabolites was achieved with linear calibration curves of standards.^[154] Finally, concentrations were normalized to the amount of biomass.

Production of DHAP was induced by the addition of DHA. Therefore, 20 mM DHA were added after 5.8 h and another 10 mM DHA after 24.5 h and sampling done as described above. Untransformed *E. coli* BL21(DE3) was used to compare intracellular DHAP concentrations.^[154]

G VI.7 Characterization of mutant GDHs from *B. subtilis*

GDH: NAD(P)H dependent GDH from *Bacillus subtilis*
Gene size: 786 bp AA: 261 SDS-PAGE: 34 kDa (calculated)

G VI.7.1 Cloning and enzyme expression

Two GDH mutants were kindly donated by Prof. Andy Bommarius from the California Institute of Technology, CA, USA: the double mutated *GDH*_{2xBs} (E170K and Q252L) and the seven times mutant *GDH*_{7xBs} (P45A, N46E, F155Y, E170K, V227A, W230F, and Q252L) from *B. subtilis*. Both engineered GDHs were cloned into pET28a and contained N-terminal 6xHis tags.^[332]

The GDH mutants were expressed in *E. coli* BL21(DE3) transformants according to Vázquez-Figueroa *et al.* without modification of the enzyme production protocol. Briefly, LB-Miller medium supplemented with Kan ($50 \mu\text{g}\cdot\text{mL}^{-1}$) was inoculated with 1% (v/v) preculture grown at 37°C , 200 rpm overnight. Expression was induced at $\text{OD}_{590} = 0.3$ – 0.5 with 0.1 mM IPTG at 37°C , 150 rpm for 20 h (**Figure E-4A**).

G VI.7.2 Purification of GDH mutants

G VI.7.2.1 Purification by HS

The two engineered GDHs showed increased thermostability compared to the wild type GDH from *B. subtilis*.^[332] Therefore, transformants were cultivated as before; cell lysis by lysozyme and subsequent HS purification performed as outlined in G I.15.1 but not lyophilized.

G VI.7.2.2 Purification by IMAC

E. coli BL21(DE3) containing the GDH-encoding plasmids were cultivated as before. Cells were centrifuged ($8\ 000 \times g$, 4°C for 10 min), the supernatant discarded, and the cell pellet resuspended in 50 mM Tris-HCl (pH 7.5) containing of 20% (ω/v) glycerol and 0.1 μM PMSF as protease inhibitor. Cells were lysed by sonication as described in G I.5. Purification by IMAC and subsequent elution was also performed in the presence of 20% (ω/v) glycerol, but otherwise followed the procedure outlined in G I.15.2. Combined fractions were snap-frozen in liquid N_2 and samples lyophilized under high vacuum (Christ Gamma 2-20 lyophilizer).

G VI.7.3 Photometric NADP⁺ consumption assay

The activity of purified GDH mutants was assayed by the NADP⁺ dependent oxidation of D-glucose. A commercially available GDH from *Pseudomonas* sp. (Sigma) was employed as the positive control.

Glycerol concentrates (see G VI.7.2.2) were 1:5-diluted in 50 mM Tris-HCl (pH 7.5) and lyophilisates dissolved in buffer prior to measurement. The protein content was determined by Bradford assay (G I.7). The assay conditions were as follows: 50 mM NaH₂PO₄ buffer (pH 8.0) containing D-glucose (100 mM) and GDH (0.05 mg·mL⁻¹) were incubated in a quartz cuvette at 37°C for 1 min (CPS temperature controller, Shimadzu). NADP⁺ was added with a plastic spatula from a 100 mM stock to a final concentration of 1 mM. It was quickly mixed with the spatula and the production of NADPH followed at $\lambda = 340$ nm, 37°C for 30 s (UV spectrophotometer UV-1800, Shimadzu).^[332]

Activities of the commercial GDH and GDH mutants purified by either HS or IMAC are plotted in **Figure E-5A**.

G VI.7.4 Preliminary evaluation of storage conditions

P. Wolf performed the optimization of isolation, purification, and the preliminary evaluation of storage conditions for *GDH_{2xBs}*.^[522]

Briefly, lyophilisates after HS (see G VI.7.2.1) were stored at 4°C. The glycerol concentrates produced after lyophilization in G VI.7.2.2 were stored at either 4°C or -20°C. For comparison, freshly purified GDH mutants were prepared. For subsequent activity measurements, lyophilisates were dissolved in 50 mM Tris-HCl (pH 7.5) and glycerol stocks 1:5-diluted in buffer. The total amount of protein determined by Bradford assay (G I.7). Photometric assays were carried out in duplicates as described in G VI.7.3. Activities of freshly purified *GDH_{2xBs}* was compared to differently stored preparations (see above). Results are plotted in **Figure E-5B**.

G VI.7.5 *GDH_{2xBs}* as cofactor regenerating enzyme *in vitro*

For NAD(P)H regeneration, *GDH_{2xBs}* (0.05 mg·mL⁻¹) and 100 mM D-glucose were routinely used in biotransformations. The addition of 0.1 mM NAD(P)⁺ started the reactions.

The applicability of *GDH_{2xBs}* as cofactor regenerator in enzymatic reduction reactions was tested with *YqjM* and *CAR_{Ni}*, for example, which were also purified by IMAC (G VI.8.2 and G VI.3.1.2, respectively).

G VI.8 Characterization of EREDs

YqjM: NAD(P)H dependent ERED from *Bacillus subtilis*
Gene size: 1 017 bp AA: 338 SDS-PAGE: 37 kDa (experimental)

G VI.8.1 Expression of *YqjM* from different vectors

YqjM was produced from the following constructs:

- pET28b(+) conveying resistance to Kan; N-terminal 6xHis tag (38 kDa)^[552]
- pHT conveying resistance to Amp; N-terminal 6xHis tag and TEV cleavage site (38 kDa)^[526]
- pET21b(+) conveying resistance to Amp; untagged (38 kDa)^[553]
- pSF1 conveying resistance to Amp; untagged (37 kDa)^[554]

All constructs were sent for Sanger sequencing (Microsynth) and the *yqjM* sequences and their genetic context confirmed (H I.1.10.5). The optimized expression protocols are summarized in **Table G-23**. Briefly, precultures were prepared in LB-Miller medium supplemented with the appropriate antibiotic at 37°C, 200 rpm overnight. TB medium containing the appropriate antibiotic was inoculated with 1% (v/v) preculture and cultivated at 20°C,

120 rpm for 8 h. Protein production was induced with 0.1 mM IPTG and expression performed for additional 16 h. Cells were harvested (6 000 x g, 4°C for 15 min). The pellet was resuspended in 50 mM Tris-HCl (pH 7.5) containing 0.1 mM FMN and 0.1 mM PMSF. Cell lysis was performed as in G I.5 and protein fractions analyzed by SDS-PAGE (G I.8; **Figure G-26**).

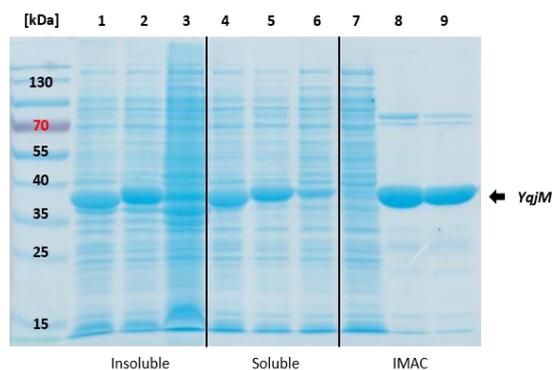


Figure G-26. Expression of *YqjM* from different constructs. SDS-PAGE analysis of *YqjM* samples produced in *E. coli* BL21(DE3) from pSF1 (**1** and **4**), pET28b(+) (**2** and **5**), and pHT (**3** and **6**). *YqjM* produced from pHT was purified by IMAC: flow-through (**7**), eluate (**8**), and concentrate (**9**). Sample loading normalized to 10 µg total protein per lane.

G VI.8.2 Purification of N-terminally 6xHis-tagged *YqjM* by IMAC

After cell lysis in the presence of 0.1 mM FMN and PMSF, cell debris were removed by centrifugation (16 000 x g, 4°C for 25 min). Purification of *YqjM* from the supernatant was done by IMAC and followed the procedure outlined in G I.15.2 except that the binding buffer contained 10 µM FMN. Pooled fractions were concentrated (10 kDa MWCO; UFC901024, Millipore) and stored at 4°C (**Figure G-27A**). Purity was checked by SDS-PAGE analysis (G I.8; **Figure G-26**, **Figure G-27A**, and **Figure E-6**).

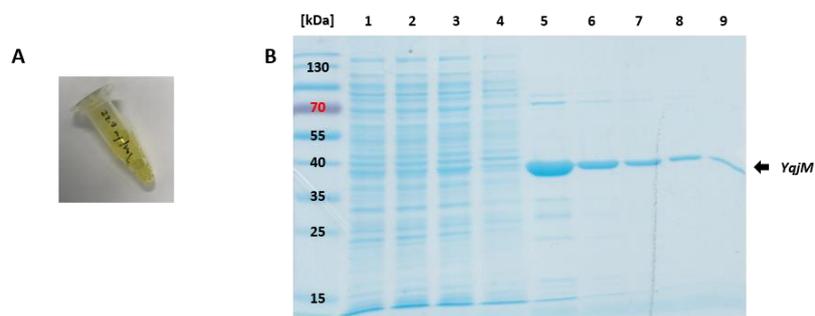


Figure G-27. Purification of *YqjM* by IMAC. (A) *YqjM* concentrate. (B) SDS-PAGE analysis fractions produced by IMAC from *YqjM*: Flow-throughs from loading (**1–3**), washing (**4**), and eluates (**5–9**). Sample loading normalized to 10 µg total protein per lane except lanes **8** and **9** due to low protein content of <math><0.4 \text{ mg}\cdot\text{mL}^{-1}</math> (sample preparation: 10 µL sample + 15 µL SDS sample buffer; loading: 20 µL).

G VI.8.3 (Assymmetric) reductions *in vitro*

The *YqjM* concentrate (22.7 mg·mL⁻¹ according to Bradford assay) was applied to the α, β-unsaturated C=C bonds in the substrates **10** and **12**. Screenings were carried out in 50 mM Tris-HCl (pH 7.5) containing substrate (10 mM), 2.5% (v/v) ACN as cosolvent, FMN (0.1 mM), and purified *YqjM* (0.5 mg·mL⁻¹) at 25°C (250 rpm) for 20 h. Recycling of NADP⁺ (0.1 mM) was achieved in the presence of D-glucose (100 mM) and purified *GDH*_{2xBs}

(0.05 mg·mL⁻¹) as described in G VI.7.5. NADP⁺ was added last to the biotransformation mixture. Sampling for GC/FID analysis was done after adding the NADP⁺ (t₀^{*}), 2 h, and 20 h. The reduction of **10** is shown in **Figure E-7**.^[522]

G VI.9 Preliminary expression of enzymes for future modular pathway extension

Considering the expansion of the mini-pathway and utilize aldehyde intermediates to access alternative product classes such as primary amines (G VI.9.1) or different carboligation products (G VI.9.2), enzymes of the two classes were expressed in course of this thesis (see also chapter F).

G VI.9.1 From aldehyde intermediates to terminal amines: ω-transaminases

Plasmids harboring ω-transaminases (ω-TAs) provided by Prof. Uwe T. Bornscheuer from the University of Greifswald, Germany. The optimized expression conditions for ω-TAs are given in **Table G-23**.

Enzymes encoded on pET22b(+) plasmids were propagated on LB-Amp plates from permanent cultures stored at -80°C (*3FCR*, *3GJU*, *3HMU*, *3i5T*, and *AspFum*). *VfIH6*, which is encoded on pET24a, was propagated on LB-Kan plates. Enzyme production was performed in TB medium supplemented with the appropriate antibiotic and induced at OD₅₉₀ = 0.7 with 0.1 mM IPTG at 20°C, 200 rpm for 20 h.

Enzymes encoded on pGASTON were propagated on LB-Amp plates. Enzyme production was performed in TB medium supplemented with Amp at OD₅₉₀ = 0.7 with 0.2% (ω/v) L-Rhm at 20°C, 200 rpm for 20 h.

CFEs were prepared as in G I.5 in the presence of 1 mM PLP during sonication. SDS-PAGE analysis confirmed the soluble expression of ω-TAs (G I.8; **Figure G-28**). The substrate scope toward the aldehydes from B I was not tested due to time constraints.

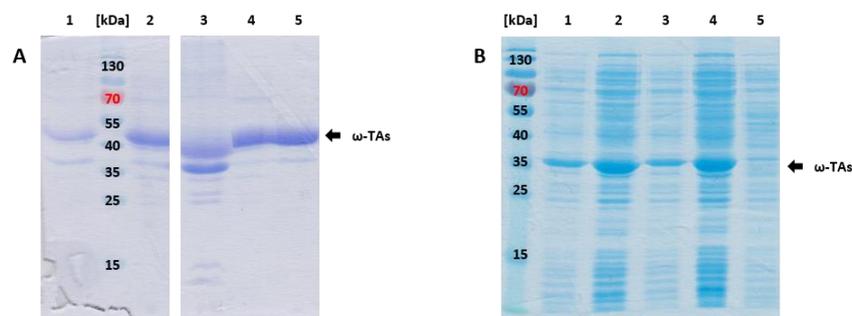


Figure G-28. Expression of ω-TAs. Enzymes were overexpressed in *E. coli* BL21(DE3). (A) Soluble fractions containing *3HMU* (1), *3i5T* (2), *3FCR* (3), *3GJU* (4), and *VfIH6* (5). Sample loading normalized to 10 μg total protein per lane. (B) Soluble fractions containing *AspFum* (1 and 2), *NeoFis* (3 and 4), and CFE from untransformed host strain (5). Cultivation in LB-Miller (1 and 3) and TB medium (2 and 4). Sample loading normalized to OD₅₉₀ = 7.0.

P. Schaaf successfully tested structurally different acetophenone substrates with lyophilized whole cells^[31, 555] expressing ω-TAs to produce chiral amines.^[549]

G VI.9.2 Alternative carboligations by a pyruvate decarboxylase (PDC) from *A. pasteurianus*

The pET22b(+)_{pdc_{Ap}} plasmid was kindly donated by Prof. Dörte Rother from the Forschungszentrum Jülich GmbH, Germany. The expression protocol for the PDC from *Acetobacter pasteurianus* (*A. pasteurianus*) was adapted: A preculture of *E. coli* BL21(DE3) harboring pET22b(+)_{pdc_{Ap}} was grown in 20 mL LB-0.8G supplemented

with Amp in a 100 mL baffled shaking flask at 37°C, 275 rpm overnight. The protein was produced in LB-5052 supplemented with Amp in baffled 1 L shaking flasks with a filling volume of 200 mL. It was inoculated with 0.2% (v/v) preculture, cultivated at 37°C, 150 rpm for 4 h, the temperature lowered to 20°C, and cultivated for another 20 h.

CFEs were prepared as in G I.5 and SDS-PAGE analysis confirmed the soluble expression of *PDC_{Ap}* (G I.8; experimental: 60 kDa). The coupling of *AlkJ* and *PDC_{Ap}* and the ligation of **1c** and different donor molecules^[556] *in vivo* was not tested due to time constraints.

G VI.10 Expression of miscellaneous proteins

G VI.10.1 Oleate hydratase *OhyA2*

The *ohyA2* gene from *Stenotrophomonas maltophilia* was donated by Prof. Jin-Byung Park from the Ewha Womans University of Seoul, South Korea. The pET28a(+)*_ohyA2* was introduced into *E. coli* BL21(DE3) and *OhyA2* expressed according to Jeon *et al.* without modification.^[557] According to SDS-PAGE analysis (G I.5 and G I.8), the oleate hydratase was exclusively expressed in soluble fractions visible as a prominent band at 68 kDa (data not shown).

Additionally, permanent cultures of *E. coli* DH5 α harboring pET28a(+)*_ohyA2* were prepared and stored at -80°C for further use.

G VI.10.2 Outer-membrane protein *AlkL*

The *alkL* gene was amplified from the pGEc47 cosmid^[289] (see G VI.2.4.1) with the following primer pairs:

AlkL_NdeI *fwd*: 5'-GCGCATGAGTTTTTCTAATTATAAAG-3'

AlkL_BamHI *rev*: 5'-GCGGGATCCTTAGAAAACATATG-3'

The PCR mixture was prepared as in **Table G-45**. Thermal cycle conditions were adapted from the *alkJ* insert amplification and are given in **Table G-46**.

Table G-45. PCR mixture for *alkL* insert amplification

	<i>Pfu</i> ⁺ polymerase mix	Final concentration
5.0 μ L	10X <i>Pfu</i> buffer	1X
2.0 μ L	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 μ L	AlkL_NdeI <i>fwd</i> (5 μ M)	0.25 μ M
2.5 μ L	AlkL_BamHI <i>rev</i> (5 μ M)	0.25 μ M
0.5 μ L	pGEc47 (139 ng· μ L ⁻¹)	1.39 ng· μ L ⁻¹
1.0 μ L	DMSO	2% (v/v)
0.5 μ L	<i>Pfu</i> ⁺ polymerase (5 U· μ L ⁻¹)	2.5 U
36.0 μ L	nuclease-free water	-

Table G-46. Thermal cycle conditions for *Pfu*⁺ amplifying the *alkL* insert

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1
Denaturation	95	45 s	30
Annealing	48	30 s	
Extension	72	42 s	
Terminal extension	72	5 min	1
Hold	12	∞	1

PCR products were purified with the innuPREP PCRpure Kit from Analytik Jena and double-digested with *NdeI* and *BamHI* (NEB) according to the supplier at 37°C for 2 h. The target vector pCOM^[558] was digested accordingly. Restriction enzymes were inactivated at 80°C for 20 min. Both the trimmed insert and the linearized vector were purified with the innuPREP PCRpure Kit from Analytik Jena as before. Ligation was performed as outlined in **Table G-17** at 25°C for 20 min. Competent *E. coli* DH5 α and BL21(DE3) cells were transformed with 8 μ L of ligation mixture according to G 1.9.2. After recovery, 150 μ L cell suspension were plated on LB-Amp agar plates and incubated at 37°C overnight.

From the *E. coli* BL21(DE3) cell suspension, a LB-Miller preculture (4 mL) supplemented with Amp was inoculated (200 μ L) and grown at 37°C, 200 rpm overnight. Expression conditions were adapted from Smits *et al.* as follows: The main culture was inoculated with 1% (v/v) preculture, cultivated at 37°C with shaking (200 rpm) until an OD₅₉₀ = 0.5–1.0 was reached. Protein production was induced by 0.05% (w/v) dicyclopropylketone (DCPK) and expression continued at 30°C, 200 rpm for 6 h.^[558] SDS-PAGE analysis of the preliminary expression did not show a protein band corresponding to the calculated size of *AlkL* (24.9 kDa; data not shown) and was not further studied in this thesis.

G VII Construction of vectors for the coexpression of pathway enzymes

G VII.1 Plasmids harboring the *alkJ* and the *fsa1-A129S* gene in different genetic configurations

In the following, the detailed cloning procedure for vectors for the coproduction of *AlkJ* and *Fsa1-A129S* from different genetic configurations will be described.

G VII.1.1 Construction of pKA1_alkJ:fsa1-A129S in operon configuration (pOPE)

The DNA fragment containing *fsa1-A129S* was amplified by PCR using the DNA polymerase *Pfu*⁺ and the pET16b vector containing the *fsa1-A129S* gene, which was kindly provided by Prof. Dr. Wolf-Dieter Fessner, TU Darmstadt, Germany, with the following primer pair:

OPE1A129SRBS *fwd*: 5'-GGATCCGGCTGCTAACTCTAGAAATAATTTTGTTTAAC-3'

OPE1A129S *rev*: 5'-CAGCAGCCAACCTCAGCTTTAAATCGACGTTCTGCCAAAC-3'

The pKA1_alkJ backbone was amplified from the previously assembled pKA1_alkJ template (G VI.2.4.1) using OptiTaq with the following primers:

OPE1pAJ3A129S fwd: 5'-CTAGAGTTAGCAGCCGGATCCTTACATGCAGACAGCTATC-3'

OPE1pAJ3A129S rev: 5'-GGCAGAACGTCGATTTAAAGCTGAGTTGGCTGCTGCCAC-3'

The optimal annealing temperatures for both insert and backbone amplification were determined by gradient PCR as outlined in G I.13. For the insert amplification, the first gradient PCR (45–65°C) showed an unspecific amplification product at 3 kb besides the expected fragment with a size about 0.75 kb (**Figure G-29A**). The PCR was optimized by lowering the temperature gradient (40–52°C), finally giving the optimized annealing temperature of 44.4°C (**Figure G-29B**).^[435]

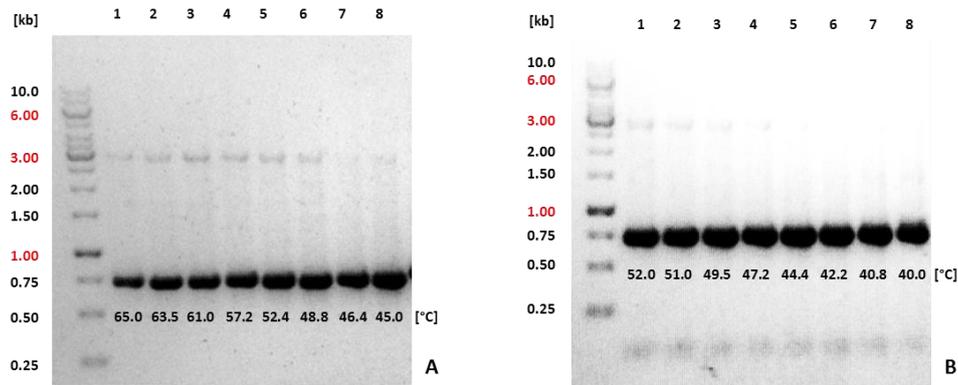


Figure G-29. Annealing temperature optimization of the *fsa1-A129S* insert for OPE plasmid assembly. Gradient PCR with annealing temperatures from (A) 45–65°C and (B) 40–52° with the unspecific 3.00 kb PCR product disappearing at lower temperatures; 0.85% (w/v) agarose. (B) was adapted from T. Wiesinger *et al.* (2017).^[435]

For the backbone amplification, gradient PCR (45–65°C) showed an optimized annealing temperature of 57.2°C. A very weak, unspecific PCR product with a size >10 kb was detectable at lower temperatures (**Figure G-30**).

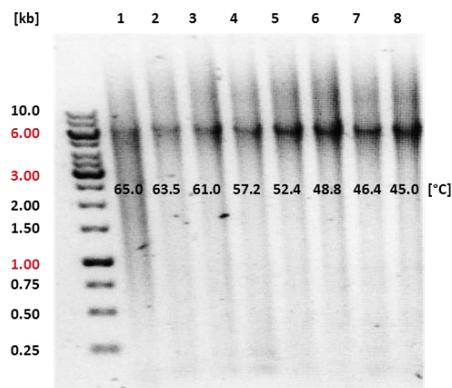


Figure G-30. Annealing temperature optimization of the pKA1_alkJ backbone for OPE plasmid assembly. Gradient PCR with annealing temperatures from 45–65°C; 0.85% (w/v) agarose. The figure was adapted from T. Wiesinger *et al.* (2017).^[435]

The preparation of each PCR mixture and the optimized thermal cycle conditions are summarized in **Table G-47** and **Table G-48**, respectively.

Table G-47. PCR reaction mixtures for the construction of the OPE plasmid

<i>Pfu</i> ⁺ PCR mix		Final concentration	Opti <i>Taq</i> PCR mix		Final concentration
5.0 µL	10X <i>Pfu</i> buffer	1X	5.0 µL	10X Pol B buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	OPE1A129SRBS <i> fwd</i> (5 µM)	0.25 µM	2.5 µL	OPE1pAJ3A129S <i> fwd</i> (5 µM)	0.25 µM
2.5 µL	OPE1A129S <i> rev</i> (5 µM)	0.25 µM	2.5 µL	OPE1pAJ3A129S <i> rev</i> (5 µM)	0.25 µM
1.0 µL	pET16b_ <i> fsa1-A129S</i> (100 ng·µL ⁻¹)	2.0 ng·µL ⁻¹	1.0 µL	pKA1_ <i> alkI</i> (128 ng·µL ⁻¹)	2.5 ng·µL ⁻¹
1.0 µL	DMSO	2% (v/v)	1.0 µL	DMSO	2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U	0.5 µL	Opti <i>Taq</i> polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-	35.5 µL	nuclease-free water	-

Table G-48. Optimized thermal cycle conditions for the construction of the OPE plasmid

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (Opti <i>Taq</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	44.4	30 s	30	Annealing	57.2	30 s	30
Extension	72	45 s		Extension	72	7.5 min	
Terminal extension	72	3 min	1	Terminal extension	72	10 min	1
Hold	4	∞	1	Hold	4	∞	1

After *DpnI* digestion and purification, the 1:1 and 4:1 mixtures were prepared (**Table G-18**) for subsequent DNA fragment assembly.

G VII.1.1.1 Successful pOPE assembly from two linear DNA fragments by SLiCE

Since FC (G II.2.1.3) did not yield putative correct assemblies post transformation of RbCl-competent *E. coli* TOP10, insert and backbone amplification were repeated as summarized in **Table G-47** and **Table G-48**. The 1:1 and the 4:1 mixtures were prepared (7.5 µL + 7.5 µL and 12 µL + 3 µL, respectively), incubated in the presence of *DpnI* at 37°C for 30 min and *DpnI* heat-inactivated. The mixtures were purified with the QIAquick® PCR Purification Kit (28104, QIAGEN) and the SLiCE reaction carried out according to **Table G-49** with both mixtures at 37°C for 1 h. The reactions were left at RT for 1 h and stored at 4°C overnight. Freshly prepared RbCl-competent *E. coli* BL21(DE3) cells (50 µL) were transformed with 5 µL of SLiCE reaction mixture as described in G I.9.3. Putative positive clones were picked from LB agar plates supplemented with Cam (37 µg·mL⁻¹).^[435]

Table G-49. SLiCE reaction mixture for pOPE assembly with two fragments

SLiCE assembly mix		Final concentration
1.0 µL	10X SLiCE buffer	1X
7.0 µL	DNA fragment mix	-
1.0 µL	ATP (10 mM)	1 mM
1.0 µL	SLiCE (from TOP10)	10% (v/v)

Colony PCR was performed for the insertion of the *fsa1-A129S* gene and plasmid DNA isolated from positive clones (OPE1, OPE4–5, OPE7; **Figure G-31**). Isolated plasmid DNA was digested with the restriction enzyme *NcoI* at 37°C for 2 h as given in **Table G-50** to verify successful plasmid assembly (**Figure G-32**).^[435]

Table G-50. *NcoI* control digestion of putative OPE plasmids

<i>NcoI</i> digestion mix		Final concentration
1.0 µL	10X Tango buffer	1X
1.5–3.0 µL	DNA (50–100 ng·µL ⁻¹)	≈15 ng·µL ⁻¹
0.5 µL	<i>NcoI</i> (10 U·µL ⁻¹)	5 U
5.5–7.0 µL	nuclease-free water	-

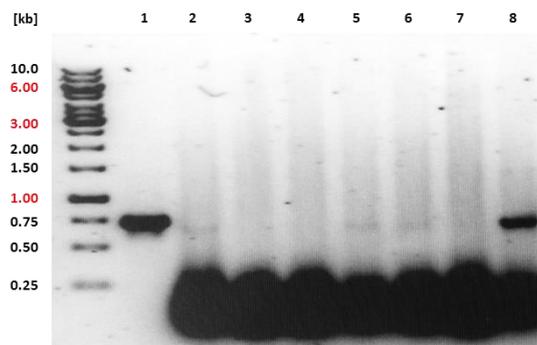


Figure G-31. Colony PCR for the detection of the *fsa1-A129S* in clones containing the putative OPE plasmid after SLiCE assembly. Control PCR of the *fsa1-A129S* gene from the pET16b_ *fsa1-A129S* template (**1**) and colony PCR of seven putative positive clones OPE1–7 (**2–8**). The *fsa1-A129S* insert could be detected in the clones OPE1 (**2**), OPE4–5 (**5–6**), and OPE7 (**8**).

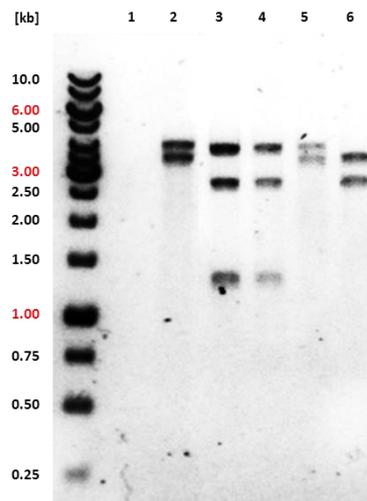


Figure G-32. Control digestion with *NcoI* of putative OPE plasmids assembled by SLiCE. Digestion mixtures containing no plasmid DNA (**1**), the parent pKA1_ *alkJ* plasmid (**2**), and four putative pKA1_ *alkJ*: *fsa1-A129S* target vectors with the two genes in operon configuration (**3–6**). The plasmid DNA of the clones OPE1 and OPE4 gave the expected DNA fragment pattern (**3–4**), whereas OPE5 could be considered false positive in the previous colony PCR (**5**) and OPE7 gave an unexpected DNA fragment pattern after digestion with *NcoI* (**6**).

Finally, the sequence of the plasmid from clone OPE1 was confirmed by Sanger sequencing. *E. coli* BL21(DE3) transformants were used for subsequent expression studies.

G VII.1.2 Construction of pKA1_alkJ::fsa1-A129S in pseudo-operon configuration (pPOP)

For the construction of the POP plasmid harboring the *alkJ* and the *fsa1-A129S* gene in pseudo-operon configuration, the *fsa1-A129S* gene was amplified by PCR using the *Pfu*⁺ DNA polymerase and the pET16b_ *fsa1-A129S* template with the following primer pair:

PO1A129SPROM *fwd*: 5'-GTAAGGATCCGGCTGCTAACCGATCCCGCGAAATTAATAC-3'

OPE1A129S *rev*: 5'-CAGCAGCCAACCTCAGCTTTAAATCGACGTTCTGCCAAAC-3'

The pKA1 backbone including the *alkJ* gene was amplified using Opti*Taq* and the previously assembled pKA1_ *alkJ* plasmid. The following primer pairs were used:

PO1pAJ3A129S *fwd*: 5'-GTATTAATTCGCGGGATCGGTTAGCAGCCGGATCCTTAC-3'

OPE1pAJ3A129S *rev*: 5'-GGCAGAACGTCGATTTAAAGCTGAGTTGGCTGCTGCCAC-3'

The optimal annealing temperatures for both insert and backbone amplification were determined by gradient PCR as outlined in G I.13. The gradient PCR was performed at 40–52°C, giving the optimized annealing temperature of 40.0°C (**Figure G-33**).^[348]

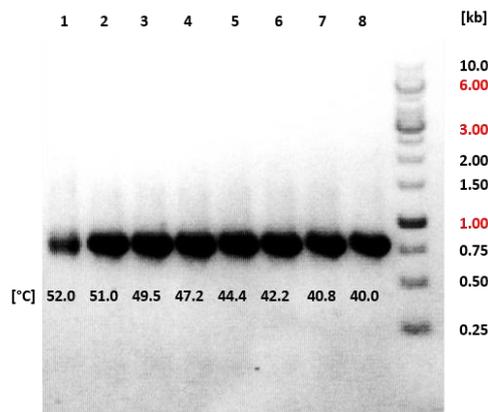


Figure G-33. Optimized annealing temperature of the *fsa1-A129S* insert for POP plasmid construction. Gradient PCR with annealing temperatures from 40–52°; 0.85% (w/v) agarose. The figure was adapted from T. Bayer *et al.* (2017).^[348]

For the backbone amplification, gradient PCR (40–52°C) yielded an optimized annealing temperature of 51.0°C (**Figure G-34**).

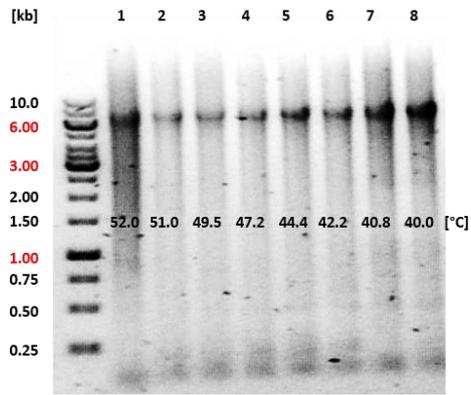


Figure G-34. Optimized annealing temperature of the pKA1_alkJ backbone for POP plasmid construction. Gradient PCR with annealing temperatures from 50–70°C; 0.85% (w/v) agarose. The figure was adapted from T. Bayer *et al.* (2017).^[348]

The preparation of each PCR mixture and the optimized thermal cycle conditions are summarized in **Table G-51** and **Table G-52**, respectively.^[348]

Table G-51. PCR reaction mixtures for the assembly of the POP plasmid

<i>Pfu</i> ⁺ PCR mix		Final concentration	<i>OptiTaq</i> PCR mix		Final concentration
5.0 µL	10X <i>Pfu</i> buffer	1X	5.0 µL	10X Pol B buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	PO1A129SPROM <i>fwd</i> (5 µM)	0.25 µM	2.5 µL	PO1pAJ3A129S <i>fwd</i> (5 µM)	0.25 µM
2.5 µL	OPE1A129S <i>rev</i> (5 µM)	0.25 µM	2.5 µL	OPE1pAJ3A129S <i>rev</i> (5 µM)	0.25 µM
1.0 µL	pET16b_ <i>fsa1-A129S</i> (100 ng·µL ⁻¹)	2.0 ng·µL ⁻¹	1.0 µL	pKA1_ <i>alkJ</i> (128 ng·µL ⁻¹)	2.5 ng·µL ⁻¹
1.0 µL	DMSO	2% (v/v)	1.0 µL	DMSO	2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U	0.5 µL	<i>OptiTaq</i> polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-	35.5 µL	nuclease-free water	-

Table G-52. Optimized thermal cycle conditions for the assembly of the POP plasmid

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (<i>OptiTaq</i>)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	4 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	40.0	30 s	30	Annealing	51.0	30 s	30
Extension	72	45 s		Extension	72	7.5 min	
Terminal extension	72	5 min	1	Terminal extension	72	10 min	1
Hold	4	∞	1	Hold	4	∞	1

G VII.1.2.1 Successful pPOP assembly by FC

The two DNA fragments were further processed and assembled as outlined in the general FC procedure (G II.2.1). Transformation of *E. coli* TOP10 cells was successful and six clones were randomly picked and colony PCR

performed according to G I.14 in 8 μ L PCR mixture. The presence of the *fsa1-A129S* insert with the expected size could be detected with the PO1A129SPROM *fwd* and OPE1A129S *rev* primer pair in four of the tested putative positive clones (**Figure G-35**, lane 3–6).^[348]

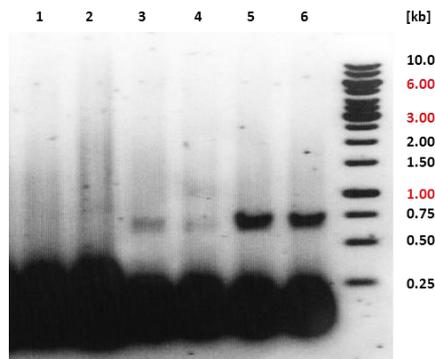


Figure G-35. Colony PCR of putative positives containing the POP plasmid assembled by FC. Detection of the the *fsa1-A129S* insert after colony PCR in four out of six clones (**3–6**); 1% (ω/v) agarose.

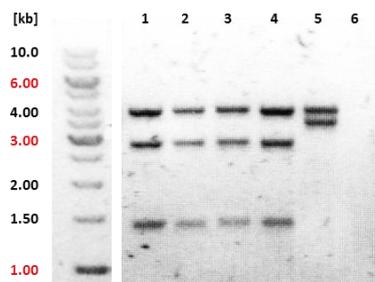


Figure G-36. Control digestion with *NcoI* of putative POP plasmids assembled by FC. Digestion mixtures containing the plasmid DNA of putative positive clones POP3–6 (**1–4**), the parent pKA1_alkJ plasmid (**5**), and no plasmid DNA (**6**). All four putative pKA1_alkJ:*fsa1-A129S* target vectors with the two genes in pseudo-operon configuration (**1–4**) gave the expected DNA fragment pattern analyzed on 0.85% (ω/v) agarose. The figure was adapted from T. Bayer *et al.* (2017).^[348]

The *E. coli* TOP10 clones positively tested for the insertion of the aldolase gene (**Figure G-35**, lane 3–6) were cultivated in LB-Miller medium supplemented with Cam overnight (37°C, 200 rpm), the plasmid DNA isolated and administered to control digestion by *NcoI* (**Figure G-36**, lane 1–4). The plasmid DNA of clone POP5 (**Figure G-35**, lane 5 for colony PCR and **Figure G-36**, lane 3 for *NcoI* control digestion) was finally confirmed by Sanger sequencing.^[348]

E. coli BL21(DE3) cells were transformed with pPOP (clone: POP5) for subsequent expression studies (see G VII.1.4.1).

G VII.1.2.2 Successful POP plasmid assembly from two DNA fragments by SLiCE

The following SLiCE reaction served as a proof of concept study for the adaption of the SLiCE protocol published by Zhang and coworkers.^[97] The linear DNA fragments containing the *alkJ* gene together with the pKA1 backbone and the *fsa1-A129S* gene, respectively, were amplified by PCR as described in G VII.1.2. Template DNA was digested by *DpnI* and purified as before. Plasmid assembly followed the SLiCE protocol as outlined in G II.2.2.4. The SLiCE reaction mixtures were prepared as follows:

G VII.1.3.2 Introduction of terminator sequences by Florida cloning utilizing a unique *Bam*HI restriction site
The B0011 single terminator was amplified using the DNA polymerase *Pfu*⁺ and the pUC57_B0014 template with the following primer pair:

B0011 *fwd*: 5'-CCGGGATCCAGAGAATATAAAAAGCC-3'

B0011 *rev*: 5'-CGCGGATCCAAATAATAAAAAAGCCGG-3'

The B0014 double terminator was amplified with *Pfu*⁺ from the pUC57_B0014 template with the following primers:

B0014 *fwd*: 5'-CCGGGATCCGGCTGCTAACTC-3'

B0014 *rev*: 5'-GGCGGATCCGGCTGCTAACAAATAATAAAAAAG-3'

The preparation of each PCR mixture and the thermal cycle conditions are summarized in **Table G-54** and **Table G-55**, respectively.

Table G-54. PCR reaction mixture for B0011 and B0014 amplification

	B001x PCR mix	Final concentration
5.0 µL	10X <i>Pfu</i> ⁺ buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	B001x <i>fwd</i> (5 µM)	0.25 µM
2.5 µL	B001x <i>rev</i> (5 µM)	0.25 µM
1.0 µL	pUC57_B0014 (100 ng·µL ⁻¹)	2.0 ng·µL ⁻¹
1.0 µL	DMSO	2.0% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-

Table G-55. Optimized thermal cycle conditions for B0011 and B0014 amplification

PCR step (B0011: <i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (B0014: <i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	57.5	20 s	30	Annealing	57.5	20 s	30
Extension	72	10 s		Extension	72	16 s	
Terminal extension	72	1 min	1	Terminal extension	72	1 min	1
Hold	12	∞	1	Hold	12	∞	1

Construction of pMON1 and pMON4 containing the B0011 and the B0014 terminator, respectively, followed the Florida cloning procedure as outlined in G II.1. Briefly, PCR was performed at optimal annealing temperature ($T_a = 57.5^\circ\text{C}$) as determined by gradient PCR (**Figure G-38**) and PCR products purified with the innuPREP PCRpure Kit from Analytik Jena.

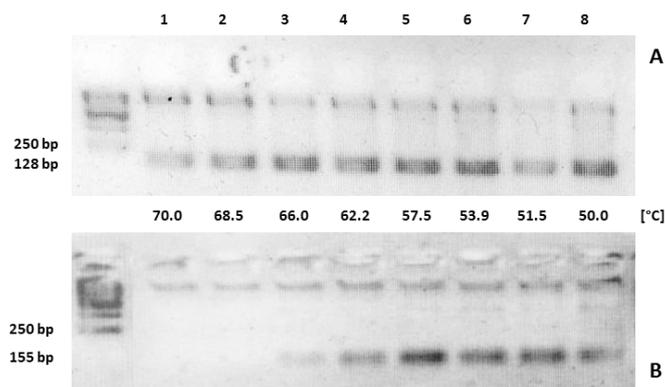


Figure G-38. Gradient PCR of B0011 and B0014. Determination of the optimal annealing temperature for the amplification of (A) the B0011 and (B) the B0014 terminator sequences analyzed on 4% and 3% (w/v) agarose, respectively. Temperature gradient ranging 50–70°C as indicated above. (A) was adapted from T. Wiesinger *et al.* (2017).^[435]

The POP plasmid was isolated from a single colony of an *E. coli* DH5 α transformant using the GeneJET Plasmid Miniprep Kit from Thermo Scientific. The vector (3 μ g) was linearized with *Bam*HI (Promega) in a total reaction volume of 40 μ L at 37°C for 2.5 h. The restriction enzyme was inactivated at 65°C for 20 min. The 5' overhangs of the linearized vector were dephosphorylated by calf intestinal alkaline phosphatase (CIAP; Promega). Therefore, 5 μ L of 10X CIAP reaction buffer (50 mM Tris-HCl, 1 mM MgCl₂, 0.1 mM ZnCl₂, 1 mM spermidine; pH 9.3) and nuclease-free water (4 μ L) were added directly to the digestion mixture, before adding 1 μ L CIAP (1 U $\cdot\mu$ L⁻¹). The resulting mixture was incubated at 37°C for 30 min. Again, 1 μ L CIAP was added and the mixture incubated at 37°C for 20 min and heated up to 56°C for 10 min. CIAP stop buffer (300 μ L; pH 7.5; 10 mM Tris-HCl, 1 mM EDTA, 200 mM NaCl, 0.5% (w/v) SDS) was added. Phenol-chloroform extraction was performed as suggest by Promega. DNA precipitation was performed by adding 0.1 volume of 3 M NaOAc (pH 5.2) and 2.5 volumes of ice-cold absolute ethanol (EtOH) and incubation at -20°C overnight. DNA pelleting was done by centrifugation (17 000 \times g, 4°C for 10–15 min). The pellet was washed with 1 mL ice-cold 70% (v/v) EtOH and dried completely by evaporation at 30°C, <30 mbar for 30 min (Christ Speedvac RVC 2-25 CD plus and Alpha 2-4 LD plus). The linearized and dephosphorylated vector DNA was dissolved in 30 μ L nuclease-free water for subsequent ligation.

Prior to ligation, the PCR products containing the B0011 or the B0014 sequence were trimmed with *Bam*HI under the same conditions as the target vector. The restriction enzyme was heat-inactivated, the DNA column-purified and eluted with 30 μ L nuclease-free water. For ligation, vector and insert DNA were mixed in volumetric ratios of 1:1 and 1:4. The fast ligation mix was prepared as given in **Table G-56**, spun down and incubated at 25°C for 20 min. For the transformation of RbCl-competent *E. coli* BL21(DE3) cells, 5 μ L of ligation mix were used. Transformation was done as described in G 1.9.3. Cells were plated on LB agar plates supplemented with Cam (37 μ g \cdot mL⁻¹). The plasmid DNA of single colonies was isolated as before. To test for the insertion of terminator sequences, plasmid DNA was digested with *Nco*I (Thermo Scientific) at 37°C for 2 h. *Nco*I was heat-inactivated at 65°C for 20 min and the DNA fragments analyzed by DNA gel electrophoresis on 1% (w/v) agarose (**Figure G-39**).

Table G-56. Ligation mixture for the construction of MON plasmids

	T4 DNA ligase mix	Final concentration
2.0 μ L	10X T4 DNA ligase buffer	1X
1.0 μ L	linearized vector DNA (20–100 ng $\cdot\mu$ L ⁻¹)	1–5 ng $\cdot\mu$ L ⁻¹
1 / 4.0 μ L	trimmed insert DNA	varying
0.5 μ L	T4 DNA ligase (5 Weiss U $\cdot\mu$ L ⁻¹)	2.5 Weiss U
15.5 / 12.5 μ L	nuclease-free water	-

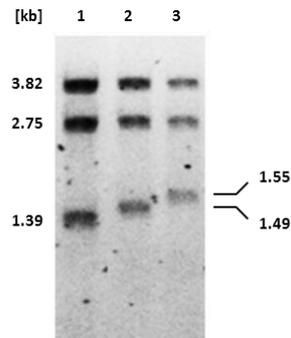


Figure G-39. *NcoI* digestion of MON1 and MON4 plasmids. Digestion of the parent POP plasmid (1), the MON1 (2), and the MON4 (3) plasmids containing the single terminator B0011 and the double terminator B0014, respectively. The figure was adapted from T. Wiesinger *et al.* (2017).^[435]

G VII.1.4 Validation of vector functionality by protein production analysis and biotransformations

For subsequent studies, *E. coli* BL21(DE3) transformants containing pOPE, pPOP, pMON1, or pMON4 were used (**Figure G-40C–F**). For comparison, untransformed *E. coli* BL21(DE3) and cells harboring the two-plasmid system (pKA1_alkJ and pET16b_fsa1-A129S) were used (**Figure G-40A** and **Figure G-40B**, respectively).

G VII.1.4.1 Coexpression studies of *AlkJ* and *Fsa1-A129S*

Coproduction of the mini-pathway was initially performed in M9-N* medium supplemented with Cam. Precultures were prepared in M9-N* medium as well and the main culture inoculated with 1% (v/v) preculture. Cultivation was performed at 37°C, 200 rpm until an OD₅₉₀ of 0.5 was reached. Protein production was induced by adding 0.5 mM IPTG, the temperature lowered to 25°C, and coexpression of *AlkJ* and *Fsa1-A129S* performed for 21 h. CFEs were prepared by sonication as in G I.5. SDS-PAGE analysis confirmed the coproduction of both enzymes in strains containing the single plasmids or the two-plasmid system (G I.8; **Figure G-40**).

Mini-pathway expression was successfully adapted for autoinduction in LB-5052 (AIM) as described in G VI.3.1.1 (data not shown).

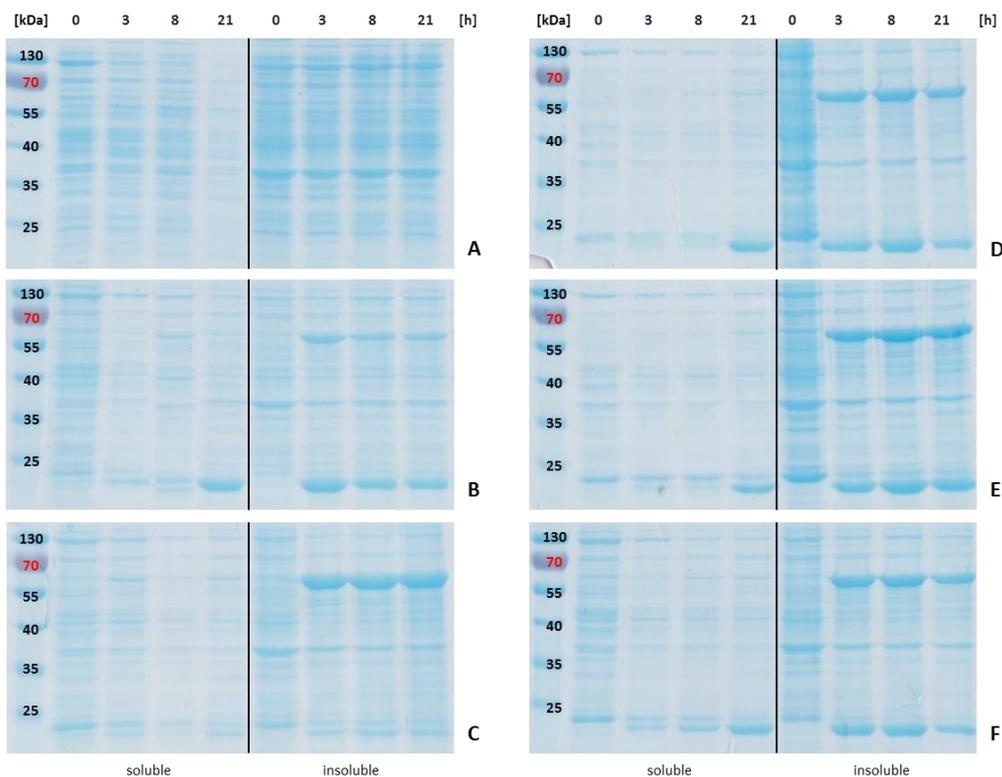


Figure G-40. Mini-pathway expression from different constructs over time. SDS-PAGE analysis of soluble and insoluble fractions from (A) untransformed *E. coli* BL21(DE3), (B) cells harboring the two-plasmid system (pKA1_alkJ + pET16b_fsa1-A129S), (C) pOPE, (D) pPOP, (E) pMON1, and (F) pMON4. *AlkJ* and *Fsa1-A129S* visible as protein bands at 57 kDa and 23 kDa, respectively. Sample loading normalized to OD₅₉₀ = 7.0.

G VII.1.4.2 Pathway validation with pPOP and DHA concentration screening

For pathway validation, *E. coli* BL21(DE3) harboring pPOP (clone: POP5) was used. Prior to screening, DHA dimer (0.9 g, 5 mmol) was dissolved in 9 mL RCM and filled up to 10 mL. For monomerization, the resulting solution was shaken (250 rpm; InforsHT Multitron 2 Standard) at 37°C for 3 h. The final concentration was 1 M of the DHA monomer.^[348]

The mini-pathway was expressed as described above (Figure G-40D), RCs prepared, and the initial DHA screening performed under standard conditions as described in G III.2 with the substrate **2b**. Biotransformations were carried out in the absence of DHA (0 eq) and in the presence of DHA monomer (5 eq, 10 eq, and 20 eq). Results of the HPLC measurements are shown in Figure D-29 (see A I for pathway schemes).

G VII.1.5 Changing context of pMON4: Increasing the spacer sequences in pMON5 and pMON6

According to SDS-PAGE, *AlkJ* and *Fsa1-A129S* were readily produced in *E. coli* BL21(DE3) harboring pMON1 or the pMON4 (Figure D-28 and Figure G-40E–F). However, functional testing of *AlkJ* under standard screening conditions (G III.2) revealed that the substrates **2b** and **4b** were not or hardly oxidized to **2c** and **4c**, respectively (Figure E-3B–C).

In pMON1 and pMON4, the terminator sequences B0011 and B0014 were placed between the coding regions of the ADH and the aldolase, respectively. Cloning of B0011 in the pMON1 (pKA1_alkJ::B0011::fsa1-A129S)^[435] resulted in a 16 bp spacer sequence; cloning of B0014 in the pMON4 (pKA1_alkJ::B0014::fsa1-A129S) also contained 16 bp flanking the terminator sequence (Figure D-25C). The detailed cloning procedure is described in G VII.1.3.2.

To increase the spacing between the gene-coding regions and the synthetic B0014 terminator sequence, PCR primer pairs were designed to insert T_{syn} with flanking spacers of 50 bp and 100 bp, giving rise to pMON5 (pKA1_alkJ::B0014₅₀::f_{sa1}-A129S) and pMON6 (pKA1_alkJ::B0014₁₀₀::f_{sa1}-A129S), respectively (**Figure E-1**). Primers also encoded *BglIII* restriction sites, whereas the target pPOP vector was digested with *BamHI* to insert the B0014 variants. By utilization of the *BamHI/BglIII* isocaudamer, ligation of target vector and B0014 insert resulted in a scar sequence that could not be cleaved by either of the original restriction enzymes.^[6] Subcloning was performed by C. Wokurek as described below.^[519]

G VII.1.5.1 Construction of pMON5 and pMON6

For the construction of pMON5, the B0014 double terminator was amplified with *Pfu*⁺ from the pUC57_B0014 template with the following primers:

MON5_BglIII fwd: 5'-GCGAGATCTCAGTGAATTCGAGCTCGGTAC-3'

MON5_BglIII rev: 5'-GCGAGATCTGATTACGCCAAGCTTGCATGCAG-3'

For the construction of pMON6, the B0014 double terminator was amplified with *Pfu*⁺ from the pUC57_B0014 template with the following primers:

MON6_BglIII fwd: 5'-GCGAGATCTGATTAAGTTGGGTAACGCCAG-3'

MON6_BglIII rev: 5'-GCGAGATCTGAATTGTGAGCGGATAACAATTC-3'

The preparation of each PCR mixture is summarized in **Table G-57**. The thermal cycle conditions were identical to the construction of pMON1 and pMON4 (**Table G-55**) except longer extension times for insert amplification:

- 19 s for B0014₅₀ (pMON5)
- 22 s for B0014₁₀₀ (pMON6)

Table G-57. PCR reaction mixture for B0014₅₀ and B0014₁₀₀ amplification

	MONx PCR mix	Final concentration
5.0 µL	10X <i>Pfu</i> ⁺ buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	MONx_BglIII fwd (5 µM)	0.25 µM
2.5 µL	MONx_BglIII rev (5 µM)	0.25 µM
1.0 µL	pUC57_B0014 (100 ng·µL ⁻¹)	2.0 ng·µL ⁻¹
1.0 µL	DMSO	2.0% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-

Assembly of pMON5 and pMON6 was in accordance to the construction of the other plasmids containing the *alkJ* and *f_{sa1}-a129S* gene in monocistronic configurations (see G VII.1.3.2) and followed the Florida cloning procedure as outlined in G II.1. PCRs at $T_a = 57.5^\circ\text{C}$ yielded single amplification products, which were purified with the innuPREP PCRpure Kit from Analytik Jena and analyzed on 1.5% (w/v) agarose (**Figure G-41**).

Ligation was performed as in **Table G-56** at 25°C for 20 min. After transformation of RbCl-competent *E. coli* BL21(DE3) (G I.9.3), cells were plated on LB agar plates supplemented with Cam (37 µg·mL⁻¹). Reisolated plasmid DNA of single colonies was digested with *NcoI* (Thermo Scientific) at 37°C for 2 h to confirm the insertion of the terminator sequences (**Figure G-42**).

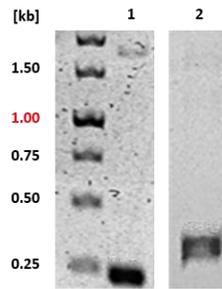


Figure G-41. PCR products containing *B0014₅₀* and *B0014₁₀₀*. Purified amplicons containing the B0014 terminator with flanking 50 bp spacers (1) and 100 bp spacers (2) and BglIII restriction sites; 1.5% (ω/v) agarose. The figure was adapted from C. Wokurek *et al.* (2017).^[519]

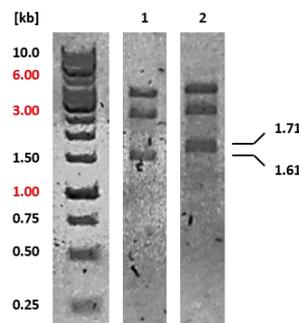


Figure G-42. *NcoI* digestion of MON5 and MON6 plasmids. Digestion of the MON1 (1), and the MON4 (2) plasmids containing the B0014 terminator with flanking 50 bp and 100 bp spacer sequences, respectively. The figure was adapted from C. Wokurek (2017).^[519]

G VII.1.5.2 Testing for the restoration of *AlkJ* activity in biotransformations

Coproduction of *AlkJ* and *Fsa1-A129S* from the newly assembled constructs pMON5 and pMON6 was performed in AIM as described in G VI.3.1.1 except that expression was done at 25°C. SDS-PAGE analysis confirmed the production of *AlkJ* and *Fsa1-A129S* (Figure E-2). RCs were prepared and the substrates 2b and 4b tested under standard screening conditions (G III.2). *AlkJ* activity could be restored and results are described in E I.1 (Figure E-3E and Figure E-3F).

G VII.2 Vector construction for the coexpression of *AlkJ* and the DHAP-dependent aldolase *FucA*

In the following, the detailed cloning procedure for vectors for the coproduction of *AlkJ* and *FucA* in different genetic configurations will be described.

G VII.2.1 Construction of pKA1_alkJ::fucA in pseudo-operon configuration

For the construction of the target plasmid harboring the *alkJ* and the *fucA* gene in pseudo-operon configuration, the *fucA* insert was PCR amplified using the *Pfu*⁺ DNA polymerase and the pKK_ *fucA* template^[486] with the following primer pair:

POPFUC1A *fwd*: 5'-GGAGATATACCATGGAACATGGAACGAAATAAACTTG-3'

POPFUC1A *rev*: 5'-CAGCAGCCAACCTCAGCTTTACTCTTCAATTCGTAAC-3'

The pKA1 backbone including the *alkJ* gene and the regulatory elements (P_{T7} and RBS) for *fucA* expression were amplified using *OptiTa*q and the previously assembled pPOP plasmid. The following primer pairs were used:

pPOPFUC1A *fwd*: 5'-GTTTATTCGTTCCATGTTCCATGGTATATCTCCTTC-3'
 pPOPFUC1A *rev*: 5'-CGAATTGAAGAGTAATAAAGCTGAGTTGGCTGCTGCCAC-3'

The optimal annealing temperatures for both insert and backbone amplification were determined by gradient PCR as outlined in G I.13. For the *fucA* insert amplification, the gradient PCR (45–65°C) resulted in a single amplification product of expected size (0.65 kb) over the whole temperature gradient (**Figure G-43A**, result shown for $T_a = 56.4^\circ\text{C}$ after PCR product purification). Backbone amplification yielded a single PCR product of expected size (7.20 kb; **Figure G-43B**).

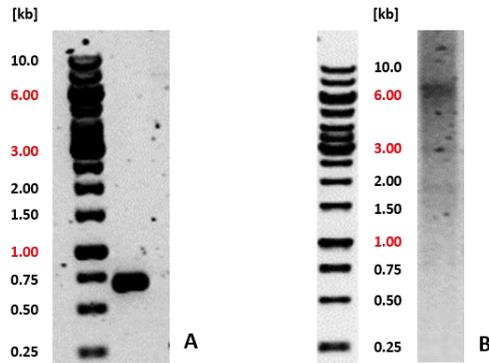


Figure G-43. Optimized annealing temperatures for the construction of pKA1_alkI::fucA. (A) Optimized annealing temperature (54.6°C) of the *fucA* insert. (B) Gradient PCR (45–65°C) of the pKA1 backbone including the *alkI* gene. (C) Optimized PCR cycle conditions with annealing at 48.8°C for 25 s; 0.85% (w/v) agarose. The figure was adapted from J. Reiterlehner (2017).^[3]

The preparation of each PCR mixture and the optimized thermal cycle conditions are summarized in **Table G-58** and **Table G-59**, respectively.

Table G-58. PCR reaction mixtures for the assembly of pKA1_alkI::fucA

<i>Pfu</i> ⁺ PCR mix		Final concentration	OptiTaq PCR mix		Final concentration
5.0 μL	10X <i>Pfu</i> buffer	1X	5.0 μL	10X Pol B buffer	1X
2.0 μL	dNTP mix (5 mM each)	0.2 mM each dNTP	2.0 μL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 μL	POPFUC1A <i>fwd</i> (5 μM)	0.25 μM	2.5 μL	pPOPFUC1A <i>fwd</i> (5 μM)	0.25 μM
2.5 μL	POPFUC1A <i>rev</i> (5 μM)	0.25 μM	2.5 μL	pPOPFUC1A <i>rev</i> (5 μM)	0.25 μM
1.0 μL	pKK_ <i>fucA</i> (60 ng·μL ⁻¹)	1.2 ng·μL ⁻¹	1.0 μL	pPOP (100 ng·μL ⁻¹)	2.0 ng·μL ⁻¹
1.0 μL	DMSO	2% (v/v)	1.0 μL	DMSO	2% (v/v)
0.5 μL	<i>Pfu</i> ⁺ polymerase (5 U·μL ⁻¹)	2.5 U	0.5 μL	OptiTaq polymerase (5 U·μL ⁻¹)	2.5 U
35.5 μL	nuclease-free water	-	35.5 μL	nuclease-free water	-

Table G-59. Optimized thermal cycle conditions for the assembly of pKA1_alkJ::fucA

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (<i>Opti</i> Taq)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	54.6	30 s	30	Annealing	48.8	25 s	30
Extension	72	40 s		Extension	72	7.5 min	
Terminal extension	72	3 min	1	Terminal extension	72	10 min	1
Hold	12	∞	1	Hold	12	∞	1

Assembly was performed by FC according to G II.2.1.3. The plasmid DNA of five putative positive clones was isolated as in G I.10. Subsequently, plasmid DNA was subjected to *Bgl*II (R0143S, NEB) control digestion according to G I.11. Plasmid DNA from all five clones yielded the expected restriction pattern (**Figure D-30A**). The insertion of the *fucA* gene was confirmed by Sanger sequencing (clone #1).

Table G-60. *Bgl*II digestion mix

	<i>Bgl</i> II digestion mix	Final concentration
1.0 µL	10X NEBuffer™ 3.1	1X
2.5 µL	DNA (80–100 ng·µL ⁻¹)	20–25 ng·µL ⁻¹
1.0 µL	<i>Bgl</i> II (10 U·µL ⁻¹)	1 U·µL ⁻¹
5.5 µL	nuclease-free water	-

Expression studies were performed with *E. coli* BL21(DE3) transformants harboring pKA1_alkJ::fucA (clone #1). Cultivation and induction of enzyme production was performed in AIM as described in G VI.3.1.1. Coproduction of *AlkJ* and *FucA* was tested at 25°C with shaking (150 rpm) and analyzed by SDS-PAGE (G I.8; **Figure D-30B**).

G VII.2.2 Construction of pKA1_fucA::alkJ in pseudo-operon configuration

For the construction of the target plasmid harboring the *alkJ* and the *fucA* gene, the *alkJ* insert was PCR amplified using the *Pfu*⁺ DNA polymerase and the previously assembled pKA1_alkJ template^[348] with the following primer pair:

POPFUCALK *fwd*: 5'-CAAAGCCCGAAAGGATAATACGACTCACTATAGG-3'

POPFUCALK *rev*: 5'-CAGCCAACCTCAGCTTTACATGCAGACAGCTATCATG-3'

The pKA1 backbone including the *fucA* gene for *fucA* expression were amplified using *Opti*Taq and the previously assembled pKA1_fucA plasmid.^[509] The following primer pairs were used:

pPOPFUCALK *fwd*: 5'-CTGTCTGCATGTAAAGCTGAGTTGGCTGCTGCCAC-3'

pPOPFUCALK *rev*: 5'-GTGAGTCGTATTATCCTTCGGGCTTTGTTAG-3'

The optimal annealing temperatures for both insert and backbone amplification were determined by gradient PCR as outlined in G I.13. For both the *alkJ* insert and the pKA1_fucA backbone amplification, gradient PCRs (45–65°C) resulted in single PCR products of expected sizes (1.7 kb and 6.2 kb; **Figure G-44A** and **Figure G-44B**,

respectively). The PCR mixtures and optimized thermal cycle conditions are given in **Table G-61** and **Table G-62**, respectively.

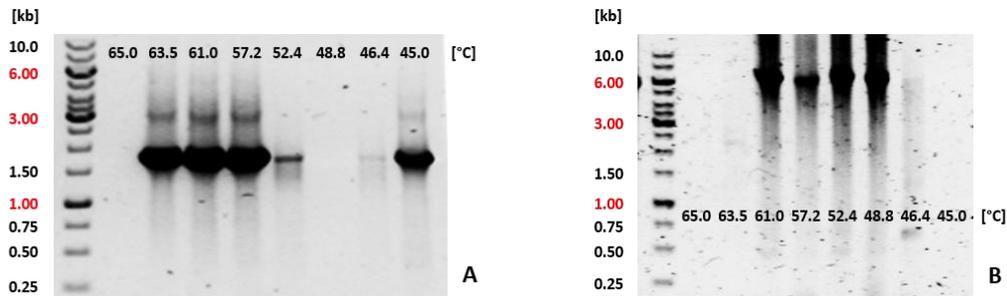


Figure G-44. Annealing temperature optimization for the construction of pKA1_fucA::alkJ. (A) Gradient PCR (45–65°C) of the *alkJ* insert resulting the optimal $T_a = 52.4^\circ\text{C}$. (B) Gradient PCR of the pKA1 backbone including the *fucA* gene giving a $T_a = 57.2^\circ\text{C}$. The figure was adapted from J. Reiterlehner (2017).^[3]

The preparation of each PCR mixture and the (optimized) thermal cycle conditions are summarized in **Table G-58** and **Table G-59**, respectively.

Table G-61. PCR reaction mixtures for the assembly of pKA1_fucA::alkJ

<i>Pfu</i> ⁺ PCR mix		Final concentration	OptiTaq PCR mix		Final concentration
5.0 µL	10X <i>Pfu</i> buffer	1X	5.0 µL	10X Pol B buffer	1X
2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP	2.0 µL	dNTP mix (5 mM each)	0.2 mM each dNTP
2.5 µL	POPFUCALK <i>fwd</i> (5 µM)	0.25 µM	2.5 µL	pPOPFUCALK <i>fwd</i> (5 µM)	0.25 µM
2.5 µL	POPFUCALK <i>rev</i> (5 µM)	0.25 µM	2.5 µL	pPOPFUCALK <i>rev</i> (5 µM)	0.25 µM
1.0 µL	pKA_alkJ (128 ng·µL ⁻¹)	2.5 ng·µL ⁻¹	1.0 µL	pKA1_fucA (100 ng·µL ⁻¹)	2.0 ng·µL ⁻¹
1.0 µL	DMSO	2% (v/v)	1.0 µL	DMSO	2% (v/v)
0.5 µL	<i>Pfu</i> ⁺ polymerase (5 U·µL ⁻¹)	2.5 U	0.5 µL	OptiTaq polymerase (5 U·µL ⁻¹)	2.5 U
35.5 µL	nuclease-free water	-	35.5 µL	nuclease-free water	-

Table G-62. Optimized thermal cycle conditions for the assembly of pKA1_fucA::alkJ

PCR step (<i>Pfu</i> ⁺)	Temperature [°C]	Time	No. of cycles	PCR step (OptiTaq)	Temperature [°C]	Time	No. of cycles
Initial denaturation	95	5 min	1	Initial denaturation	95	5 min	1
Denaturation	95	30 s		Denaturation	95	30 s	
Annealing	54.6	30 s	30	Annealing	48.8	30 s	30
Extension	72	90 s		Extension	72	4.75 min	
Terminal extension	72	5 min	1	Terminal extension	72	7 min	1
Hold	12	∞	1	Hold	12	∞	1

Assembly was performed by FC according to G II.2.1.3. The plasmid DNA of six putative positive clones was isolated as in G I.10. Subsequently, plasmid DNA was subjected to *NcoII* (R0193, NEB) control digestion according

to G I.11. Plasmid DNA from all six clones yielded the expected restriction pattern (**Figure D-33A**). The insertion of the *alkJ* gene was confirmed by Sanger sequencing (clone: JRE2-1).

Expression studies were performed with *E. coli* BL21(DE3) transformants harboring pKA1_ *fucA::alkJ* (clone: JRE2-1). Cultivation and induction of enzyme production was performed in AIM as described in G VI.3.1.1. Coproduction of *AlkJ* and *FucA* was tested at 25°C with shaking (150 rpm) and analyzed by SDS-PAGE (G I.8; **Figure D-33B**).

G VIII Expression of multiple pathway enzymes in a single host cell

G VIII.1 Coproduction of *AlkJ*, *Fsa1-A129S*, *CAR_{Ni}*, and *PPtase_{Ec}*

For coproduction of the target enzymes, competent *E. coli* BL21-Gold(DE3) cells harboring pETDuet-1_ *pptase_{Ec}::car_{Ni}* were transformend with pKA1_ *alkJ::fsa1-A129S* (pPOP) as described in G I.9.3. Optimized coexpression of *AlkJ*, *CAR_{Ni}*, *PPtase_{Ec}*, and *Fsa1-A129S* could be achieved by autoinduction as described in G VI.3.1.1 and confirmed by SDS-PAGE analysis (**Figure E-15A**). The engineered strain was referred to as CRAS (containment of reactive aldehyde species) strain. All enzymes were produced at increased yields compared to systems cultivated in TB medium supplemented with Amp + Cam and induced by IPTG (0.5 mM IPTG at OD₅₉₀ = 0.5, subsequent expression at 20°C, 150 rpm for 20 h).

An example protein gel is shown in **Figure E-15A**.

The functionality of the CRAS strain was demonstrated employing RCs and standard conditions. The conversion of the model substrates **2b** and **2c** was monitored over time (0–24 h) in the absence of DHA (**Figure E-19A**) and the presence of DHA. The latter produced the target aldol **2e** (**Figure E-19B–C**).

G VIII.2 Coproduction of *AlkJ* and DHAP-dependent aldolases

For coproduction of the target enzymes, competent *E. coli* BL21(DE3) cells were cotransformed with pKA1_ *alkJ* and either pKK_ *fruA*, pKK_ *fucA*, or pKK_ *rhuA* as described in G I.9.3. The preculture was prepared in 4 mL LB-Miller supplemented with Cam + Amp and grown at 37°C, 200 rpm overnight. The main culture was inoculated with 1% (v/v) preculture (37°C, 200 rpm) and grown until OD₅₉₀ = 0.3 before 0.1 mM ZnCl₂ was added. Cells were further grown to an OD₅₉₀ = 0.5 at 30°C, 130 rpm. Protein coproduction was induced with 0.5 mM IPTG and performed at 25°C, 130 rpm for 21 h.

E. coli BL21 (DE3) harboring the single enzyme expression vectors (pKA1_ *alkJ*, pKK_ *fruA*, pKK_ *fucA*, or pKK_ *rhuA*) were cultivated accordingly as expression controls. Cells were harvested and lysed as in G I.5. SDS-PAGE analysis was performed as in G I.8 (**Figure G-45**).

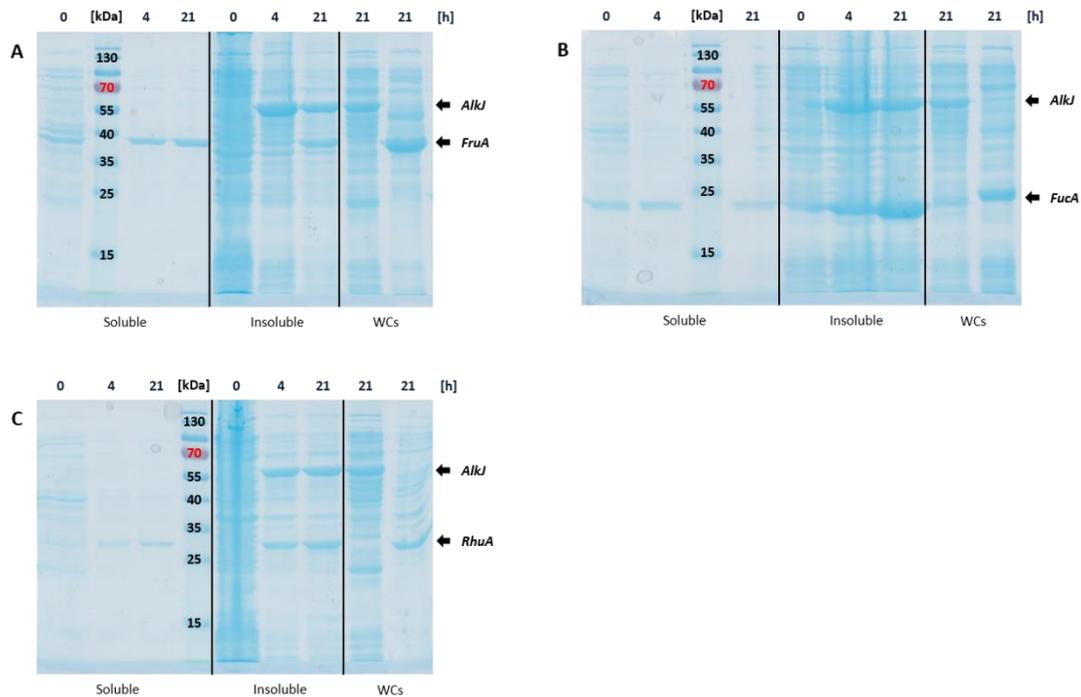


Figure G-45. Coproduction of *AlkJ* and DHAP-dependent aldolases in one cell. *AlkJ* in *E. coli* BL21(DE3) coexpressing (A) *FruA*, (B) *FucA*, and (C) *RhuA*. SDS-PAGE analysis of soluble and insoluble fractions collected over time (0–21 h) and single enzyme expression controls of whole cells (WCs) after 21 h expression. Sample loading normalized to $OD_{590} = 7.0$.

G VIII.3 Preparation of LEPs coexpressing *FucA* with other pathway enzymes

For coproduction of the target enzymes, competent *E. coli* BL21(DE3) cells were cotransformed with different combinations of plasmids as described in G I.9.3.

- pKA1_*fucA*^[509] and either pET26(+)_*phoN-Sf* or pCDFDuet-1_*yqaB* as described in G I.9.3. The resulting strains will be referred in this thesis as FucP (D III.2.2) and FucY (E I.6.2), respectively.
- pKA1_*fucA*^[509] and pRSETa_*dhaK* yielding the FucD strain, which could not be tested in this thesis since *DhaK* would not express anymore at later stages of this project (see E I.6.1).

The preculture was prepared in 4 mL LB-0.8G supplemented with Cam + Kan or Cam + Str and grown at 37°C, 275 rpm overnight. The main culture was inoculated with 1% (v/v) preculture and grown in the presence of 0.1 mM ZnCl₂ at 30°C, 200 rpm for approximately 8 h until an $OD_{590} > 1.5$ was reached. Protein production was induced with 1 mM IPTG and expression performed for another 12 h and not longer.

Example protein gels are shown in **Figure E-24**.

G VIII.4 Coproduction of *AlkJ*, *FucA*, and *PhoN-Sf*

Since coproduction of target enzymes was unsatisfying by autoinduction (G VI.3.1.1), *E. coli* BL21(DE3) transformants harboring pKA1_*fucA::alkJ*^[3] and pET26b(+)_*phoN-Sf* (referred to as AFucP strain) were cultivated in accordance to G VIII.3. Unfortunately, LEPs did not exhibit *AlkJ* activity experiments aiming at the validation of the DHAP-dependent aldolase pathway (see D III.2.2 and **Table D-6**).

G VIII.5 Coproduction of *FucA*, *YqaB*, and *DhaK*

E. coli BL21(DE3) transformants harboring pKA1_ *fucA*^[509], pCDFDuet-1_ *yqaB*, and pRSETa_ *dhaK* were prepared according to G I.9.3 and will be referred to as FYD strain in this thesis. Since the FYD strain harbors three plasmids with different resistance markers, antibiotics were employed at half of the standard working concentrations – Cam (18.5 $\mu\text{g}\cdot\text{mL}^{-1}$), Str (12.5 $\mu\text{g}\cdot\text{mL}^{-1}$), and Amp (50 $\mu\text{g}\cdot\text{mL}^{-1}$) – in subsequent fermentations. The expression medium from Wu *et al.* was adapted and is given in (Table G-63).^[43] Similar as in G VIII.3 and G VIII.4, expression conditions were adapted from the Wang group.^[293]

Briefly, a single colony of the recombinant FYD strain was inoculated in 12 mL FYD medium (Table G-63) supplemented with antibiotics and incubated in baffled shaking flasks (see G VI.6.2) at 37°C with vigorous shaking (350 rpm) for 8–10 h. Then, 1 mL of the preculture was transferred into 29 mL FYD medium in a 500 mL baffled shaking flask. The cells were grown at 37°C, 350 rpm for about 4 h to reach an $\text{OD}_{590} > 1.5$, followed by the addition of IPTG to 1 mM to induce the enzyme expression. The cells were grown for 12–15 h at 30°C to reach late exponential phase and used immediately. An example protein gel showing increased expression levels can be seen in Figure G-46, lane 2).

Table G-63. Constituents of FYD medium

FYD medium		10X M9 salts		100X Trace element solution	
100 mL	10X M9 salts	5.0 g	NaCl	0.18 g	ZnSO ₄ · 7 H ₂ O
10.0 g	yeast extract	75.0 g	Na ₂ HPO ₄ · 2 H ₂ O	0.12 g	CuCl ₂ · 2 H ₂ O
3.0 mL	1 M MgSO ₄	30.0 g	KH ₂ PO ₄	0.12 g	MnSO ₄ · 2 H ₂ O
1.0 mL	1 M CaCl ₂	24.0 g	(NH ₄) ₂ SO ₄	0.18 g	CoCl ₂ · 6 H ₂ O
0.6 mL	0.1 M FeCl ₃ · 6 H ₂ O			0.03 g	H ₃ BO ₃
2.0 mL	1 mM thiamine-HCl			0.025 g	Na ₂ MoO ₄ · 2 H ₂ O
10.0 mL	100X trace element solution			0.084 g	Na ₂ EDTA · 2 H ₂ O
100 mL	20% (w/v) glucose				

All stock solutions were prepared and sterilized separately. Trace element solution, 1 mM thiamine-HCl, 0.1 M FeCl₃ and 20% (w/v) glucose were sterilized by filtration (sterile syringe filter, 0.2 μm cellulose acetate). For preparation of FYD medium, 10.0 g yeast extract were weighed in, 100 mL 10X M9 salts added, filled up to 784 mL with dH₂O, and autoclaved. Subsequently, all other components were added under sterile conditions.

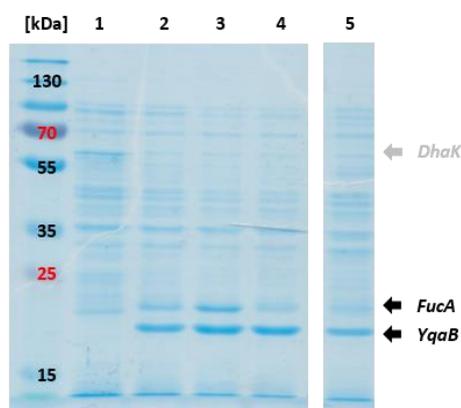


Figure G-46. Increased expression levels in the FYD medium. SDS-PAGE analysis of whole cell samples of untransformed *E. coli* BL21(DE3) (1) and the FYD strain cultivated in FYD medium containing 6.0 $\text{g}\cdot\text{L}^{-1}$ (2) or 10.0 $\text{g}\cdot\text{L}^{-1}$ yeast extract (3), cultivated in LB-Miller medium as in G VIII.3 (4), and in AIM as in G VI.3.1.1. *E. coli* BL21(DE3) treated as sample (2). Sample loading normalized to $\text{OD}_{590} = 3.5$. *DhaK* not expressing due to unknown reason (see E I.6.1).

G IX Downstream processing by SPE

To isolate polyhydroxylated compounds from biotransformation mixtures, RCs were centrifuged ($8\,000 \times g$, 4°C for 15 min), and the supernatant directly transferred by an external pump (BESTA) onto a flash cartridge (BUCHI Sepacore, 25 g, C18) with continuous flow of $10 \text{ mL}\cdot\text{min}^{-1}$ of 5% (v/v) MeOH (MeOH and water, HPLC grade). For elution of target aldol adducts, the mixture was pumped through a Grace REVELERIS X flash chromatography system with integrated ELS/UV-Vis detection. Pure products were manually collected with 95% (v/v) MeOH (Figure G-47).

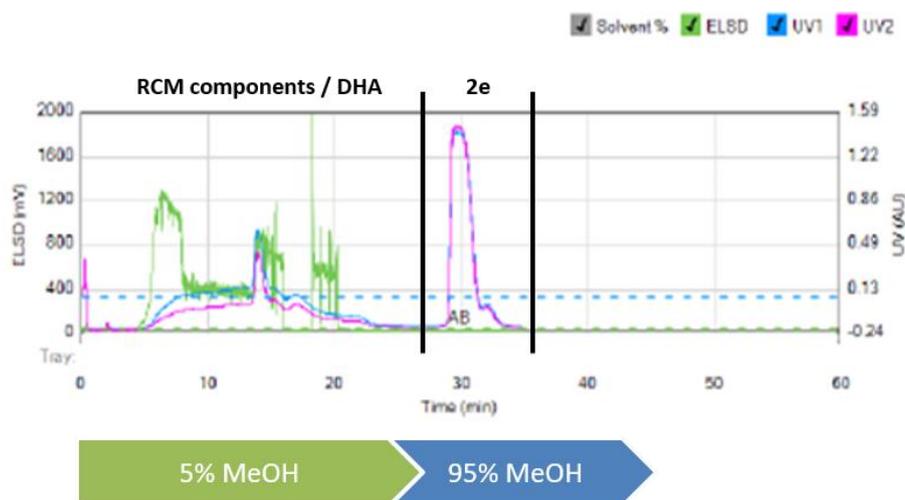


Figure G-47. SPE chromatogram from the purification of aldol **2e**. Recorded chromatogram of SPE of the target compound **2e**; ELSD (green), UV-Vis (blue: 200 nm; pink: 205 nm). The figure was adapted from T. Wiesinger *et al.* (2017).^[435]

H Appendix

H I Gene sequences and plasmid maps

In the following, gene sequences used in this work will be given with indications of start (light green) and stop codons (dark red) for ORFs. Miscellaneous features such as protein tags (dark yellow) or protease cleavage sites (orange) are indicated separately. Restriction sites used for cloning are underlined.

Vector maps were created with Geneious version 6.1 (Biomatters) available from <http://www.geneious.com>. Maps include the plasmid backbone (black), ORIs (light blue) and antibiotic markers (turquoise). For IPTG-induced gene expression, maps feature regulatory sequences (grey) such as *lac* operator (*lacO*) sequences and the *lacI* gene, which encodes the *lac* repressor protein. Promoters (e.g., T7 prom) and RBS (e.g., T7_transl_en_RBS) given in shades of green; transcriptional terminators (e.g., T7 term) in red. The restriction sites for enzymes commonly encountered in this thesis are included: *Bam*HI, *Bgl*II, *Hind*III, *Nco*I, *Nde*I, *Not*I, *Pst*I, *Xba*I, and *Xho*I. Restriction sites given as base number in brackets.

H I.1.1 Maps of empty vectors routinely used in this thesis

H I.1.1.1 pCDFDuet-1

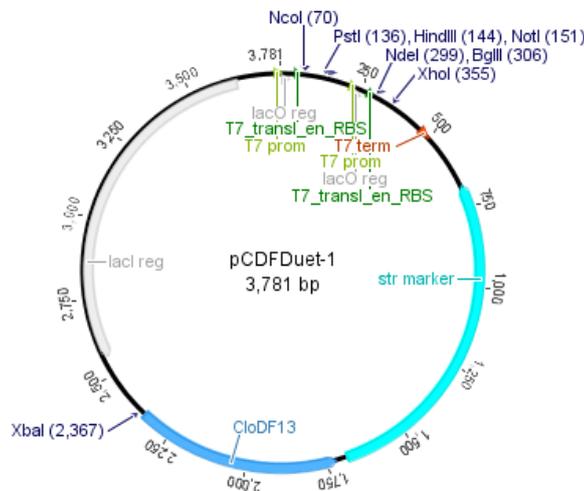


Figure H-1. pCDFDuet-1. The vector features a CloDF13 ORI (compatible with pBR322 and p15A ORIs), two MCS (1: *Nco*I; 2: *Nde*I), and conveys resistance to Str.

H I.1.1.2 pET16b

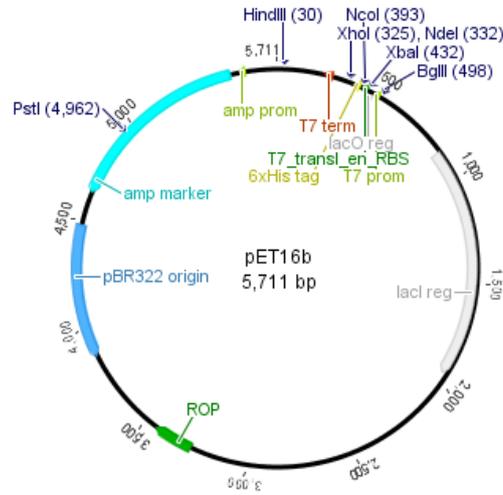


Figure H-2. pET16b. The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (untagged: *NcoI*; N-terminal 6xHis tag: *NdeI*), and conveys resistance to Amp.

H I.1.1.3 pET21b(+)

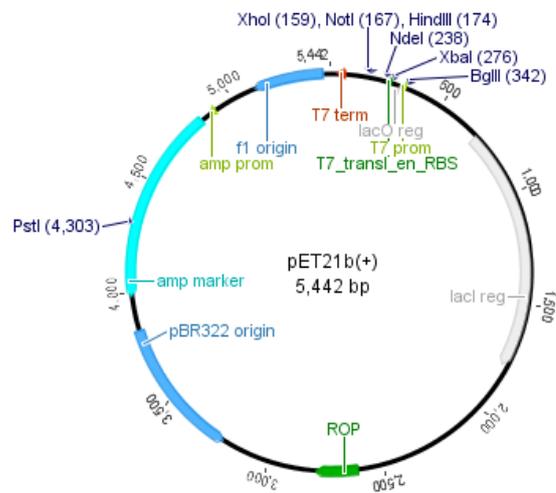


Figure H-3. pET21b(+). The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (*NdeI*), and conveys resistance to Amp.

H I.1.1.4 pET22b(+)

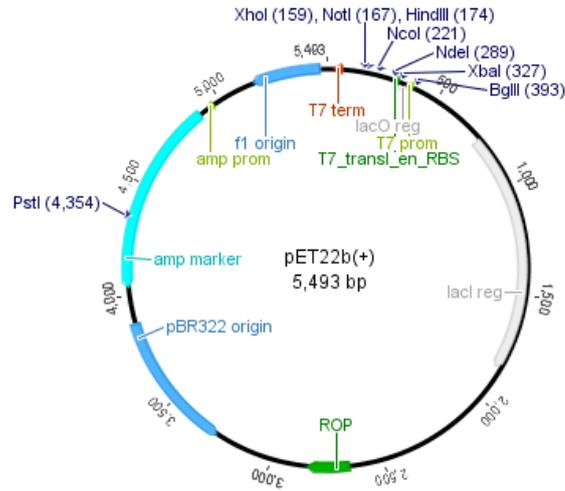


Figure H-4. pET22b(+). The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (*NdeI*), and conveys resistance to Amp.

H I.1.1.5 pET26b(+)

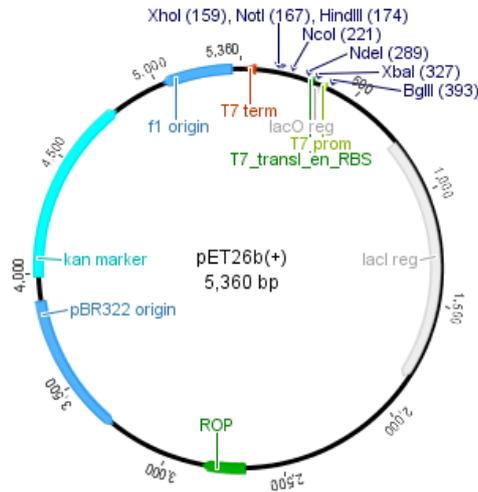


Figure H-5. pET26b(+). The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (*NdeI*), and conveys resistance to Kan.

H I.1.1.6 pET28a(+)

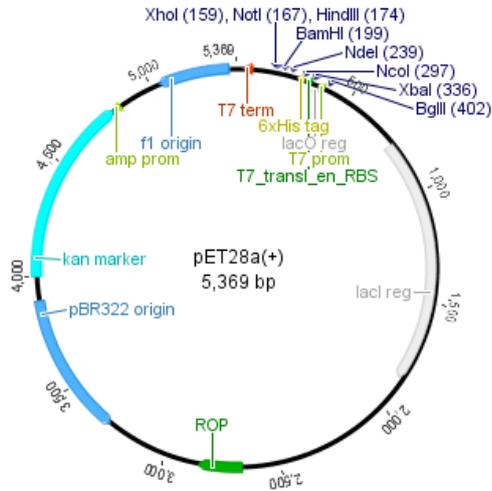


Figure H-6. pET28a(+). The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (untagged: *NcoI*; N-terminal 6xHis tag: *NdeI*), and conveys resistance to Kan.

H I.1.1.7 pETDuet-1

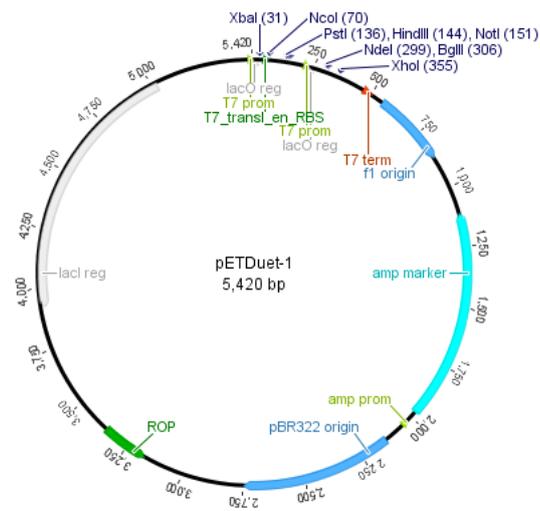


Figure H-7. pETDuet-1. The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), two MCS (1: *NcoI*; 2: *NdeI*), and conveys resistance to Amp.

H I.1.1.8 pKA1

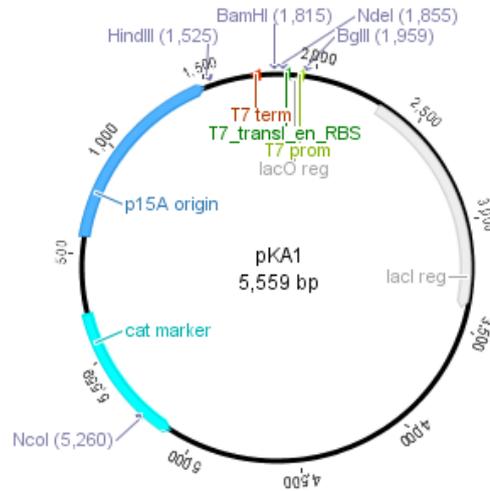


Figure H-8. pKA1. The vector is a fusion vector of pET11a and pACYC. It features a p15A ORI (compatible with pBR322 and CloDF13 ORIs), one MCS (*NdeI*), and conveys resistance to Cam.

H I.1.1.9 pKK223-3

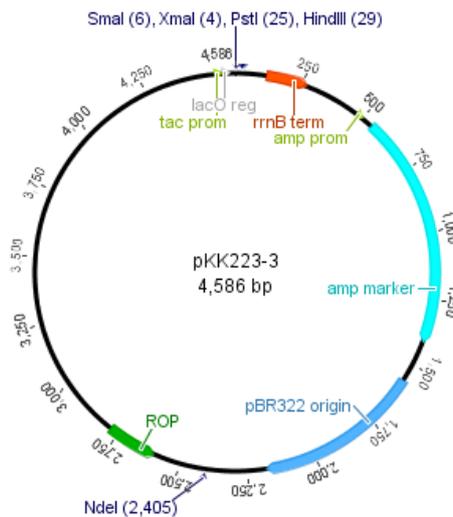


Figure H-9. pKK223-3. The vector features a pBR322 ORI (compatible with p15A and CloDF13 ORIs), one MCS (*Smal*, *XmaI*), and conveys resistance to Amp.

H I.1.2 Esterases

H I.1.2.1 B2S: pET28a(+)_bs2

The plasmid map of the backbone is shown in **Figure H-6** (H I.1.1.6). The *bs2* gene was subcloned into pET28(+) by *NdeI/BamHI*. The translated enzyme bears a C-terminal 6xHis tag.

NdeI BS2 →

CATATGACTCATCAAATAGTAACGACTCAATACGGCAAAGTAAAAGGCACAACGGAAAACGGCGTACATAAGTGAAAGGCATCCCCTATGCCAAGCCGCTGTGCGGACAATGGCGTTTTAAAGCACCTGAGCCGCCTGAAGTGTGGGAAGATGTGCTTGATGCCACAGCGTACGGCTCTATTGCCCGCAGCCGCTGTGATTTGCTGTCACTTTCTATACTAGACTGCCCCGCCAGTCCGAGGATTGCTTGATGTCAATGTATTTGCGCCTGACACCCCAAGTAAAAATCTTCTGTGCATGGTGTGGATTCACGGAGGCGCTTTTATCTAGGAGCGGGCAGTGACCATTGTATGACGGATCAAACCTTGCAGCACAGGGAGAAGTCAATTGTCTTACATTGAACATATCGGCTGGGGCCGTTTTGGCTTTTGCACCTGTCTTCAATTAATGAGGCGTATTCTGATAACCTGGGCTTTTAGACCAAGCCGCCGCTGAAATGGTGCAGAGAATATTTACAGCTTTGGCGGTGATCCGATAACGTAACAGTATTTGGAGAATCCGCCGGCGGATGAGCATTGCCCGCTGCTTGCTATGCCTGCGCAAAGGCCTGTTCCAGAAAGCAATCATGGAAAGCGGCGCTTCTCGAACGATGACGAAAGAACAAGCGGCGAGCACCTCGGCAGCCTTTTAC

```

AGGTCCTTGGGATTAACGAGGGCCAACCTGGATAAATGCATACGGTTCTGCGGAAGATTGCTAAAAGCGGCTGATCAGCTTCGGATTGCAGAAAAAGA
AAATATCTTTCAGCTGTTCTCCAGCCCGCCCTTGATCCGAAAACGCTGCCTGAAGAACCAGAAAAAGCGATCGCAGAAGGGGCTGCTCCGGTATTCCGC
TATTAATTGGAACAACCCGTGATGAAGGATATTTATTTTACCCCGGATTGACAGCTTCATTCTCAGGAAACGCTTGCAGCGCTCGAGTATTTACTAG
GGAAGCCGCTGGCAGAGAAAAGTTGCCGATTGTATCCGCTTCTGGAAGCCAAATTCACATGATGACTGATTTATTTTGGAGCCCTGCCGTCGCT
ATGCATCCGCACAGTCTCATTACGCCCTGTCTGGATGTACAGTTCGATTGGCACCCGAAGAAGCCCGCTACAATAAAGCGTTTCACGATTAGAGCTTC
CTTTGTCTTTGGAATCTGGACGGATTGGAACGAATGGCAAAGCGGAGATTACGGATGAGGTGAAACAGCTTCTCACAGATAACAATCAGCGTGGAT
CACGTTCCGCAAAAACAGGAAACCAAGCACCGAAGCTGTGAATTGGCTGCGTATCATGAAGAAACGAGAGAGACGCTGATTTTAGACTCAGAGATTACG
ATCGAAAACGATCCCGAATCTGAAAAAAGGCAGAAGCTATTCCTTCAAAGGAGAAGGATCCCATCATCATCATCATATTGA
    ← BS2 BamHI                               6xHis
    
```

H I.1.2.2 *PEST*: pET28a(+)_{pest}

The plasmid map of the backbone is shown in **Figure H-6** (H I.1.1.6). The *pest* gene was subcloned into pET28(+) by *NdeI/BamHI*. The translated enzyme bears a C-terminal 6xHis tag.

NdeI PEST →

```

CATATGCCTCTTAGCCCTATACTAAGGCAAATCTCCAACAGTTGGCCGCGCAGTTGCAGTTTAGACCCGACATGGACGTCAAGACGGTGAGAGAGCAGTT
TGAGAAAGTCTCCCTCATCTCTGTCAAAATGGCCAATGAGCCTATTACCGTGTGGAGGACATCACGATTCCGGGCAGGGGGCCCAATTAGGGCTAGG
GTTTATAGGCCGCGGGATGGGAGAGGTTGCCCGCGTGGTGTACTACCAGCGGGGGCTTCGTCTTGGGAGCGTGGAGACTCACGACCAGTGTGT
AGGCGGTTGGCCAACCTCTCCGGGGCAGTCTGCTCTGTGGACTACCGCTAGCCCCGAGCACAAATCCCGCCGCGCTGGAAGACGCATACGACG
CCGCCAAGTGGGTCGCCGACAACACTACGACAAGCTCGGCGTCGACAATGGGAAAATCGCCGTGGTGGGGACTCGGCGGGGGGCAACTTAGCCGCGGTG
ACGGCCATCATGGCCAGGGACAGGGGGGAGAGCTTTGTGAAATACCAAGTGTAAATCTACCCCGCGTCAACCTCACTGGGTCTCCACAGTGTCTAGAG
TGGAGTACAGCGGGCCGAATACGTCATCCTCACCGCCGACTAATGGCGTGGTTGGGAGACAGTATTCTCAAAGCCGCAAGACGCCCTCAGTCCCTAT
GCCTCTCCCATATTTGCAGATTTGTCAAACCTCCGCCCCGCTGGTATAACCGCCGAGTACGACCCGCTACGCGACGAGGGAGAGCTACGCCCACTT
GTTGAAGACTAGGGGAGTTAGGGCGTGGCGGTGAGGTACAACGCGCTCATCCACGGCTTCGTAACCTTCAACCCATATTAGAAGAGGGGAGAGAGGG
AGTTTCGCAAAATTCGCGCTCAATAAAGTCGATGGCTGTGGCGGATCCCATCATCATCATCATATTGA
    ← PEST BamHI                               6xHis
    
```

H I.1.2.3 *Pfel*: pGASTON_{pfel}

The *pfel* gene was subcloned into pGASTON by *NdeI/BamHI*. The translated enzyme bears a C-terminal 6xHis tag (**Figure H-10**).

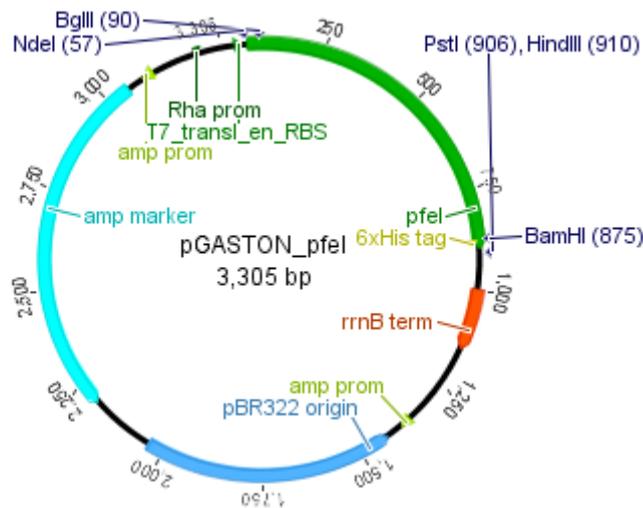


Figure H-10. pGASTON_{pfel}. The vector contains the *pfel* gene, a pBR322 ORI (compatible with p15A and CloDF13 ORIs), and conveys resistance to Amp.

NdeI Pfel →

```

CATATGAGCACATTTGTTGCAAAGACGGTACCCAGATCTATTTCAAGGACTGGGGCAGCGGTAACCGGTTGTTTCAGCCACGGTTGGCTACTGGATGC
CGACATGTGGGAATACCAGATGGAGTACCTCAGCAGCCGCGGCTATCGCACCATCGCTTTGACCGCCGCGCTTTGGCCGCTCGGACCAACCCTGGACC
GGCAACGACTACGACACCTTCGCCGACGACATCGCCAGTTGATCGAACACCTGGACCTCAAGGAGGTGACCCTGGTGGGCTTCCATGGCGGGCGGCG
ATGTGGCCCGCTACATCGCCGCCAGGCAGCGCACGGGTGGCCGCGCTGGTCTGCTGGCGCCGCTACCCCGCTGTTCCGCCAGAAGCCCGACTATCC
    
```

GCAGGGTGTCCCGCTCGATGTGTTGCAAGGTTCAAGACTGAGCTGCTGAAGGATCGCGCGCAGTTCATCAGCGATTTCAACGCACCGTTCTATGGCATCA
 ACAAGGGCCAGGTCGTCTCCCAAGGCGTGCAGACCCAGACCTGCAAATCGCCCTGCTGGCCTCGCTCAAGGCCACGGTGGATTGCGTCACCGCGTTGCG
 CGAAACCGACTTCCGCCGGACATGGCCAAGATCGACGTACCCACCTGGTGATCCATGGCGATGGCGACCATGTCGCGTTCGAGACCACCGGCAAA
 GTGGCGCGGAGTTGATCAAGGGCGCCGAAGTGAAGGTGTACAAGGACGCGCCCCACGGTTCGCGGTGACCCACGCCAGCAGTTGAACGAAGACCTG
 TTGGCGTCTTGAACGCGGATCCATCATCATCATCATCATTGA
 ← PfiI BamHI 6xHis

H I.1.3 ADHs

H I.1.3.1 ADH-A: pET22b(+)_adh-a

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *adh-a* gene was subcloned into pET22b(+) by *NdeI/BamHI*.

NdeI ADH-A →

CATATGAAAGCAGTCCAGTACACCGAAATTGGTAGCGAACCGTTGTTGTTGACATTCGACGCCGACGCCGGTCCGGGTGAAATCCTGCTGAAAGTCA
 CCGCGGCCGGTCTGTGTCATAGCGACATTTTTGTTATGGATATGCCGCGAGCTCAGTATGCATACGGTCTGCCGCTGACGCTGGGTACGAGGGTGTGGG
 TACCGTTCGCGAACTGGCGAAGGTGTGACCGGCTTCGGTGTGGCGATGCTGTTGCACTATGGTCCGTGGGGTTCGGTGCATGTCATGCATGCGCA
 CGTGGTGCAGAACTACTGCACGCTGCAGCAGATCTGGGTATACCCCGCCGGTCTGGGCAGCCCGGGTCTATGGCTGAATATATGATTGTGGACT
 CCGCGGCCATCTGGTTCGATCGGTGACCTGGACCCGGTGGCAGCTGCACCGCTGACGGATGCAGGTCTGACCCGTACCACGCAATTAGTCGTGTTCTG
 CCGCTGCTGGTCCGGCTCCACCGCAGTGGTTATCGGTGTGGCGGTCTGGGTACGTTGGCATTCAAATCCTGCGTCCGTGAGTGCCGACGCGTCA
 TTGCACTGGATCTGGATGACGATCGTCTGGCTCTGGCGCGAAGTTGGCGCAGATGCTGCGGTCAAATCAGGTGCCGCGCCGAGACGCAATTCGTGA
 ACTGACGGGGGTGAGGGTGCACCGCAGTTTTGACTTCGTCGGCGCGCAAAGCAGCATGATACCCTCAGCAAGTCTGGCGGTGGACGGTCAATTT
 TCTGTTGTCGGTATCCATGCTGGCGCGCACGCAAGTTGGCTTTTTCATGATTCGTTTGGCGCCTCAGTGGTTACGCGTATTGGGGACCCGCTCGGAA
 CTGATGGAAGTCTGGCACTGGCTCGTGCAGGTCTGCGATATCCACCCGAAACGTTACCCCTGGACGAAGGCCGCGCGGTATCGTCTGCGTG
 AAGGCTCAATCCGTGGTCTGGCGTCTGTCCTCCGTAAGGATCC
 ← ADH-A BamHI

H I.1.3.2 ADH-ht: pET26b(+)_adh-ht

The plasmid map of the backbone is shown in **Figure H-5** (H I.1.1.5). The *adh-ht* gene was subcloned into pET26b(+) by *NdeI/XhoI*.

NdeI ADH-ht →

CATATGAAAGCAGCAGTTGTGGAACAGTTTAAAGAACCGCTGAAAATTAAGAAAGTGGAAAAACCGACCATAGCTATGGTGAAGTTCTGGTTCGTATTA
 AAGCCTGTGGTGTGTTGTCATACCGATCTGCATGCAGCACATGGTGATTGGCCTGTTAAACCGAAACTGCCGCTGATCCGGGTGATGAAGGTGTTGGTATT
 GTTGAAGAAGTTGGTCTGGCGTTACCCATCTGAAAGTTGGTATCGTGTGGTATTCCGTGGCTGTATAGCGCATGTGGTATTGTGATTATTGTCTGAGC
 GGTCAAGAAAACCTGTGCGAACATCAGAAAAATGCAGGTTATAGCGTGGATGGTGTATGCGAATATTGCTGTCAGCAGCAGATTATGTTGTGAAAA
 TTCCGGATAATCTGAGCTTTGAAGAAGCAGCACCGATTTTTGTGCCGTTGTTACCACCTATAAAGCACTGAAAGTTACCGGTGCAAAACCGGTGAATGG
 GTTGCAATTTATGGTATTGGTGGTCTGGGCCATGTTGCAGTTCAGTATGCAAAGCAATGGGTCTGAATGTTGTTGCAAGTGGATATTGGTGTGAAAACT
 GGAAGTGGCAAAAGAACTGGGTGCAGATCTGGTGTGTTAATCCGCTGAAAGAAGATGCAGCCAAATTTATGAAAGAAAAAGTGGGTGGTGTTCATGCAGC
 AGTGTATCCGCGATTAGCAAACCGCATTTCAGAGCGCATATAATAGCATTGCTGCTGGTGGTGCATGTGTTCTGGTGGTCTGCCTCCGGAAGAAATGC
 CGATCCGATTTTTGATACCGTGTGAATGGCATTAAAATATTGGTATGATTGGGACCCGTAAGATCTGCAAGAAGCACTGCAGTTTGCAGCAGAA
 GGTAAGTTAAAACCATTAATTGAAGTGCAGCCGCTGAAAAAATTAATGAAGTGTGATCGCATGCTGAAAGGTGAGTAAATGGTCTGTTGTTCTGAC
 CCTGGAAGATAAATAACTCGAG
 ← ADH-ht XhoI

H I.1.3.3 ADH_{Lk}: pET21b(+)_adh_{Lk}

The plasmid map of the backbone is shown in **Figure H-3** (H I.1.1.3). The *adh_{Lk}* gene was subcloned into pET21b(+) by *NdeI/XhoI*.

NdeI ADH_{Lk} →

CATATGACTGACCGTTTGAAGGTAAGGTAGCAATTGTAAGTGGCGGTACCTTGGGAATTGGCTTGGCAATCGCTGATAAGTTTGTGAAAGGGCGCAA
 AGGTTGTTATTACCGCCGTCACGCTGATGTAGGTGAAAAAGCTGCCAAATCAATCGCGCGCACAGACGTTATCCGTTTTGTCCAACACGATGCTTCTGAT
 GAAGCCGGTGGACTAAGTTGTTGATACGACTGAAGAAGCATTGGCCAGTTACCACGGTGTCAACAATGCCGAATTGCGGTGAGCAAGAGTGTG
 AAGATACCACAACTGAAGAATGGCGAAGCTGCTCTCAGTTAACTGGATGGTGTCTTCTCCGTACCCGCTTGGAAATCAACGATGAAGAATAAAGGA
 CTCGGAGCATCAATCATCAATATGTCATCTATCGAAGGTTTTGTTGGTATCCAACCTGGGTGCATACAACGCTTCAAAGGTGCTGTCAGAATTATGCT
 AAATCAGCTGCCTTGGATTGCGCTTTGAAGGACTACGATGTTCCGGTTAACACTGTTTCATCCAGGTTATCAAGACACCATTGGTTGACGATCTTGAAGG

GGCAGAAGAAATGATGTCACAGCGGACCAAGACACCAATGGGTCATATCGGTGAACCTAACGATATCGCTTGGATCTGTGTTTACCTGGCATCTGACGAAT
 CTAATTTGCCACTGGTGCAGAATTCGTTGTCGACGGTGGCTACTGCTCAATAGGGATCC
 ← *ADH_{Lk}* *Bam*HI

H 1.1.3.4 *LK-ADH_{N-Int}*: pTYB21_intein-CBD:*adh_{Lk}*

The *adh_{Lk}* gene was subcloned into pTYB21 by *Sap*I/*Nco*I (G VI.2.1.3). The resulting construct contained the ADH with an N-terminal intein-CBD tag for purification and subsequent self-cleavage of the tag (**Figure H-3**). The intein cleavage site is indicated by (▼) in the sequence below.

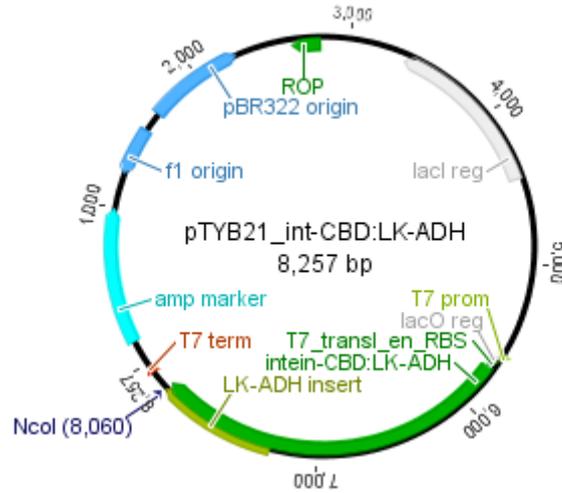


Figure H-11. pTYB21_intein-CBD:*adh_{Lk}*. The vector contains the *adh_{Lk}* gene with an intein-CBD tag fused to its N-terminus. The plasmid features a pBR322 ORI and conveys resistance to Amp.

Intein-CBD →

ATGAAAATCGAAGAAGGTAACACTGGTAATCGGTTCCCTGGAGGGTGTGCTTTGCCAAGGGTACCAATGTTTTAATGGCGGATGGGTCTATTGAATGTATTGA
 AAACATTGAGGTTGGTAATAAGGTCATGGGTAAGATGGCAGACCTCGTGAGTAATTAATGGCCAGAGGAAGAGAACTATGTACAGCGTCGTGCA
 GAAAAGTCAGCACAGAGCCACAAAAGTGACTCAAGTCGTGAAGTGCCAGAATTACTCAAGTTTACGTGTAATGCGACCCATGAGTTGTTGTTAGAACA
 CCTCGTAGTGTCCGCCGTTTGTCTCGTACCATAAGGGTGTGCAATATTTGAAGTTATTACTTTTGTAGATGGGCCAAAAGAAAGCCCCGACGGTAGAATT
 GTTGAGCTTGTCAAGGAAGTTTCAAAGAGCTACCAATATCTGAGGGGCTGAGAGAGCCAACGAATTAGTAGAATCCTATAGAAGGCTTCAAATAAAG
 CTTATTTGAGTGGACTATTGAGGCCAGAGATCTTCTCTGTTGGGTTCCATGTTCTGTAAGGCTACTACCAGACTTACGCTCCAATCTTTATGAGAATGA
 CCATTTTTCGACTACATGCAAAAAAGTAAAGTTTCACTCACCATTGAAGTCCAAAAGTACTTGCTTATTACTTGGTTTATGGATTGGTGATGGATTGTCT
 GACAGGGCAACTTTTTCGGTTGATTCCAGAGATACTTCTTGTGGAACGTGTTACTGAATATGCTGAAAAGTTGAATTTGTGCGCCGAGTA AAGGACAG
 AAAAGAACCACAAGTTGCCAAAAGTGAATTTGACTCTAAAGTTGTCAGAGGTGCTAGCACAAAATCCTGGTGTATCCGCTTGGCAGGTCAACACAGCTT
 ATACTGCGGGACAATTGGTCACATAAACGGCAAGACGTATAAATGTTGCAGCCCCACACCTCCTTGGCAGGATGGGAACCTCAACGTTCTCGCTTG
 TGGCAGCTTCAAGGTGGCCACGGTGGTATTGCAATAATCTAATACTGAGAATCCATTATGGGACGCTATTGTTGGCTTAGGATTCTTGAAGGACGGTGT
 CAAAAATATCCTTCTTCTGTTCTACGGACAATATCGGTACTCGTGAACATTTCTGCTGGTCTAATTGATTCTGATGGCTATGTTACTGATGAGCATGGT
 ATTAAGCAACAATAAAGACAATTCATACTTCTGTCAGAGATGGTTGGTTCCCTGCTCGTTCTTAGGCTTAGTAGTCTCGGTTAACGCAGAACCTGCTA
 AGGTTGACATGAATGTCACCAACATAAAAATTAGTTATGCTATTTATATGCTGGTGGAGATGTTTTGCTTAACGTTCTTTCGAAGTGTGCCGGCTCTAAA
 AATTCAGGCCTGCTCCCGCTGCTTTTGCACGTGAGTGCCGCGGATTTTATTCGAGTTACAAGAATTGAAGGAAGACGATTATTATGGGATTACTTTAT

← *Intein-CBD* ▼ *ADH_{Lk}* →

CTGATGATTCTGATCATCAGTTTTTGTCTGGATCTCAGGTTGTTGTACAGAACATGACTGACCGTTTTGAAAGGTAAGGTAGCAATTGTAACCTGGCGGTACCT
 TGGGAATTGGCTTGGCAATCGCTGATAAGTTTGTGAAGAAGGCGCAAAGTTGTTATTACCGCCGTCACGCTGATGTAGGTGAAAAAGCTGCCAAATC
 AATCGGCGGCACAGAGCTTATCCGTTTTGTCCAACAGATGCTTCTGATGAAGCCGGCTGGACTAAGTTGTTGATACGACTGAAGAAGCAATTTGGCCAG
 TTACCAGGTTGTCAACAATGCCGGAATTGCGGTGAGCAAGAGTGTGAAGATACCACAACGAAGAATGGCGCAAGCTGCTCAGTTAACTTGGATGGT
 GTCTTCTCGGTACCGTCTTGGAAATCAACGATGAAGAATAAAGGACTCGGAGCATCAATCATCAATATGTCATCTATCGAAGGTTTTGTTGGTGATCCA
 ACTCTGGGTGCATACAACGCTTCAAAGGTGCTGTCAGAATATGCTAAATCAGCTGCCTTGGATTGCGCTTGAAGGACTACGATGTTCCGGTTAACT
 GTTCATCCAGGTTATCAAGACACATTGTTGACGATCTTGAAGGGGCAAGAAGAAATGATGTCACAGCGGACCAAGACACCAATGGGTGATATCGGTG
 AACCTAACGATATCGCTTGGATCTGTGTTTACTGGCATCTGACGAATCTAAATTTGCCACTGGTGCAGAATTCGTTGTCGACGGTGGCTACTGCTCAAT
 AGCCATGG

*Nco*I

← *ADH_{Lk}*

H I.1.3.5 *ADH_{Rr}*: pRR_*adh_{Rr}*

The pRR resembles the pKA1 backbone, which is shown in **Figure H-8** (H I.1.1.8). The *adh_{Rr}* gene was subcloned into pRR by *NdeI/BamHI*.

NdeI ADH_{Rr} →

```
CATATGAAAGCGCTGCAGTATACCGAAATTGGCAGCGAACCGGTGGTGGATGTGCCGACCCCGCGCCGGGCCGCGAAATTCTGCTGAAAGTG
ACCGCGCGGGCCTGTGCCATAGCGATATTTTGTGATGGATATGCCGCGGAACAGTATTTATGGCCTGCCGCTGACCCCTGGGCCATGAAGCGGTGG
GCACCGTGGCGGAACTGGGCGCGGGCGTGACCGGCTTTGAAACCGCGCATGCGGTGGCGGTGTATGCCCGTGGGGCTGCGGCGCGTGCATGCGTGC
GCGCGTGGCCGTGAAACTATTGCACCCGTGCGGCGGAACTGGGCATTACCCCGCCGGCCTGGGCGAGCCCGGGCAGCATGGCGGAATATATGATTGTG
GATAGCGCGCTCATCTGGTGGCATTGGCGATCTGGATCCGCTGGCGCGGTGCCGCTGACCCGATGCGGGCCTGACCCCGTATCATGCGATTAGCCGTG
TGCTGCCGCTGCTGGGCCCGGCAGCACCGCGGTGGTGAATTGGCTGGGCGGCCCTGGGCCATGTGGGCATTGAGATTCTGCGTGCCTGGCGCGGCGC
GTGTGATTGCGGTGGATCTGGATGATGATCTGCGCGTGGCGCGTGAAGTGGGCGCGGATGCGGCGGTGAAAGCGGCGCGGGCGCGGGGATGCG
ATTCGTGAAGTACCGGCGCGGAAGCGCGACCGCGTGTGTTGATTTGTGGCGCGCAGAGCACCATGATACCGCGCAGAGGTGGTGGCGATTGAT
GGCCATATTAGCGTGGTGGCATTGATGCGGGCGCGCATGCGAAAGTGGGCTTTTTATGATCCGTTTGGCGCGAGCGTGGTACCCCGTATTGGGGCA
CCCGTAGCGAACTGATGGATGTGGTGGATCTGGCGGTGCGGGCGTCTGGATATTCAACCGAAACCTTTACCTGGATGAAGCGCCGACCGCGTATCG
TCGCTGCGTGAAGGCAGCATTCGTGCGCGTGGCGTGGTGGTGGCGGCTAAGGATCC
```

← *ADH_{Rr} BamHI*

H I.1.3.6 *AlkJ*: pKA1_*alkJ*

The *alkJ* gene was inserted into pKA1 by FC (G VI.2.4.1) because the *alkJ* gene contains various commonly used restriction enzyme sites utilized in molecular cloning (**Figure H-12**).

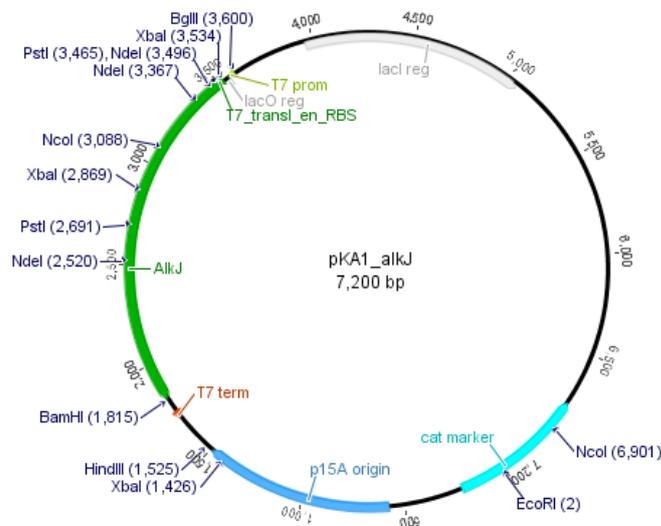


Figure H-12. pKA1_*alkJ*. The vector contains the *alkJ* gene, a p15A ORI (compatible with pBR322 and CloDF13 ORIs), and conveys resistance to Cam.

AlkJ →

```
ATGTACGACTATATAATCGTTGGTGGTGGATCTGCAGGATGTGTGCTTGTCTAATCGTCTTTCGGCCGACCCCTCTAAAAGAGTTTGTACTTGAAGCTGGG
CCGCGAGATACGAATCCGCTAATTCATATGCCGTTAGGTATTGCTTTGCTTTCAAATAGTAAAAAGTTGAATTGGGCTTTTCAAAGTGCACAGCAAAT
CTCAACGGCCGGAGCCTTTTCTGGCCAGGAGAAAACGTTAGGTGGTTCAAGCTCAATCAACGCAATGGTCTATATCCGAGGGCATGAAGACGATTACC
ACGCATGGGAGCAGGCGCGCCGCGCTACTGGGGTTGGTACCGGGCTCTTGAGTTGTTCAAAGGCTTGAATGCAACACGCGATTGATAAGTCCGAGC
ACCATGGGGTTGACGGAGAATTAGCTGTTAGTGATTTAAAATATATCAATCCGCTTAGCAAAGCATTGTCGCAAGCCGGCATGGAGGCCAATATTAATTC
AACGGAGATTTCAACGCGAGTACCAGGACGCGTAGGGTTCTATCAAGTAACCCAAAAAATGGACAACGCTGGAGCTCGGCGCGTGCATTCTTGACG
GTGTACTTTCCAGACCAATCTAGACATCACTACTGATGCGCATGCATCAAAAATCTTTTTGAAGACCCTAAGCGGTTGGTGTCTTATATAAAGAAAA
ATATGCACCATCAAGTCAAGACAACGAGTGGTGGTGAAGTACTTCTAGTCTTGGCGAGTCGCGACGCCTCACCTCTAATGCTTCTGGTGTGGGGCT
GCAGCCGAGCTTAAGGAACATGGTGTCTCTAGTCCATGATCTTCTGAGGTGGGGAAAAATCTTCAAGATCATTGGACATCACATTGATGTGCGCAGC
AAATTCGAGAGAGCCGATAGTGTGCTTTCTTTCATCCCTGTTGCTCGGGTTTGTTCATATGTGTTTAAAGCGAGGGGTTTCTCACTAGTAAAC
GTGGCAGAGTCGGGTGGTTTTGTA AAAAGTCTCTGATCGTGCATCGGCCAATTTGCAAGTTTCTTTCTTCAAATCTTAAAGATCACGGTCGAAAA
```

ATAGCGGGTGGTTATGGTTATACGCTACATATATGTGATCTTTTGCCTAAGAGCCGAGGCAGAAATTGGCCTAAAAAGCGCCAATCCATTACAGCCGCTTT
AATTGACCCGAACATCTTAGCGATCATGAAGATATTAACCATGATTGCGGGTATTAAGATAGGGCGCGCTATTTTGCAGGCCCATCGATGGCGAAGC
ATTTAAGCATGAAGTAGTACCGGGCCAGGCTGTTAAACTGATGATGAAATAATCGAAGATATTCTAGCGGAGCTGAGACTATATACCATCCGGTAGGT
ACTTGTAGGATGGGTAAAGATCCAGCGTCAGTTGTTGATCCGTGCCTGAAGATCCGTGGGTTGGCAAATATTAGAGTCGTTGATGCGTCAATTATGCCGCA
CTTGGTCGCGGTAACACAAACGCTCCAATTATGATTGCAGAAAATGCGGCAGAAAATAATTATGCGGAATCTTGATGTTGAAGCATTAGAGGCTAGC
GCTGAGTTTCTCGCGAGGGTGCAGAGCTAGAGTTGGCCATGATAGCTGTCTGCATG**TAA**
← *AlkI*

H 1.1.4 CARs and *Pptase_{Ec}*

H 1.1.4.1 *CAR_{Ni}*: pETDuet-1_*pptase_{Ec}*::*car_{Ni}*

The plasmid map of the backbone is shown in **Figure H-7** (H 1.1.1.7). The *pptase_{Ec}* gene was cloned into MCS-1 (*NcoI/HindIII*) and the *car_{Ni}* gene into MCS-2 (*NdeI/XhoI*) of the pETDuet-1 vector (G VI.3.1.1). The translated *CAR_{Ni}* bears an N-terminal 6xHis tag.

NdeI Pptase_{Ec} →
CCATGGTTCGATATGAAAACACGATACCTCCCTCCCTTTGCGGACATACGCTGCATTTTGTGAGTTCGATCCGCGCAATTTTGTGAGCAGGATTTACT
CTGGCTGCCGCACTACGCACAACGCTGACGCTGAAACGTAACAGAGCATTAGCCGGACGCGATCGCTGCTTTTATGCTTTGCGGGAATATG
GCTATAAATGTGTGCCGCAATCGCGGAGCTACGCCAACCTGTCTGGCCTGCGGAGGTATACGCCAGTATTAGCCACTGTGGGACTACGGCATTAGCCGT
GGTATCTCGTCAACCGATTGGCATTGATATAGAAGAAATTTTTCTGTACAACCGCAAGAGAATTGACAGACAACATTATTACACCAGCGGAACACGAGC
GACTCGCAGACTGCGGTTTAGCCTTTCTGTGGCCTGACTGACTGCAATTTCCGCCAAAGAGAGCGCATTAAAGGCAAGTGAGATCCAACATGATGAGGT
TTTCTGGACTATCAGATAATTAGCTGGAATAAACAGCAGGTATCATTCATCGTGAGAATGAGATGTTTGTGCTGACTGGCAGATAAAAGAAAAGATAGT
CATAACGCTGTGCCAACACGATTAAT**AAGCTT**
← *Pptase_{Ec}* *HindIII*

NdeI CAR_{Ni} → 6xHis
CATATGTCGTACTAC**CATCA****CCATCA****CCATCA****CA**CGAGTGGATTACCCGGATGAGCGGCTACAGCGCCGATTGCACAGTTGTTGCAGAAGATGAGCAGGT
CAAGGCCGCACGTCGCTCGAAGCGGTGAGCGCGGGGTGAGCGCGCCGGTATGCGGCTGGCGCAGATCGCCGCACTGTTATGGCGGGTTACGCCGA
CCGCCCGCCCGGGCAGCGTGCCTGAACTGAACACCGACGACGCGGCGCCGACCTCGCTGCGGTTACTTCCCGATTGAGACCATCACTATC
GCGAACTGTGGCAGCGAGTGGCGAGGTTGCCGCGCCTGGCATCATGATCCCGAGAACCCTTGGCGCGAGGTGATTTGTCGCGCTGCTCGGCTTAC
CAGCATCGACTACGCCACCCTGCACTGGCCGATACACCTCGCGCGGTTACCGTGGCGTTGACAGCCAGCGCGGGGTGCCAGCTGATCGTATCC
TCACCAGACTTCCCGCGGCTGCTCGCCTCGACCCCGGAGACCTCGATGCGCGGCTCGAGTGCCTACTCGCGGCACACACCGGAACGACTGGTGGT
CTTCGACTACACCCCGAGGACGACGACGCGTGGCCCTTGAATCCGCCCGCCCGCCTTGGCGACGCGGGCAGGTTGGTATCGTCAAAACGCTC
GATGCCGTGCTGCCGGGGCCGCGACTTACCGCCGCGCCACTGTTCCGTTCCCGACACCGACGACGACCCGCTGGCCCTGCTGATCTACACCTCCGGCAG
CACCGGAACGCCGAAGGGCGCGATGTACACCAATCGTTGCGCGCCACGATGTGGCAGGGGAACTCGATGTGACAGGGGAACTCGAACGGGTGGGA
TCAATCTCAACTACATGCCGATGAGCCACATCGCCGTCGATATCGCTGTTCCGGCTGCTCGCTCGCGGTGGCACCCGATACTTCGCGGCCAAGAGCGAC
ATGTCGACACTGTTGAAAGACATCGGCTTGGTACGTCACCCAGAGATCTTCTCGTCCGCGCGTGTGCGACATGGTCTTCCAGCGCTATCAGAGCGAGCT
GGACCGCGCTCGGTGGCGGGCCGACCTGGACACGCTCGATCGGGAAGTGAAGCCGACCTCCGGCAGAATACTCGGTGGGCGCTTCTGTTGGC
GGTCTGCGGACGCGCCGCTGGCCGCGGAGATGAAGACGTTTATGGAGTCCGTCCTCGATCTGCCACTGCACGACGGGTACGGGTGACCGAGGCGGG
CGAAGCGTGTCTGACAACAGATCCAGCGCCGCGGCTGCTGATTACAAGTCTGTCGAGTGGCCGAACTGGGTTACTTCCGACCGACCGGCCG
CATCCGCGCGTGTGAGTGTGTTGAAGGCGGAGACACGATCCGGGCTACTACAAGCGCCCGAGGTACCGCGGAGATCTTCCAGCAGGACGGCTTCT
ACAAGACCGCGATATCGTGGCCGAGCTCGACACGATCGGCTGTTATGTCGACCGTCCGCAACAATGTGCTCAAAGTGTGCGAGGGCGAGTTCGTGAC
CGTCCGCTATCGAGGCGGTGTTGCCAGCAGCCGCTGATCCGGCAGATCTTCTACGCGCAGCAGCAACGTTCTATCTGCTGCGGGTATCGTCCC
CACCGACGACGCGTGGCGGGCCGACACCGCCACTTGAATCGGCACTGGCCGAATCGATTGAGCGCATCGCAAGGACGCGAACCTGCAGCCCTAC
GAGATCCGCGGATTTCTGATCGAGACCGGCTTCCACATCGCAACGGACTGCTCTCCGGCATCGCAAGCTGCTGCGCCCAATCTGAAGGAACG
CTACGGCGCTCAGTGGAGCAGATGTACACCGATCTCGCAGAGGCCAGGCGGATGAGTCTGCGCTGCGCCGCGAAGCCGCGACCTGCCGTTGCTC
GAAACCGTCAAGCGGCGAGGAAAGCGATGCTCGGCTGCTCCGCGATATGCTCCGACGCGCACTTACCGACCTGGCGCGGATTCCCTTCCG
CGTGTGTTCTGAACCTGCTGCACGAGATCTTGGGGTCAAGGTGCCGTTGGTGTGCTGCTGACCCGCGCAACGAGCTGCGCGATCTGGCGAATTA
CATTGAGCGGAACGCAACTCGGCGCGAAGCGTCCACCTTACCTCGTGCACGCGCGGTTCCGAGATCCGCGCCGCGATCTGACCTCGACAAG
TTCATCGATGCCGACCTGGCCGCGCCGACAGCATTCCGACGCGCGGTTGCCAGCGCAGACGGTGTGCTGACCGCGCGAAGCGGTACCTCGGCC
GGTCTGCTGCTGGAATGGCTGGAGCGGCTGGACAAGCGGGTGGCACGCTGATCTGCTGCTGCGCGGTAGTGACGCGCGCGGCCGTAACGGC
TGGACTCGCGTTCGACAGCGGATCCCGCCTGCTCGAGCACTACCAGCAACTGGCCGACGACCTGGAAGTCTCGCGGTGATATCGCGGACCC
GAATCTCGTTCGACGACGCGACTTGGCAGCGTTGGCCGAAACCTGACCTGATGTCATCCCGCGGTTGGTCAACCAGTCTTCCCTACACCC
AGCTGTTCCGCCCCAATGCTGTCGCGCACCGCCGAAATCGTCCGTTGGCGATCACGGCGCGGCAAGCCGGTCACTACCTGTCGACCGTGGAGTGGC
CGACAGGTGACCCGCGGAGTATCAGGAGGACAGCGACTCCGCGAGATGAGCGCGGTGCGCGTGTGCGCGAGAGTTACGCCAACGGTACGGCA
ACAGCAAGTGGCGGGGAGGTCTGCTGCGGAAAGCACAGATCTGTGTGGTTCGCGGTGCGGTTCCGTTCCGACATGATCTGGCGCACAGCC
GGTACGCGGGTCAAGTCCAGGACGTTTACCCGGCTGATCTCAGCTGGTCCGACCGCATCGCGCGTACTCGTTTACCGAACCGACGCG

GACGGCAACCGCAGCGGGCCCACTATGACGGCTTGCCGGCGGACTTACGGCGGGCGGATCACCAGCTCGGCATCCAAGCCACCGAAGGCTCCGG
 ACCTACGACGTGCTCAATCCGTACGACGATGGCATCTCCCTCGATGAATTCGTGACTGGCTCGTGAATCCGGCCACCCGATCCAGCGCATCACCAGCTAC
 AGCGACTGGTTCCACCGTTTCGAGACGGCGATCCGCGCTGCCGAAAAGCAACGCCAGGCTCGGTGCTGCCGTTGCTGGACGCTACCGCAACCCCT
 GCCCGCGGTCCGCGCGGATACTCCCGCCAAGGATTTCAAGCGGGTGAACAACAGCCAAAATCGGTCCGGAACAGGACATCCCGATTGTCGCG
 GCCACTGATGATAAGTACGTACGATCTGGAAGTCTCAGCTGCTTAACTCGAG

← *CAR_W* *Xho*I

H I.1.4.2 *CAR_{Mm}*: pETDuet-1_ *pptase_{Ec}::car_{Mm}*

The plasmid map of the backbone is shown in **Figure H-7** (H I.1.1.7). The *pptase_{Ec}* gene was cloned into MC-1 (*Nco*I/*Hind*III) and the *car_{Mm}* gene into MC-2 (*Nde*I/*Xho*I) of the pETDuet-1 vector (G VI.3.1.1). The sequence of *pptase_{Ec}* is given in H I.1.4.1. The translated *CAR_{Mm}* bears an N-terminal 6xHis tag.

*Nde*I *CAR_{Mm}* → 6xHis

CATATGTCGCCAATCATCACCATCACCATCACACGCGTGAAGAGCGGCTCGAGCGCCGATCCAGGACCTCTACGCCAACGACCCGAGTTCGCCCGCGC
 CAAACCCGCCACGCGATCACCAGCAATCGAGCGCCGGTCTACCGCTACCCAGATCATCGAGACCGTATGACCGGATACGCCGATCGCCGGCT
 CTCGCTCAGCGCTCGGTGAATTCGTGACCGACCGCGCACCGGCCACACCAGCTGCGACTGCTCCCCACTTCGAAACCATCAGCTACGGCGAGCTTTG
 GGACCGCATCAGCGACTGCGGACGTGCTAGCACCAGACAGCGGTGAAACCGGGGACCGGGTCTGCTTGTGGGCTTCAACAGCGTCTGACTACGCC
 ACGATCGACATGACTTTGGCGGGCTGGCGCGGTGGCCGTACACTCGAGACCGCGCGGATAAACCAGTGCAGCCGATCGTCCGCGAGACCCAG
 CCCACCATGATCGCGCCAGCGTGCAGCACTCGTACGCCACCGAATTGGCTCTGTCGGTGCAGACCGTACCCGAGTCTGGTTCGACACCACCG
 GCAGTTGACGCACACCGCGAGCGTGAATCCGCCGGGAGCGCTGGCGGCTCGGCGGTCTGAAACCTGGCCGAGGCCATCGCGCGGGCGA
 CGTGCCCCGCGGTGCTCCGCGGCTCGGCGCCGCGACCGATGTGCCGACGACTCGTCCGCTACTGATCTACCTCGGGCAGCACGGGTGCGCCC
 AAGGGCGGATGATCCCCGACGCAACGTTGCGACCTTGGCGCAAGCGCACCTGGTTCGAAGCGGCTACGAGCCGTCGATCAGCTGAATTCATGC
 CAATGAGCCACGTATGGCCGCCAAATCTGTACGGCAGCTGTGAATGGCGCACCGCTACTTCGTCGAAAAGCGATCTCCACCTTGTTCGAA
 GACCTGGCGTGTGCGGCCACCGAGTACCTCGTCCGCGGTGGGACATGGTTCGACGAGTTTCAGAGTGAGTGCACCGCCGCTGGTGC
 ACGGCGCCGACCGGTGCGCTCGAAGCCAGGTCAAGGCCGAGATACGCAACAGCTGCTCGGTGGACGGTATACAGCGACTACCGGCTCCGCCC
 CTATCTCCGACGAGATGAAGCGTGGTTCGAGGAGTGTGACATGATCTGGTTCGAGGGTACGGCTCCACCGAGCCGGATGATCTGATCGACG
 GAGCCATTCGGCGCCGCGGTAAGTTCGACTACAAGTGGTTCGATGTTCCGACCTGGGTTACTTCTGACCGACCGGCCACATCCGCGGGCGAGTTGCT
 GGTCAAGACCGATAGTTTGTCCGGGCTACTACCAGCGAGCCGAAGTACCGCCGACGTGTTTCGATGCTGACGGTCTACCGGACCGGCGACATCATG
 GCCGAGTTCGGCCGAACAGTTCGTGATCTCGACCGCCGAACAACGTGTTGAAGTGTGCGAGGGCGAGTTCGTACCGTCTCAAATCGAAGCGG
 TGTTTGGGACAGCCACTGGTACGGCAGATCTACATACGGCAACAGCGCCGTCCTACTGTTGGCGGTGATCGTCCACCCAGGAGGCGCTGGA
 CGCCGTGCTGTCGAGGAGCTCAAGGCGCGGCTGGGCGACTCGTGAAGAGGTGCAAGGCGCCGCGCTGCAGTCTACGAGATCCCGCGGACTT
 CATCATGAAACAACCATGGACGCTGGAGAAGGCGCTGCTACCGGCATCCGCAAGTTGGCAGGCGCAGCTGAAAAAGCATTACGGCGAGCTTCTC
 GAGCAGATCTACCGGACCTGGACACGGCCAGGCCGACAACTGCGCTCGTGCGCCAAAGCGGTGCCGATGCGCCGTTGCTGGTACGGTGTGCCGT
 GCGGCGCCGCGCTGTTGGCGGCGAGCGCTCTGACGTCCAGCCGATGCGCACTTACCGATTTGGCGGCGACTCGTGTCCGCGCTGCTTACCA
 ACCTGTGACGAGATCTTCGACATCGAAGTGCCTGGTGGCGTCTGTCAGCCCCGCAACGACTTGCAGGCCCTGGCCGACTACGTCGAGGCGGCTCG
 CAAACCCGGTCTGACGGCCGACCTTCGCTCGGTCCAGGCGCCTCGAATGGCAGGTACCGAGGTGATGCCGGTACCTGTCCCTGGACAAATTC
 ATCGATGCCCAACCTGGCCGAAGTCCCGGCTGCCGCGCAACACCAAGTGCACCGTGTGCTGACCGGCGCCACCGGCTTCTCGGGCGCT
 ACCTGGCCCTGGAATGGTGGAGCGGATGGACCTGGTTCGACGGCAAACTGATCTGCTGGTCCGGGCCAAGTCCGACACCGAAGCACGGGCGCGGCTGG
 ACAAGACGTTGACAGCGCGACCCGAACTGCTGGCCACTACCGCGACTGGCCGCGACCACTCGAGGTGCTCGCCGGTGAACAAGGCGAAGCCG
 ACCTCGGACTGGACCGGACAGCTGGCAACGCTGGCCGACCGTGCACCTGATCGTACCCCGCGCCCTGGTCAACCAGTACTGCCATACAGCCA
 GCTGTTCCGGCCCAACGCGCTGGGACCGCCGAGTGTGCGGCTGGCGCTACCTCCAAGATCAAGCCCTACAGCTACACTCGACAATCGGTGTCGCC
 GACCAGATCCCGCGTTCGCGTTCACCGAGGACCGGACATCCGGTTCATAGCGCCACCCGCGGTCGACGACAGCTACGCCAATGGTACTCGAACA
 GCAAGTGGGCGGCGAGGTGCTGTTGCGGAGGCGCATGACCTGTGTCGCTGCGGTTGCGGTGTTCCGCTGCGACATGATCTGGCCGACACCACATG
 GGCGGACAGCTCAATGTGCCGACATGTTACCCGGATGATCTGAGCCTGGCGGCCACCGGATCGCGCCGGTTGTTCTATGAGCTTGGCGCCGAC
 GGCGCCCGCAACGCGCCACTATGACGGTCTGCCGTGAGTTCATCGCCGAGGCGATTCGACTTTGGGTGCGCAGAGCCAGGATGGTTCCACACGT
 ATCAGTGTGAACCCCTACGACGACGGCATCGGACTCGACGAGTTCGTCGACTGGCTCAACGAGTCCGGTTGCCCATCCAGCGCATCGTACTATGGC
 GACTGGTGCAGCGCTTCGAAACCGCACTGCGCGCACTGCCGATCGGCGAGCGCACAGCTCACTGCTGCCGCTGTTGCAACAATATCGGAGCCGGAGC
 GGCCCGTCCGCGGTCGATCGCCCTACCGATCGTTCGGGCGAGCGTGAAGAGGCCAAGATCGGCCCGACAAAGACATTCCGCACGTGCGCGCGCC
 GATCATCGTGAAGTACGTACGACCTGCGCCTACTCGGCCTGCTTAACTCGAG

← *CAR_{Mm}* *Xho*I

H I.1.5 Aldolases

H I.1.5.1 *Fsa1*: pET16b_ *fsa1*

The plasmid map of the backbone is shown in **Figure H-2** (H I.1.1.2). The *fsa1* gene was subcloned into pET16b by *Nco*I/*Xho*I.

NdeI Fsa1 →

```
CCATGGAACTGTATCTGGATACTTCAGACGTTGTTGCGGTGAAGGCGCTGTACGTATTTTTCCGCTGGCGGGTGTGACCACTAACCCAAGCATTATCGCC
GCGGGTAAAAAACCGCTGGATGTTGTGCTTCCGCAACTTCATGAAGCGATGGGCGGTCAGGGGCGTCTGTTTGCCAGGTAATGGCTACCACTGCCGAAG
GGATGGTTAATGACGCGCTTAAGCTGCGTTCTATTATTGCGGATATCGTGGTAAAAGTTCCGGTGACCGCCGAGGGGCTGGCAGCTATTAAGATGTAAA
AGCGGAAGGGATTCCGACGCTGGGAACCGCGGTATATGGCGCAGCACAAGGGCTGCTGTGCGGCTGGCAGGTGCGGAATATGTTGCGCCTTACGTAA
TCGTATTGATGCTCAGGGCGGTAGCGGCATTACAGACTGTGACCGACTTACACCAGTTATTGAAAATGCATGCGCCGACGGCGAAAAGTCTGGCAGCGAGT
TTCAAACCCCGCTCAGGCGCTGGACTGCTTACTGGCAGGATGTGAATCAATTACTCTGCCACTGGATGTGGCACAACAGATGATTAGCTATCCGGCGGT
TGATGCCGCTGTGGCGAAGTTTGAGCAGGACTGGCAGGGAGCGTTTGGCAGAACGTCGATTTAACTCGAG
```

← *Fsa1 XhoI*

H I.1.5.2 *Fsa1-A129S*: pET16b_ *fsa1-A129S*

The plasmid map of the backbone is shown in **Figure H-2** (H I.1.1.2). The *fsa1-A129S* gene was subcloned into pET16b by *NcoI/XhoI*. The mutated gene contains a single bp exchange (G ▶ T), which is highlighted in red in the sequence below.

NcoI Fsa1-A129S →

```
CCATGGAACTGTATCTGGATACTTCAGACGTTGTTGCGGTGAAGGCGCTGTACGTATTTTTCCGCTGGCGGGTGTGACCACTAACCCAAGCATTATCGCC
GCGGGTAAAAAACCGCTGGATGTTGTGCTTCCGCAACTTCATGAAGCGATGGGCGGTCAGGGGCGTCTGTTTGCCAGGTAATGGCTACCACTGCCGAAG
GGATGGTTAATGACGCGCTTAAGCTGCGTTCTATTATTGCGGATATCGTGGTAAAAGTTCCGGTGACCGCCGAGGGGCTGGCAGCTATTAAGATGTAAA
AGCGGAAGGGATTCCGACGCTGGGAACCGCGGTATATGGCGCAGCACAAGGGCTGCTGTGCGGCTGGCAGGTGCGGAATATGTTTCGCCTTACGTAA
TCGTATTGATGCTCAGGGCGGTAGCGGCATTACAGACTGTGACCGACTTACACCAGTTATTGAAAATGCATGCGCCGACGGCGAAAAGTCTGGCAGCGAGT
TTCAAACCCCGCTCAGGCGCTGGACTGCTTACTGGCAGGATGTGAATCAATTACTCTGCCACTGGATGTGGCACAACAGATGATTAGCTATCCGGCGGT
TGATGCCGCTGTGGCGAAGTTTGAGCAGGACTGGCAGGGAGCGTTTGGCAGAACGTCGATTTAACTCGAG
```

← *Fsa1-A129S XhoI*

H I.1.5.3 *FucA*: pKA1_ *fucA*

The *fucA* gene was inserted into pKA1 by FC (G VI.4.1.2) by replacing the *alkJ* gene in the parent pKA1_ *alkJ*. The resulting vector is shown below.

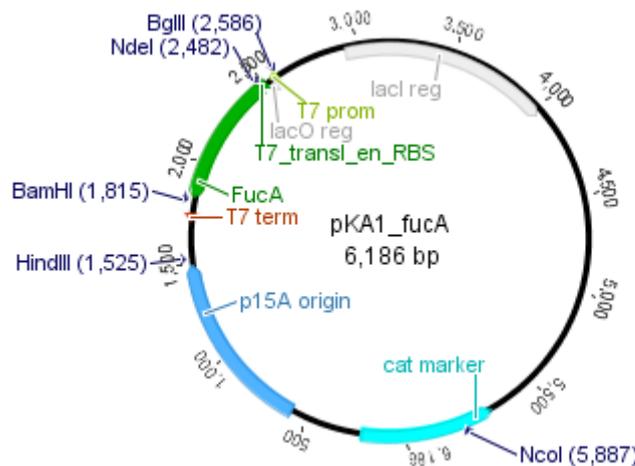


Figure H-13. pKA1_ *fucA*. The vector contains the *fucA* gene, a p15A ORI (compatible with pBR322 and CloDF13 ORIs), and conveys resistance to Cam.

FucA →

ATGGAACGAAATAAACTTGCTCGTCAGATTATTGACACTTGCCTGGAAATGACCCGCTGGGACTGAACCCAGGGGACAGCGGGGAACGTCAGTGTACGTT
 ATCAGGATGGGATGCTGATTACGCCTACAGGCATTCCATATGAAAACTGACGGAGTCGCATATTGTCTTTATTGATGGCAACGGTAAACATGAGGAAGG
 AAAGTCCCCTCAAGCGAATGGCGTTTCCATATGGCAGCCTATCAAAGCAGACCCGGATGCCAACGCGGTTGTTCAACATCATGCCGTTTATTGCACGGCAG
 TTTCCATTCTAACCGATCGATCCCCGCTATTCACTACATGATTGCGGGCGCTGGCGGTAATTCTATTCTTGCAGCCTTATGCGACCTTTGAAACACGCGA
 ACTTCTGAACATGTTGCGCTGGCTCTCAAAAATCGTAAGGCACTTTGTTACAACATCATGGGCTTATCGCTTGTGAGGTGAATCTGAAAAAGCGTTATG
 GCTGGCGCATGAAGTTGAAGTGTGGCGCACTTTACCTGACGACCTGGCGATTACGGACCCGGTGCCAGTGTGAGCGATGAAGAGATTGCCGTAGTG
 CTGGAGAAATCAAAACCTATGGGTTACGAATTGAAGAGTAA

← *FucA*

H I.1.5.4 *FucA*: pKK223-3_ *fucA*

The plasmid map of the backbone is shown in **Figure H-9** (H I.1.1.9). The vector containing the *fucA* gene was provided by M. Kickstein from the TU Dortmund, Germany. The *fucA* sequence is given in H I.1.5.3.

H I.1.5.5 *FruA*: pKK223-3_ *fruA*

The plasmid map of the backbone is shown in **Figure H-9** (H I.1.1.9). The vector containing the *fruA* gene was provided by M. Kickstein from the TU Dortmund, Germany.

FruA →

ATGAACCAAGAACAATTTGACAAAATTAATAATGGTAAAGGATTTATCGCAGCATTGGACCAAAGCGGTGGTAGTACTCCGAAAAGCGCTAAAAGATTATG
 GCGTTGAAGAAAATGAATACTCTAACGATGAAGAAAATGTTCAACCTTGTACACGATATGCGTACTCGTATCATTACTTCACCTGCATTTAACGGAGAAAAA
 TCTTAGGTGCGATTCTATTCGAACAACTATGGACCGTGAAGTTGAGGGCAAATACACAGGTTTATTTAGCAGATAAAGGTATCGTTCCATTCTTGAAA
 GTCGACAAAGTTTGGCTGAAGAAGCTGACGGCGTTCAATTAATGAAACCTATTCCAGACTTAGATAAAATTATTAGATCGTGCGAACGAACGTTGATCTT
 CGGTACTAAAATGCGTTTAAACATCTTAGAAAATAATAAAGAAGCAATTGAAAAAGTTGTTAAACAACAATTTGAAGTTGCAAAAAGAAATCATTGCACTG
 GTCTAGTACCAATTATCGAACCTGAAGTTAACATCAATGCTAAAGACAAAGAAGCTATCGAAGCTAACTTAGCTGAAGCAATCAAAGCTGAATTAGATAAC
 TTGAAAAAAGATCAATATGTAATGTTGAAATTAACCTATTCCAATAAAGTGAATGCTTACAGCGAATTAATGAACATCCACAAGTAATCCGCGTGGTTGCA
 TTATCTGGTGGTTACAGCCGACGAAGCAAAACAAAATCTTGAACAAAATGATGGTTAATCGCAAGCTTCTCACGTGCATTAGTATCTGACTTAAACGCA
 CAACAATCAGATGCAGAATTCAATGAAAATTACAAGAAGCTATCGATACAATCTTCGATGCTTCAGTAAACAAGCTTAA

← *FruA*

H I.1.5.6 *RhuA*: pKK223-3_ *rhuA*

The plasmid map of the backbone is shown in **Figure H-9** (H I.1.1.9). The vector containing the *rhuA* gene was provided by M. Kickstein from the TU Dortmund, Germany.

RhuA →

ATGCAAAAACATACTCAGTCTGTTTGTCCAGGGAATGATCAAAGCCACCACCGACGCTGGCTGAAAGGCTGGGATGAGCGCAACGGCGGCAACCTGA
 CGCTACGCCTGGATGACGCCGATATCGACCATATCACGACAATTTCCACCAACAACCGCGCTATATCCGCTCAGCCAGCCATGCCTTTACTGGCAAATA
 CACCGTTTATTGTCACCGGCTCGGGCAAATCTCCGTAACGTCCAGCTTATCTGCGGCTAACTTAGGCATCGTAAAAGTCGACAGCGACGGCGCGGGC
 TACCACATCTTTGGGGTTAACCAACGAAGCCGTCCTCCAACTCCGGCTCACTTCTTCCACTGCGAGCGCATTAAAGCCACCAACGGCAAA
 GATCGGGTGATCATGCACTGCCACGCCAACCTGATCGCCCTCACCTATGTAATTGAAAACGACACCGCGGTTTCACTCGCCAATGTGGGAAGGCAG
 CACCGAGTGTCTGGTGTATTCCGGATGGCGTTGGCATTTCGCTGGATGGTGCCCGGCACGGACGAAATCGGCCAGGCGACCGCACAAAGAGATGCA
 AAAACATTCGCTGGTGTGGCCCTCCACGGCGCTTCGGCAGCGGACCGACGCTGGATGAAACCTTCGGTTAATCGACACCGCAGAAAAATCAGCAC
 AAGTATTAGTGAAGTTTATTCGATGGCGGCATGAAACAGACCATCAGCCGTGAAGAGTTGATAGCGCTCGGCAAGCGTTTCGGCGTTACGCCACTCGC
 CAGTGCCTGGCGCTGTA

← *RhuA*

H I.1.6 Phosphatases

H I.1.6.1 *PhoN-Se V78L_{fl}*: pCVD_ *phoN-V78L_{fl}*

The gene encoding *PhoN-Se V78L_{fl}* was delivered in a broad host range vector (**Figure H-14**). The sequence including the N-terminal signal peptide (SP) and the C-terminal 3xFLAG tag is given below.

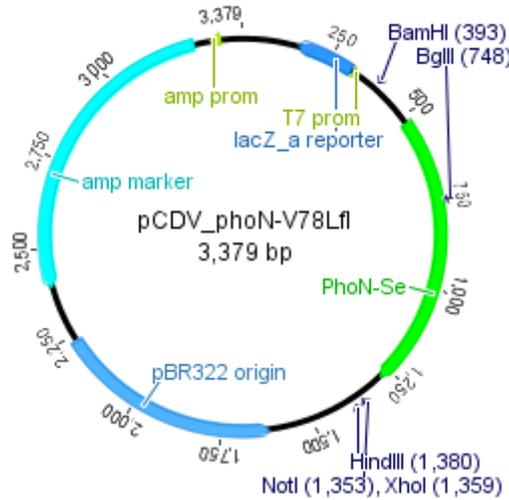


Figure H-14. pCDV_{phoN-V78Lfl}. The vector contains the gene encoding *PhoN-Se V78L_{fl}*, a pBR322 ORI, and conveys resistance to Amp. The vector is not suitable for target gene expression.

NdeI *PhoN-Se V78L_{fl}* → SP
CATATGAAAAGTCGTTATTTAGTATTTTTCTACCACTGATCGTAGCTAAATATACATCAGCAGAAACAGTGCAACCCTTTCATTCTCCTGAAGAATCAGTGA
ACAGTCAGTTCTACTTACCACCACCGCCAGGTAATGATGATCCGGCTTACCCTATGATAAAGGAGGCTTATTTAAGGGCTATGCGATAAAGGGTCCCCG
CGATGGAACAAGCTGCTGAGGATGCAGATCTGAGCGTGAAATATAGCCAGAATATTCTGCCAGTAGTGGGTGCTAAAATTAACCCAAAGATACGC
CAGAAACCTGGAATATGTTAAAGAATCTTCTGACAATGGGCGGCTACTACGCTACTGCTTCGGCAAAAAAATATTATATGCGTACCCGCCCTTTGTCTTAT
TTAATCATTCCACCTGCCGTCCTGAAGATGAGAATACTTTGCGAAAAATGGCTCTTACCCTCCGGGCATACTGCTTATGGTACACTTCTGGCATTAGTATT
ATCCGAGGCCAGACCGGAACGCGCAGGAGCTGCCAGACGCGGATGGGAGTTCGGGCAAGCAGAGTGATGCGGTGCTCACTGGCAAAGCGATG
TTGATGCTGGCCGTTATGTGGGAGCAGTAGAGTTGCAAGACTGCAACAATCCCGGCTTTTCAGAAGTCACTGGCAAAATCCGTGAGGAGCGACTACAA
AGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAGGATGACGATGATAAGTGACTCGAG
3xFLAG ← *PhoN-Se V78L_{fl}* *XhoI*

H I.1.6.2 *PhoN-Se V78L_{fl}*: pET26b(+)_{phoN-V78L_{fl}}

The plasmid in **Figure H-14** was used as PCR template for subcloning of the gene encoding *PhoN-Se V78L_{fl}* for subcloning into pET26b(+) by *NdeI/XhoI*. The phosphatase-coding sequence is given in H I.1.6.1. The target vector backbone is shown in **Figure H-5** (H I.1.1.5).

H I.1.6.3 *PhoN-Se V78L_{Sp}*: pET26b(+)_{phoN-V78L_{Sp}}

The plasmid map of the backbone is shown in **Figure H-5** (H I.1.1.5). The gene including the SP was amplified from the pCDV vector (**Figure H-14**) with appropriate primers (G VI.5.1.1) and subcloned into pET26b(+) by *NdeI/XhoI*.

NdeI *PhoN-Se V78L_{Sp}* → SP
CATATGAAAAGTCGTTATTTAGTATTTTTCTACCACTGATCGTAGCTAAATATACATCAGCAGAAACAGTGCAACCCTTTCATTCTCCTGAAGAATCAGTGA
ACAGTCAGTTCTACTTACCACCACCGCCAGGTAATGATGATCCGGCTTACCCTATGATAAAGGAGGCTTATTTAAGGGCTATGCGATAAAGGGTCCCCG
CGATGGAACAAGCTGCTGAGGATGCAGATCTGAGCGTGAAATATAGCCAGAATATTCTGCCAGTAGTGGGTGCTAAAATTAACCCAAAGATACGC
CAGAAACCTGGAATATGTTAAAGAATCTTCTGACAATGGGCGGCTACTACGCTACTGCTTCGGCAAAAAAATATTATATGCGTACCCGCCCTTTGTCTTAT
TTAATCATTCCACCTGCCGTCCTGAAGATGAGAATACTTTGCGAAAAATGGCTCTTACCCTCCGGGCATACTGCTTATGGTACACTTCTGGCATTAGTATT
ATCCGAGGCCAGACCGGAACGCGCAGGAGCTGCCAGACGCGGATGGGAGTTCGGGCAAGCAGAGTGATGCGGTGCTCACTGGCAAAGCGATG
TTGATGCTGGCCGTTATGTGGGAGCAGTAGAGTTGCAAGACTGCAACAATCCCGGCTTTTCAGAAGTCACTGGCAAAATCCGTGAGGAGCGACTCACTCGA
← *PhoN-Se V78L_{Sp}* *XhoI*

H I.1.6.4 *PhoN-Se V78L_{3xFLAG}*: pET26b(+)_{phoN-V78L_{3xFLAG}}

The plasmid map of the backbone is shown in **Figure H-5** (H I.1.1.5). The gene including the 3xFLAG tag was amplified from the pCDV vector (**Figure H-14**) with appropriate primers (G VI.5.1.1) and subcloned into pET26b(+) by *NdeI/XhoI*.

NdeI *PhoN-Se V78L_{3xFLAG}* →

CATATGAAATATACATCAGCAGAAACAGTGAACCCCTTTCATTCTCCTGAAGAATCAGTGAACAGTCACTTACTTACCACCACCGCCAGGTAATGATGAT
 CCGGCTTACCCTATGATAAGGAGGCTTATTTAAAGGGCTATGCGATAAAGGGTCCCGCGATGGAAACAAGCTGCTGAGGATGCGATCTGAGCGTGG
 AAAATATAGCCAGAATATTCTGCCAGTAGTGGGTGCTAAAATTAACCCCAAAGATACGCCAGAAACCTGGAATATGTTAAAGAATCTTCTGACAATGGGC
 GGCTACTACGCTACTGCTTCGGCAAAAAATATATATGCTACCCGCCCTTGTCTTATTAATCATTCCACCTGCCGCTCCTGAAGATGAGAATACTTTGC
 GAAAAAATGGCTCTTACCCTCCGGGCATACTGCTTATGGTACACTTCTGGCATTAGTATTATCCGAGGCCAGACCAGAACCGCGCAGGAGCTCGCCAGA
 CGCGGATGGGAGTTCCGGGCAAGCAGAGTGATATGCGGTGCTCACTGGCAAAGCGATGTTGATGCTGGCCGTTATGTTGGGAGCAGTAGAGTTTGAAGA
 CTGCAACAATCCCGGCTTTCAGAAAGTCACTGGCAAATCCGTGAGGAGCGGACTACAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAA
GGATGACGATGATAAGTGACTCGAG 3xFLAG
 ← *PhoN-Se V78L_{3xFLAG}* *XhoI*

H I.1.6.5 *PhoN-Sf*: pET26b(+)_{*phoN-Sf*}

The plasmid map of the backbone is shown in **Figure H-5** (H I.1.1.5). The *phoN-Sf* gene was subcloned into pET26b(+) by *NdeI/XhoI*.

NdeI *PhoN-Sf* →

CATATGAAAAGACAGCTTTTTACTCTTAGTATTGTCGGGGTATTTCTTTAAATACCTTTGCATCAATTCCTCCGGGAAATGATGTGACAACAAAACCTGACC
 TTTACTACCTGACAAACGATAATGCTATTGACAGTCTGGCATTATTACCGCTCCACCACAAATCGGAAGTATTGCCTTCTGAATGACCAGGCCATGTATG
 AAAAGGGCGTTTTATTGCGAAATACTGAAAGAGGGAAAGCTGGCAGCTGAGGATGCCAATCTCAGTAGTGGTGGCGTGGCAAATGATTTTTCTGCTGCTTT
 TGGTTCTCCGATAACGGCTAAAGATTACCCAGAATTGCATAAGTTACTGACAAATATGATTGAGGATGCCGAGACCTGGCAACCGGTTTCCAGCAAAAGAAT
 ATTATATGCGTATTCGACCTTTTGTCTTTTATGGTGTCTACGTGTAATAACAAAAGAACAGGATACATTATCCAGAAATGGCTCTTATCCATCAGGTCATAC
 ATCGATTGGTTGGGCAACAGCATTGTATTGTCGAAATAAATCCGGCACGCCAGGATACCATTCTGAAACGGGGTTATGAACTGGGGGACAGCAGGGTT
 ATTTGCGGTTACTCTGGCAAAGTATGTTGACGACGACGATTGTTGGCTGCTATTGTAGCAACACTGCACTCAAATCTGTGTTCCAGGCCAATAA
 CAAAAGCGAAAGATGAATTCGCTAACAAATCAGAAAAATAACTCGAG
 ← *PhoN-Sf* *XhoI*

H I.1.6.6 *YqaB*: pCDFDuet-1_{*yqaB*}

The plasmid map of the backbone is shown in **Figure H-1** (H I.1.1.1). The *yqaB* gene was subcloned into the MCS-1 in pCDFDuet-1 by *NcoI/BamHI*.

NcoI *YqaB* →

CCATGGGATACGAGCGTTATGCAGGTTAATTTTTGATATGGATGGCACAACTCTGGATACGGAGCTACGCACCGTAAAGCGTGGCGCGAAGTATTAGG
 GCACTACGGTCTCAGTACGATATTAGGCGATGATTGCGCTTAATGGATCGCCACCTGGCGTATTGCTCAGGCAATTATTGAGCTGAATCAGGCCGATCT
 CGACCCGATCGCTTAGCGCGTAAAAAACAAGCAGTAAGAAGTATGCTGCTGGATAGCGTCAACCGCTTCTCTTGTGATGTGGTAAAAAGTTGG
 CATGGTCTGCGCCAAATGGCTGAGGAACGGGGAGTAAAGCGCCATCGTGAAGCATTGCTGGCGCACCTGGGATTACGCCATTATTTGACGCCGTCG
 TCGCTGCCGATCACGTCAAACACCATAAACCCGCGCCAGACACATTTTTGTTGTGCGCGCAGCGTATGGCGTGAACCGACGCGATGTGTGGTCTTTGAA
 GATGCCGATTTGCTATTAGGCGGCCCTGCAGCAGGCATGGACGCCGTGGATGTTGCTGCTGTGAGGATCC
 ← *YqaB* *BamHI*

H I.1.7 Mutant GDHs

H I.1.7.1 *GDH_{2xBs}*: pET28a(+)_{*gdh_{2xBs}*}

The plasmid map of the backbone is shown in **Figure H-6** (H I.1.1.6). The 2x mutant gene was subcloned into the pET28a(+) by *NcoI/HindIII*. The translated enzyme bears an N-terminal 6xHis tag.

NcoI *GDH_{2xBs}* →

CCATGGGCAGCAGCCATCATCATCATCAAGCAGCGCCCTGGTCCGCGCGGCGAGCCATATGTATCCGGATTTAAAGGAAAAGTCTGCTATTACA
 GGAGCTGCTTCAGGGCTCGGAAAGGCGATGGCCATTCGCTTCGGCAAGGAGCAGGCAAAAGTGTTATCAACTATTATAGTAATAAAACAAGATCCGAACG
 AGGTAAGAAGAGAGTCAAGGCGGGCGGTGAAGCTGTTGCTGTCGAAGGAGATGTACGAAAGAGGAAGATGAAAAAATATCGTGCAACCGGCA
 ATTAAGGAGTTCCGACACTCGATATTATGATTAATAATGCCGGTCTTGAAATCTGTGCCATCTCACGAAATGCCGCTCAAGGATTGGGATAAAGTCATC
 GGCACGAACCTAACGGGTGCCTTTTAGGAAGCCGTGAAGCGATTAATATTTGCTAGAAAACGATATCAAGGAAATGTCATTAACATGTCCAGTGTGCA
 CGAAGTGATTCCTTGGCCGTTATTTGTCCACTATGCGGCAAGTAAGGCGGGATAAAGCTGATGACAAAAACATTAGCGTTGGAATACGCGCCGAAGGGC
 ATTCGCGTCAATAATATTGGGCCAGGTGCGATCAACACGCCAATCAATGCTGAAAAATTCGCTGACCCTAAACAGAAGGCTGATGTAGAAAGCATGATTCC
 AATGGGATATATCGCGAACCGGAGGAGATCGCGCAGTAGCAGCTGGCTTCTGCTGAAAGGAAAGCCAGCTACGTACAGGCATCACGTTATTCGCGGAC
 GGCGGTATGACTGTATCCTTCATTCCAGGCAGGCCGCGGTTAAAGCTT
 ← *GDH_{2xBs}* *HindIII*

H I.1.7.2 *GDH_{7xBs}*: pET28a(+)*_gdh_{7xBs}*

The plasmid map of the backbone is shown in **Figure H-6** (H I.1.1.6). The 7x mutant gene was subcloned into the pET28a(+) by *NcoI*/*HindIII*. The translated enzyme bears an N-terminal 6xHis tag.

NcoI *GDH_{7xBs}* → 6xHis

CCATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTCCGCGCGGCAGCCATATGTATCCGATTTAAAAGGAAAAGTCGTCGCTATTACAGGAGCTCAGGCTCGGAAAGGCGATGGCCATTCGCTTCGGCAAGGAGCAGGCAAAAGTGGTTATCAACTATTATAGTAATAACAAGATGCGGAAAGAGTAAAGAAGAGTCAATCAAGCGGGCGGTGAAGCTGTTGTCGTCGAAGGAGATGTCACGAAAGAGGAAGATGTAATAAATATCGTCAAACGGCAATTAAGGAGTTCGGCACACTCGATATTATGATTAATAATGCCGGTCTTGAAAATCCTGTCCATCTCACGAAATGCCGCTCAAGGATTGGGATAAAGTCATCGGCACGAACTTAACGGGTGCCTTTTAGGAAGCCGTGAAGCGATTAATATTCGTAGAAAACGATATCAAGGGAAATGTCATTAACATGTCCAGTGTGCACGAAGTGATTCCTGGCCGTATATGTCCACTATGCGGCAAGTAAAGGCGGGATAAAGCTGATGACAAAAACATTAGCGTTGGAATACGCGCCGAAGGGCATTGCGTCAATAATTTGGGCCAGGTGCGATCAACACGCCAATCAATGCTGAAAAATTCGCTGACCCTAACAGAAGGCTGATGTAGAAAGCATGATTC CAATGGGATATATCGGCAACCGGAGGAGATCGCCGACGCGCAGCCTTCCTGCTTGAAGGAAGCCAGCTACGTACAGGCATCACGTTATTCGCGGA CGGCGGTATGACACTGTATCCTTCATTCCAGGCAGGCCGCGGTAAAGCTT

← *GDH_{7xBs}* *HindIII*

H I.1.8 Transaminases

H I.1.8.1 *3FCR*: pET22b(+)*_3fcr*

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *3fcr* gene was subcloned into pET22b(+) by *NdeI*/*BamHI* and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The enzyme bears a C-terminal 6xHis tag.

NdeI *3FCR* →

CATATGCTGAAAAACGACCAACTGGACCAATGGGACCGTGATAACTTCTCCACCCGTCACGCACCTGGCGCAACATGCCCGTGGCGAATCAGCTAACCGTGTGATCAAAAACCGCGTCGGCGTTTTTATTGAAGATCGCGACGGTACGAAACTGCTGGATGCTTTCGCGGGCCTGTATTGCGTTAATGTCGGCTACGGTCGTCAGGAAATTGCCGAAGCAATCGCTGATCAAGCGCGCAACTGGCCTATTACCATAGCTATGTGGCCACGGTACCGAAGCTTCTATCACGCTGGCGAA AATGATTCTGGATCGTGCCCGAAAAACATGAGTAAAGTTTACTTTGGTCTGGCGGTTCCGACGCAAAACGAAACCAATGTCAAAGTATCTGGTATTACA ACAATATTTGGGCCCGCCGGAGAAAAAGAAAATTATCAGTCGTTGGCGCGTTATCATGGCAGTGGTCTGGTTACCGGCTCCCTGACGGGTCTGGAAT GTTTCATAAAAAATTCGATCTGCCGGTGGAAACAGGTTATTACACCGAAGCCCCGTATTACTTTCGTCGGAAGACCTGAACGACGGAAGAACAATTCG TCGCACTGTGTGGCTGAACTGGAAGCGCTGATCGAAGTGAAGCGCGGATACCATTCGCGCCTTCATCGCGCAACCGATTCTGGGTACGGGCGGTAT TGTGCCGCCCGCGCGGTTATTGGGAAGCAATCCAGACCGTCTGAATAAACATGATATTCTGTTGCGGACGAAGTGGTTACCGGCTTTGGTCGCC TGGGACAGATGTCGGTCTGATCACTATGGCCTGGAACCGGACATTATCCATCGCGAAAGTCTGACGTCAGCGTACGCCCGCTGAGCGGTTCTATT GTGTCGGATAAAGTCTGGAAGTGTGGAACAGGGCACCGACGAAACGGTCCGATCGCCATGGTTGGACGTATAGCGCACACCCGATTGGTGCAGCT GCAGGTGTGCAAATCTGAAACTGCTGGATGAACTGAACTGGTTAGCAATGCCGCGAAGTGGTGCCTACCTGAACGCAACCATGGCAGAAGCTCTGT CCCAACATGCTAATGTTGGCGATGTCCTGGCGAAGTCTGCTGTGCGCGGTGGAATTTGTTAAAGATCGTGACAGCCGACGTTTTTCGATGCCGCGAGC AAAATCGGTCCGAGATTTCTGCGAAACTGCTGGAACAAGATAAAATTATCGCGCGTCCATGCCGAGGGCGACATTCTGGGTTTTGCCCGCGGTTCTG TCTGACCCGCGCAGAAGCTGATCAAGTCTGGAAGGTACGCTGCGCGTGTCAAAGCCGTTCTGGGTTCACATCACCATCACCCACTAAAGGATCC

← *3FCR* 6xHis *BamHI*

H I.1.8.2 *3GJU*: pET22b(+)*_3gju*

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *3gju* gene was subcloned into pET22b(+) by *NdeI*/*BamHI* and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The enzyme bears a C-terminal 6xHis tag.

NdeI *3GJU* →

CATATGCTGAACCAATCGAACAACTGAACGCCTGGGATCGCGACCACTTTTTCCACCCGTCACGCACATGGGCACGCACGCACGCGCGGAAAAGTCCGA CCCGTATATGGCCGCGGTGAAGGCGTTACGGTCTGGGATAACAATGGTCGCAAATCCATTGACGCGTTTCCCGCCTGTATTGCGTGAACGTTGGCTAC GGTCGCCAGAAAATTGAGATGCTATCGCGACCAAGCTAAAAATCTGGCGTATTACCATGCCTATGTGGCCACGGTACCGAAGCGAGCATTACGCTGG CCAAAATGATTATCGATCGTGCGCGGAAAGCATGTCTCGCGTTTACTTCGGCTGAGCGGTTCTGATGCCAACGAAACCAATCAAAGTATCTGGTAC TACAACAACGTCCTGGGCCGTCGGAGAAAAAGAAAATTATCTCCGTTGGCGCGGTTATCATGGCAGCGGTTATGACCGGCTCTCTGACGGGTCTGG ACCTGTTTCATAACGCATTGCACTGCCGCTGCTCCGGTGTGTCACACCGAAGCCCCGTTACTTTCTGTCGACGGATGCGAGTATGTCCGAAGAACAGT TCAGCCAACTGTGACAGAAAAGTGAAGAAATGATTCTGGTGAAGGCCGGAACCATTCGCGCCTTATCGCGCAACCGATTCTGGGTACGGGCGG TATCGTTCGCCGCGCGGGTTATTGGGAAAAAATTCAGGCCGCTGAAAAAATACGATGTGCTGCTGGTTGCGGACGAAAGTGGTTACCGGCTTTGGT CGCCTGGGCACGATGTTGCGTTCAGATCATTATGGCATCAAACCGACCTGATTACCATCGCAAAGGCCGACGAGTGCCTACGCACCGCTGTCCGGTGT CATTGTGGCGGATCGTGTGGGACGTTCTGGTCCAAGGCTCAGACAAAAGTGGTTCGCTGGGCCATGGTTGGACCTATTCGGCACACCCGATCTGCGTG

GCAGCTGGTGTGCCAACCTGGAACCTGATTGATGAAATGGACCTGGTACCAATGCGGGCGAAACGGGTGCATATTTTCGTGCTGAACTGGCTAAAGCGG
 TTGGCGGTACAAAAATGTCGGCGAAGTGCAGCGGCGATGGTATGCTGGCGGCCGTTGAATTCGTCGAGATAAAGATGACCGTGTGTTTTCGACGCTTC
 ACAGAAAAATCGGTCCGCAAGTCGCAACCCGACTGGCAGCTTCGGGTGTGATCGGTGTCGAATGCCGAGGGCGATATTCTGGGTTTTGCCCGCCGCTG
 TGTCTGACCCGTGAACAGGCAGATATTGCTGTGAGCAAACGGCCGACGCTGTCAAATCAGTCTTCGCAAACCTGTCA**CACCACCATCACCACCACTAAGG**
ATCC ← 3*GU* 6xHis *Bam*HI

H I.1.8.3 3*HMU*: pET22b(+)_{3hmu}

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *3hmu* gene was subcloned into pET22b(+) by *Nde*I/*Bam*HI and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The enzyme bears a C-terminal 6xHis tag.

*Nde*I 3*GU* →

CATATGAGCCTGGCGACCATTACGAACCACATGCCGACGGCGAACTGCAAGCCCTGGATGCTGCCACCACCTGCACCCGTTTAGCGCAAACAATGCACT
 GGGTGAAGAAGGCACCCGTGTTATTACCGGTGCTCGCGGTGTCTGGCTGAACGATAGCGAAGGCGAAGAAATCTGGACGCCATGGCAGGTCTGTGGT
 CGTCAATATCGTTATGGTCTGATGAACTGGCAGAAGTGGCAGCAGTCAGATGCGTGAAGTCCGTTATTACAACACCTTTTTCAAACACCGCATGTTT
 CGGCTATTGCGCTGCCAAAAAAGTGGCAGAAGTGGTCCGGGCGATCTGAATCACGTGTTTTTCGCGGCGGTGGCAGCGAAGCAAACGACACCAATAT
 CCGTATGGTGCACGATTTGGCAGAACAAAGGTCAACCGGAAAAACCGTTATTATCAGCCGTAATAATGCGTACCATGGCTCTACGGTGCAGACTCTG
 CACTGGGTGGCATGGCTGGTATGCACGCGCAGTCTGGCTGATTCCGGATGTGCATCACATCAACCAACCGAATTGGTGGGCCGAAGGTGGCGATATGGA
 CCCGGAAGAATTTGGTCTGGCAGTGTGCGAAGTGGAAAGCAATTTGGAAGTGGTGAACCGTGTGGCAGCTTTATTGCTGAACCGTTTCAG
 GGTGCGGGTGGCGTATCGTTGCACCGGATTCAATTGGCCGAAATTCACGCATCTGCGATAAATACGACATTCTGCTGATCGCGGACGAAGTTATTTG
 TGGTTTTGGCCGTACCGTAATTGGTTCGGCACCCAGACGATGGGTATCCGCCGCATATTATGACGATCGAAAAAGTCTGAGTTCGGCTATGCTCCGA
 TTGGTGGCTCAATCGTGTGTGATGAAGTCGACACGATGATTGGCAAAGACGAATTAACCATGGTTATACCTACTCGGGTACCCGGTGGCAGCAGCAGTT
 GCACTGAAAAATCTGCGTATTCTGGAAGAAGAAAACATCTGGATCATGTTTCGCAATGTCGCTGCGCCGTATCTGAAAGAAAAATGGGAAGCACTGACCG
 ACCACCCGCTGGTGGTGAAGCCAAAATTGTGGCATGATGGCATCCATCGCTGACCCCGAACAAAGCGTCACGCGCCAAATTTGCATCGAAACCGGG
 TACGATTGGCTACATCTGCCGTGAACGCTGTTTCGCGAACAAATCTGATTATGCGTCATGTCGCGGATCGCATGATTATCAGTCCGCCGCTGGTATTACCC
 GGCCGAAATCGATGAAATGTTCTGTTATCCGCAAATCCCTGGACGAAGCGCAGGCCGAAATGAAAAACAGGGTCTGATGAAATCTGAAGGTAGT**CAC**
CACCACCACCACCACTAAGGATCC ← 3*GU*
 6xHis *Bam*HI

H I.1.8.4 3*i5T*: pET22b(+)_{3i5T}

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *3i5T* gene was subcloned into pET22b(+) by *Nde*I/*Bam*HI and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The enzyme bears a C-terminal 6xHis tag.

*Nde*I 3*i5T* →

CATATGAGCCTGCGTAATGATGCAACGAATGCCGCGGGTGTGTCGGTGTGCGATGCGTGACCATATCCTGCTGCCGGCCAAAGAAATGGCGAAACTGG
 GCAAATCTGCCAGCCGTTTCTGACCCATGCGGAAGGCATTTATGTCCACACGGAAGACGGTCTGCGCTGATGATGGTCCGGCAGGCATGTGGTGC
 TCAAGTGGGTTACGGTCTGCGAAATTTGTTGATGCAATGGCTCATCAAGCAATGGTCTGCGGTATGCTTCTCCGTGGTACATGGCAACCACTCCGGCAG
 CACGTCTGGCAGAAAAAATTGCTACCTGACGCCGGGTGACCTGAACCGTATCTTTTTACCACGGGGCGGTAGTACCGCAGTGGATTCCGCACTGCGTTTT
 CAGAATTTATAACAATGCTCTGGTTCGCCCGAGAAAAACGTATTATCGTGCCTATGATGGTACCATGGTAGCACCGCACTGACGGCAGCTTGTACC
 GGTCGTACGGGTAAGTGGCCGAATTTGATATTGCCCAAGACCGCATCTCTTCTGAGCTCTCCGAACCCGCGTACGCTGGCAATCGCTCACAGGAAGC
 ATTTCTGGATGACCTGTTCAAGAATTCGAAGACCGTATTGAATCTCTGGTCCGGATACCATCGCGGCTTTCTGGCGGAACCGATTCTGGCCAGTGGCG
 GTGTGATTATCCCGCCGCGCAGGTTATCATGCTGTTTTAAAGCGATTTGCGAAAAACACGACATTCTGTACATCTCGGATGAAGTGGTTACCGGCTTCGGT
 GCTGTGGCGAATGTTTTCGAGCGGAAAAAGTTTTCGGTGTCGTGCCGATATTATCACCTTTGCAAAAAGGCGTTACGTCGGGTTATGTCGCCGCTGGCGGT
 CTGGCAATCAGCGAAGCGGTGCTGGCCGATTTCCGGCGAAAACGCCAAAGTTCATGGTTACCAACGGCTATACGTAATCAATCAGCCGTTGTCATG
 CGCAGTCTGCACTGGCAAATATCGAACTGATGGAACGTGAAGTATTGTGACCAAGCCCGGAAATGGCAGATTACTTTGCCGACGCTGGCATCTCTGC
 GTGATCTGCGGGTGTGGCAGAAACCCGAGTGTGGCCTGGTGGTGTGCGTGCAGTGTCTGCTGGACCCGACCCGCTGACGCGTACGCGGGAAGATA
 AAGCCTTTACGCTGAAATCGACGAACGTTGCTTCAACTGGCCGCTGATTGTCCGTCGCTGGGTGATCTGTGTGATCTGCGCCGCTGATTATCAGCC
 GTGCCAGATTGATGAAATGGTGGCAATTATGCGCAAGCTATCACGAAGTTAGCGCGGCCACGGCTGACGGCGAAAGAACCGGCTGCTGTGCGAAG
 GTAGT**CACCACCACCACCACCACTGATAAGGATCC**
 ← 3*i5T* 6xHis *Bam*HI

H 1.1.8.5 *AspFum*: pET22b(+)_{aspfum}

The plasmid map of the backbone is shown in **Figure H-4** (H 1.1.1.4). The *aspfum* gene was subcloned into pET22b(+) by *NdeI/BamHI* and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The enzyme bears a C-terminal 6xHis tag.

NdeI AspFum →

CATATGGCATCTATGGATAAAAGTTTTTAGTGGTTACTACGCGGCCAGAAAAGTCTGGAACGTAGTGATAATCCGTTCCAGCAAAGGTATTGCCTATGTTGAAGGCAAACCTGGTCTGCCGAGTGATGCACGCATCCCGCTGCTGGATGAAGTTTTATGCATAGTGATCTGACCTACGATGTTATTAGCGTGTGGGATGGCCGTTTCTTCGCCTGGATGATCACCTGCAGCGCATCCTGGAAAGTCTGCGATAAAATGCGTCTGAAAATTTCCGCTGGCACTGAGCTCTGTTAAAAACATTCTGGCAGAAATGGTGGCGAAAAGCGGCATTCTGTGATGCGTTCGTGGAAGTTATTGTTACCCGTTGCTGACCGGTTCGTGGTTCTAAACCGGAAAGATCTGTATAACAATAACATTTACCTGCTGGTTCGCGGTATATCTGGGTGATGGCGCCGAAAATCAGCTGCATGGCGGTGAAGCCATTATTACCCGTACCGTTCGTCGCACCCCCCGGGTGCAATTTGATCCGACCATTAAAACTGCAGTGGGGTGATCTGACCAAAGCCCTGTTGAAGCCATGGATCGTGGTGCAACCTATCCGTTCCTGACGGATGGCGATACCAATCTGACGGAAGGCTCTGTTTTCAATATCGTTCGGTGAAGGTTCCGGTGAAGCCAGGCGTATCATAGCGATGAAATTTTATGTGTACCACGGCCGGCGTATTATGCCGATCACCTGCTGGATGGTCAGCCGGTTAATGATGGTCAAGTGGGCCGATTACCAAGAAAATTTGGGATGGCTATTGGGAAATGCACTATAACCCGGCGTACAGCTTCCCGTGGATTACGGCTCTGGTTCAGGATCCCATCATCATCATCATCATTGAGGATCC

← *AspFum* 6xHis BamHI

H 1.1.8.6 *AspTer*: pGASTON_{aspter}

The *aspter* gene was subcloned into pGASTON by *NdeI/BamHI* and provided by Uwe T. Bornscheuer from the Ernst-Moritz-Arndt University Greifswald, Germany. The plasmid map (**Figure H-15**) and the sequence are given below. The enzyme bears a C-terminal 6xHis tag.

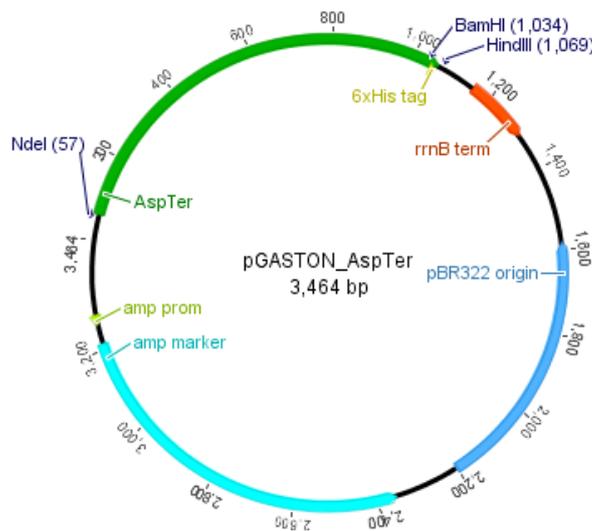


Figure H-15. pGASTON_{aspter}. The vector contains the gene encoding *AspTer*, a pBR322 ORI (compatible with p15A and CloDF13 ORIs), and conveys resistance to Amp.

NdeI AspTer →

CATATGGCAAGCATGGATAAAAGTTTTTGCCGGTTATGCAGCACGTACGGCAATTCTGGAAGCACCGAAACCACCAATCCGTTTGCAAAGGTATTGCATGGTTGAAGGTGAAGTGGTCCGCTGGCAGAAGCACGTATCCGCTGCTGGATCAGGGTTTTATGCATAGCGATCTGACCTATGATGTTCCGAGCGTTTGGGATGGTCTTTTTCTGCTGGATGATCATATTACCCGCTCGGAAGCCAGCTGTACCAAATGCGTCTGCGTCTGCCGCTGCCTCGTGATCAGGTTAAACAAATTCGTGTTGAAATGGTTGCCAAAAGCGGTATTCGTGATGCATTTGTGGAAGTATTGTTACCCGTTGCTGAAAGGTGTTCTGTTGACCGTCCGGAAGATATCGTGAATAATCTGTATATGTTGTGCAGCCGTATGTTGGTTATGGAACCGGATATGCAGCGTGTGGTGGTAGCGAGTTGTTGCACGTACCGTTCGTCGTGTTCCGCTGGTCAATTTGATCCGACCGTTAAAAATCTGCAGTGGGGTGATCTGGTTCGTGGTATGTTTGAAGCAGCAGATCGTGGTGCAACCTATCGTTTTCTGACCGATGGTATGCACATCTGACCGAAGGTAGCGTTTTAACATTGTGCTGGTGAAGATGGTGTCTGTATACCCGGATCGTGGTGTCTGCAGGGTTACACGTAAGCGTGATTAATGCAGCAGAAGCCTTTGGTATTGAAGTGCCTGTTGAATTTGTTCCGTTGAACTGGCATATCGCTGTGATGAATTTTTATGTGTACCACCGCAGCGGTATTATGCCGATTACCCCTGGATGGTATGCCGGTTAATGGTGGTCAGATTGGTCCGATTACCAAAAAATTTGGATGGCTATTGGCAATGCATTATGATGCAGCCTATAGCTTTGAAATTGATTATAATGAACGCAATGGATCCCATCATCATCATCATCATTGAGGATCC

← *AspTer* 6xHis BamHI

AAAGCAATCGACGTGGTGATGAATGAAGGGCTGGCTGAGAACGTCCGCCCTTCCCCCGTTTCGAGGAAAGGCTGAAACATATCGCCGAGCGCCG
AACATCGGTGAATATCGCGGCATCGGCTTCATGTGGGCGCTGGAGGCTGTCAAGGACAAGGCAAGCAAGACGCCGTTTCGACGGCAACCTGTGGTTCAGC
GAGCGTATCGCAATACCTGCACCGATCTGGGGCTGATTTGCCGGCGCTTGGTCAGTCCGTCGCTTTGTCCGCCCTTTATCCTGACCGAGGGCGAGATG
GATGAGATGTTTCGATAAACTCGAAAAAGCCCTTGATAAGGTCTTTGCCGAGGTTGCCTGACTCGAG
← *VfH6* *BamHI*

H I.1.9 ‘Core modules’ constructed for the coexpression of pathway enzymes

H I.1.9.1 pOPE: *pKA1_alkJ:fsa1-A129S*

pOPE was constructed by employing SLICE (G VII.1.1.1) to assemble the linear fragments coding for the *pKA1_alkJ* backbone and the *fsa1-A129S* insert in operon configuration (Figure D-25A). The plasmid map (Figure H-17) and the pathway coding region are given below.

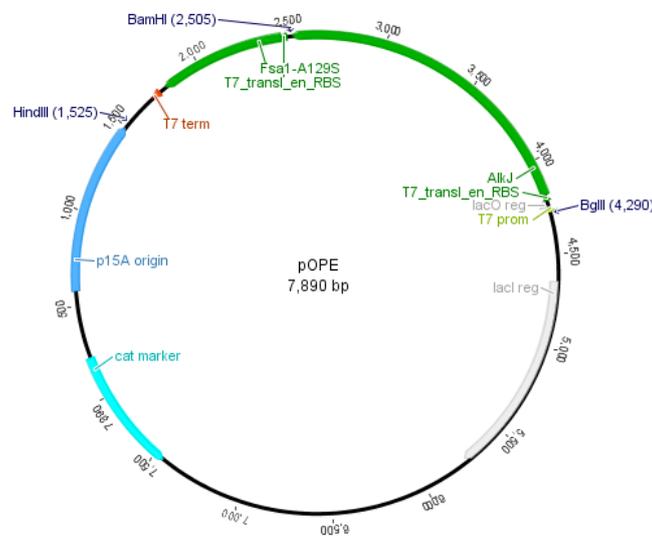


Figure H-17. pOPE. The vector assembled in this thesis contains the operon encoding *alkJ* and the *fsa1-A129S*. The plasmid features a p15A ORI (compatible with pBR322 and CloDF13 ORIs) and conveys resistance to Cam.

BglII *P_{T7}* *lacO* *RBS*
AGATCTCGATCCC GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCGCCCTCTAGAAATAATTTTGT**TAACTTTAAGAAGGAG**
AlkJ → ← *AlkJ* *BamHI* *RBS* *Fsa1-A129S* →
ATATACATATGTAC...//...ATGTAAGGATCCGGCTGCTAACTCTAGAAATAATTTTGT**TAACTTTAAGAAGGAG**ATATACCTGGAACTGTATCTGGACT
TCAGACGTTGTTGCGGTGAAGGCGCTGTCACGTATTTTCCGCTGGCGGTGTGACCACTAACCCAAGCATTATCGCCGCGGGTAAAAACCGCTGGATGT
TGTGCTTCCGCAACTTCATGAAGCGATGGGCGGTGAGGGCGCTGTTGCCAGGTAATGGCTACCACTGCCGAAGGGATGGTTAATGACGCGCTTAAG
CTGCGTTCTATTATTGCGGATATCGTGGTGAAGTTCCGGTGACCGCCGAGGGGCTGGCAGCTATTAAGATGTTAAAAGCGGAAGGGATTCCGACGCTGG
GAACCGCGGTATATGGCGCAGCACAAGGGCTGCTGTCGGCGCTGGCAGGTGCGGAATATGTTTCGCCTACGTTAATCGTATTGATGCTCAGGGCGGTAG
CGGCATTACAGACTGTGACCGACTTACACCGATTATTGAAAATGCATGCGCCGAGGCGAAAGTGCTGGCAGCGAGTTCAAACCCCGCTCAGGCGCTG
GACTGCTTACTGGCAGGATGTGAATCAATTACTCTGCCACTGGATGTGGCACAACAGATGATTAGCTATCCGGCGGTTGATGCCGCTGTGGCGAAGTTTGA
GCAGGACTGGCAGGGAGCGTTTGGCAGAACGTCGATTAAAGCTGAGTTGGCTGCTGC
← *Fsa1-A129S*

H I.1.9.2 pPOP: *pKA1_alkJ:fsa1-A129S*

pPOP was constructed by employing FC (G VII.1.2.1) to assemble the linear fragments coding for the *pKA1_alkJ* backbone and the *fsa1-A129S* insert including an individual *P_{T7}* in pseudo-operon configuration (Figure D-25B). The plasmid map (Figure H-18) and the pathway coding region are given below.

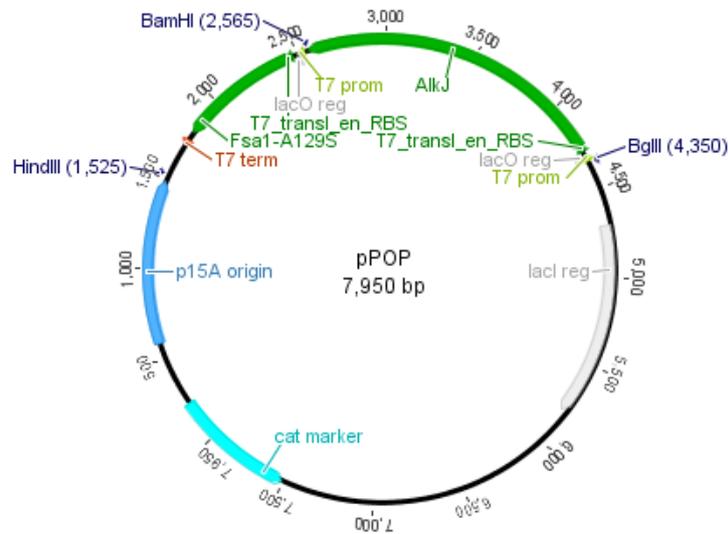


Figure H-18. pPOP. The vector assembled in this thesis contains the *alkJ* and the *fsa1-A129S* gene in pseudo-operon configuration. The plasmid features a p15A ORI (compatible with pBR322 and CloDF13 ORIs) and conveys resistance to Cam.

```

BamII      PT7      lacO      RBS
AGATCTCGATCCCGCAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTAGAAATAATTTGTTAACTTTAAGAAGGAG
      AlkJ →      ← AlkJ BamHI      PT7      lacO
ATATACATATGTAC...//...ATGTAAGGATCCGGCTGCTAACCGATCCCGCAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTCT
      RBS      Fsa1-A129S →      ← Fsa1-A129S
CTAGAAATAATTTGTTAACTTTAAGAAGGAGATATACCATGGAAGCTG...//...TCGATTAAAGCTGAGTTGGCTGCTGC
    
```

H I.1.9.3 pMON1: *pKA1_alkJ::B0011::fsa1-A129S*

pMON1 was constructed by utilizing a unique *Bam*HI restriction site to insert the synthetic terminator B0011 between the *alkJ* and the *fsa1-A129S* gene (G VII.1.3). The target vector contains the two genes in monocistronic configuration (**Figure D-25C**). The plasmid map (**Figure H-19**) and the pathway coding region are given below.

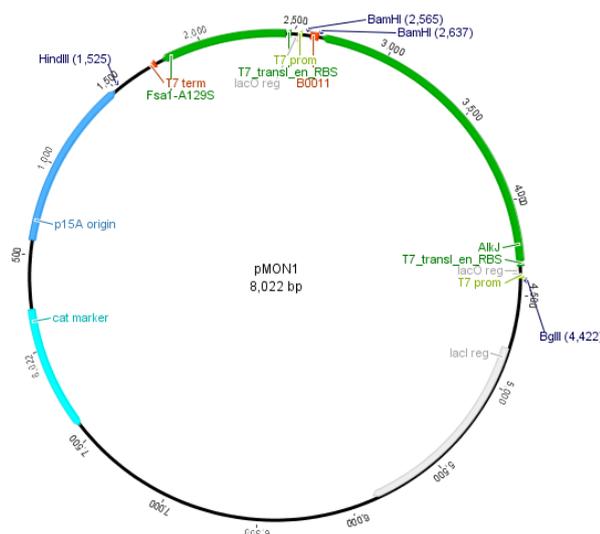


Figure H-19. pMON1. The vector assembled in this thesis contains the *alkJ* and the *fsa1-A129S* gene in monocistronic configuration by insertion of the B0011 terminator between the two coding regions. The plasmid features a p15A ORI (compatible with pBR322 and CloDF13 ORIs) and conveys resistance to Cam.

BallI P₇₇ lacO RBS
 AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTAACCTTAAGAAGGAG
 AlkJ → ← AlkJ *BamHI* 10 bp B0011 10 bp *BamHI*
 ATATACATATGTAC...//...ATGTAAAGGATCCGGCTGCTAACAAATAATAAAAAAGCCGGATTAATAATCTGGCTTTTATATTCTCTGTTAGCAGCCGGATC
 P₇₇ lacO
 CGGCTGCTAACCGATCCCGCGAAATGGCTGCTAACCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGA
 RBS *Fsa1-A129S* → ← *Fsa1-A129S*
 AATAATTTTGTTAACCTTAAGAAGGAGATATACCATGGAAGCTG...//...TCGATTAAAGCTGAGTTGGCTGCTGC

H I.1.9.4 pMON4: *pKA1_alkJ::B0014::fsa1-A129S*

pMON4 was constructed by utilizing a unique *BamHI* restriction site to insert the synthetic terminator B0014 between the *alkJ* and the *fsa1-A129S* gene (G VII.1.3). The target vector contains the two genes in monocistronic configuration (Figure D-25C). In pMON4, the B0011 sequence is replaced by B0014 but otherwise resembles the plasmid map shown in Figure H-19. The pathway coding region in pMON4 is given below.

BallI P₇₇ lacO RBS
 AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTAACCTTAAGAAGGAG
 AlkJ → ← AlkJ *BamHI* 10 bp B0014
 ATATACATATGTAC...//...ATGTAAAGGATCCGGCTGCTAACCACTACACTGGCTCACCTTCGGGTGGCCCTTCTGCGTTTATACTAGAGAGAGAATATAAA
 10 bp *BamHI*
 AAGCCAGATTATAATCCGGCTTTTTATTATTGTTAGCAGCCGGATCCGGCTGCTAACCGATCCCGCGAAATGGCTGCTAACCGATCCCGCGAAATTA
 P₇₇ lacO RBS *Fsa1-A129S* →
 TACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTAACCTTAAGAAGGAGATATACCATGGAAGCTGATCTG.../
 ← *Fsa1-A129S*
 /...AGAACGTCGATTAAAGCTGAGTTGGCTGCTGC

H I.1.9.5 pMON5: *pKA1_alkJ::B0014₅₀::fsa1-A129S*

pMON5 was constructed by utilizing the unique *BamHI* restriction site to insert the synthetic terminator B0014 with flanking 44 bp spacer sequences between the *alkJ* and the *fsa1-A129S* gene (G VII.1.3). The target vector contains the two genes in monocistronic configuration and ligation of the T_{Syn} leaves a 'scar' sequence (Figure E-1B). pMON5 contains the B0014 terminator with long flanking spacer sequences but otherwise resembles the plasmid map shown in Figure H-19. The pathway coding region in pMON5 is given below.

BallI P₇₇ lacO RBS
 AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTAACCTTAAGAAGGAG
 AlkJ → ← AlkJ *Scar* 44 bp B0014
 ATATACATATGTAC...//...ATGTAAAGGATCTAATTCGAGCTCGGTACCTCGGAATGCATCTAGATATCGGATCCTCACTGGCTCACCTTCGGGTGGCCCT
 44 bp
 TTCTGCGTTTATACTAGAGAGAGAATATAAAAAGCCAGATTATAATCCGGCTTTTTATTATTGGATCCCGGGCCGCTGACTGCAGAGGCCTGCAT
Scar P₇₇ lacO
 GCAAGCTTGGGATCTGGCTGCTAACCGATCCCGCGAAATGGCTGCTAACCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATA
 RBS *Fsa1-A129S* → ← *Fsa1-A129S*
 ACAATTCCCCTCTAGAAATAATTTTGTTAACCTTAAGAAGGAGATATACCATGGAAGCTG...//...TCGATTAAAGCTGAGTTGGCTGCTGC

H I.1.9.6 pMON6: *pKA1_alkJ::B0014₁₀₀::fsa1-A129S*

pMON6 was constructed by utilizing the unique *BamHI* restriction site to insert the synthetic terminator B0014 with flanking 94 bp spacer sequences between the *alkJ* and the *fsa1-A129S* gene (G VII.1.3). The target vector contains the two genes in monocistronic configuration and ligation of the T_{Syn} leaves a 'scar' sequence (Figure E-1B). pMON6 contains the B0014 terminator with long flanking spacer sequences but otherwise resembles the plasmid map shown in Figure H-19. The pathway coding region in pMON6 is given below.

```

BglII          PT7          lacO          RBS
AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTGTTAACTTTAAGAAGGAG
      AlkJ →          ← AlkJ  Scar
ATATACATATGTAC...//...ATGTAAGGATCTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGCCAGTGAATTCGAGCTCGGTACCT
                                         B0014
GCGAATGCATCTAGATATCGGATCCTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATATACTAGAGAGAGAATATAAAAAGCCAGATTATTA
                                         94 bp
ATCCGGCTTTTTTATTATTTGGATCCCGGGCCCGTGCAGTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGT
      Scar
TATCCGCTCAAAATGGATCTGGCTGCTAACCGATCCCGCGAAATGGCTGCTAACCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCG
                                         PT7          lacO
GATAACAATTCCCCTCTAGAAATAATTTGTTAACTTTAAGAAGGAGATATACCATGGAAGCTG...//...TCGATTAAAGCTGAGTTGGCTGCTGC
                                         RBS          FsaI-A1295 →          ← FsaI-A1295

```

H 1.1.9.7 pJRE: *pKA1_fucA::alkJ*

The plasmid was constructed by FC (G VII.2.2) and assemble the linear fragments coding for the *pKA1_fucA* backbone and the *alkJ* insert in pseudo-operon configuration (Figure D-31D). The plasmid map (Figure H-20) and the pathway coding region are given below.

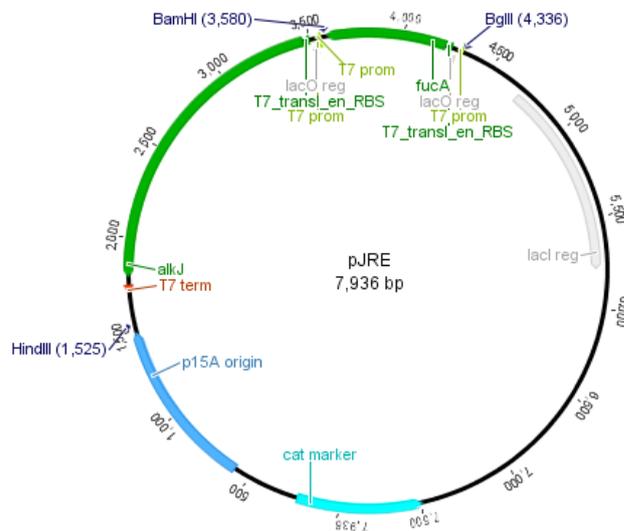


Figure H-20. *pKA1_fucA::alkJ*. The vector assembled in this thesis contains the *fucA* and the *alkJ* gene in pseudo-operon configuration. The plasmid features a p15A ORI (compatible with pBR322 and CloDF13 ORIs) and conveys resistance to Cam.

```

BglII          PT7          lacO          RBS
AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTGTTAACTTTAAGAAGGAG
      FucA →
ATATACATATGGAACGAAATAAACTTGCTCGTCAAGATTATTGACACTTGCCTGGAATGACCCGCTGGGACTGAACCAGGGGACAGCGGGGAACGTCAG
TGTACGTTATCAGGATGGGATGCTGATTACGCCTACAGGCATTCCATATGAAAAACTGACGGAGTCGCATATTGCTTTATTGATGGCAACGGTAAACATG
AGGAAGGAAAGCTCCCCTCAAGCGAATGGCGTTTCCATATGGCAGCCTATCAAAGCAGACCGGATGCCAACGCGGTTGTTACAATCATGCCGTTCAATGC
ACGGCAGTTTCCATTCTTAACCGATCGATCCCCGCTATCACTACATGATTGCGGCGGCTGGCGGTAATTCTATTCCTTGC GCGCCTTATGCGACCTTTGGAA
CACGCGAATTTCTGAACATGTTGCGCTGGCTCTCAAAAATCGTAAGGCAACTTTGTTACAACATCATGGGCTTATCGCTTGTGAGGTGAATCTGGAAAAA
GCGTTATGGCTGGCGCATGAAGTTGAAGTGTGGCGCAACTTTACCTGACGACCCTGGCGATTACGGACCCGGTGCCAGTGCTGAGCGATGAAGAGATTG
                                         ← FucA  BamHI
CCGTAGTGCTGGAGAAATCAAACCTATGGGTTACGAATTGAAGAGTAAGGATCCGGCTGCTAACAAAGCCCGAAAGGATAATACGACTCACTATAGG
      lacO          RBS          AlkJ →          ← AlkJ
GGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTGTTAACTTTAAGAAGGAGATATACATATGTAC...//...ATGTAAGCTGAGTTGGCTGCT
                                         T7
GCCACCCTGAGCAATAACTAGCATAACCCCTGGGGCTCTAACCGGCTTGTAGGGGTTTTTGGCTGA

```

H I.1.10 Other enzyme-coding genes and gene modifications

H I.1.10.1 *DhaK*: pRSETa_*dhaK*

The *dhaK* gene was subcloned into pRSETa by *NdeI*/*Bam*HI and provided by E. García-Junceda from the University of Madrid, Spain. The plasmid map (Figure H-21) and the sequence are given below.

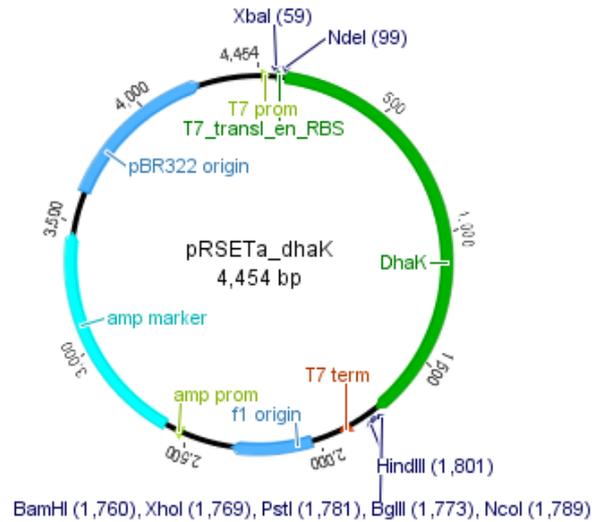


Figure H-21. pRSETa_*dhaK*. The vector contains the *dhaK* gene, a pBR322 ORI (compatible with p15A and CloDF13 ORIs), and conveys resistance to Amp.

NdeI *DhaK* →

```

CATATGTCCTCAATTCTTTTTTAATCAACGCACCCATCTCGTGAGCGACGTCATCGACGGTACGATTATCGCCAGCCCGTGGAAATAACCTGGCGCGTCTGGAA
AGCGATCCGGCCATTGCGATCGTGGTCCGTCGTGACCTCAACAAAATAACGTGGCGGTAATTCGGCGGTGGTTCAGGGCAGCAACCCGCGCACGTTG
GGTTTATCGGTAAGGCATGCTAACCGCTGCGGTTTGC GGCGACGTTTTTCGCTTCCCCGAGCGTGGATGCGGTACTGACCGCCATCCAGGCGGTAACCGGT
GAGGCGGGCTGTTTATTGATCGTGAAAAATTACACCGGTGACCGTCTTAATTTTCGGTCTCGCCGCGAGAAAGCCCGTGCCTTGGTTACAACGTTGAAAT
GCTGATTGTTGGCAGCAGACATCTCCTGCCTGATAACAAACACCCACGCGCATTGCGGGAACCATCTGGTGCATAAAATCGCAGGCTATTTTGCCGAAC
GCGGCTACAACCTGCCACCGTCTCGGTGAAGCGCAGTACGCGGCAATAACACCTTCAGCTGGGCGTTCGCGTTTCCAGCTGTCTGCCGAAGAA
GCCGACGCCGCCCGTATCATCCGGCCACGCGGAACTGGGATGGGCATTACCGCGAACCAGGCGCATCGGTTATCGACACCCAGAACAGTGGCG
AGGTGGTGAACCTGATGGTGGATAAGCTGATGGCAGCCCTGCCTGAAACCGGCCGTCTGGCGGTGATGATTAACAATCTTGGCGGCGTTTCTGTTGCCGA
AATGGCCATCATTACCGCGAAGTGGCCAGCAGCCCGTGCACCCACGATCGACTGGCTGATTGGCCCGGCTCACTGGTACCAGCTCTGGATATGAAAA
GCTTTTCACTGACGGCCATCGTCTGGAAGAAAGCATGAAAAAGCGTTACTACCGAGGTGGAACACGCAACTGGCCGACGCCGTTCCCGCGGTGA
AATCAGTTGTGTACCATCATCTCAGCGTAGCGCAGCGTGGAAATTCAGCCTTCGCGAAGCCATGGTGGCCGGATTGTTGGAATTGTCACCACAACCC
TTCCGATCTGGAGACTCATTTAATGCGCTGGACGCCAAAGTCGGCGATGGCGATACCGGTTTCAGCCTTTCGCGTGGCGCGCGTGAAATTGCCAGTCTG
TTGCATCGCCAGCAGTTGCCGCTGGATAACCTTGCACGCTGTTTCGCGCTGATTGGCGAAGCTCTGACCGTAGTGATGGGTGGTTCCAGCGGTGTGCTGAT
GTCTATTTTCTTACCGCTGCGGGGAGAACTGGAACAGGGAGCTAGCGTTGCCGAATCCCTGAATACGGGACTGGCGCAGATGAAGTTCTACGGCGGC
GCAGACGAAGGCGATCGCACCATGATTGATGCGCTGCAACCAGCCCTGACTTCGCTGCTCAGCAGCCGCAAATCTGCAGGCCGATTGACGCCGCGC
AAGCGGGAGCCGAACGAACCTGTTTGTGAGCAAAGCCAATGCCGGTTCGCGCATCGTATCTCAGCAGCGAAAGCCTGCTCGGAAATATGGACCCCGGCGC
GCACGCCGTAGCGATGGTGTAAAGCGCTAGCGGAGAGTGAGCTGGGCTAA TGACTCGAG
    
```

← *DhaK* *Bam*HI

H I.1.10.2 BLAST analysis and reannotation of pET22b(+)_{*dhaK*}

The putative *dhaK* gene subcloned into pET22b(+) according to the publication by Daniel *et al.*^[539] was submitted to BLAST analysis. Sequence identity to the *dhaK* sequence utilized by Itturate *et al.*^[363] was only 93%. BLAST analysis results in the annotation of a cyclopropane fatty acid synthase from *C. freundii* and other related genes (Figure H-22).

Sequences producing significant alignments:

Select: All None Selected: 0

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	Citrobacter freundii cyclopropane fatty acid synthase (cfa) gene, partial cds; dihydroxacetone kinase (dhaK), glycerol dehydrogenase (dhaD), transcriptional activator (d	3064	3064	100%	0.0	100%	U09771.1
<input type="checkbox"/>	Citrobacter braakii strain FDAARGOS_290, complete genome	2931	2931	100%	0.0	99%	CP022049.1
<input type="checkbox"/>	Citrobacter freundii strain SL151, complete genome	2433	2433	100%	0.0	93%	CP016952.1
<input type="checkbox"/>	Citrobacter freundii strain CAV1741, complete genome	2433	2433	100%	0.0	93%	CP011657.1
<input type="checkbox"/>	Citrobacter freundii strain CAV1321, complete genome	2433	2433	100%	0.0	93%	CP011612.1
<input type="checkbox"/>	Citrobacter freundii strain CECT 4626 dihydroxacetone kinase (dhaK) gene, complete cds	2433	2433	100%	0.0	93%	DQ473522.1
<input type="checkbox"/>	Citrobacter freundii strain B38, complete genome	2422	2422	100%	0.0	93%	CP016762.1
<input type="checkbox"/>	Citrobacter freundii strain 18-1, complete genome	2416	2416	100%	0.0	93%	CP022273.1
<input type="checkbox"/>	Citrobacter freundii strain 705SK3, complete genome	2410	2410	100%	0.0	93%	CP022151.1
<input type="checkbox"/>	Citrobacter sp. 30_2 genome	2405	2405	100%	0.0	93%	CP022311.1
<input type="checkbox"/>	Citrobacter freundii strain P10159, complete genome	2388	2388	100%	0.0	93%	CP012554.1
<input type="checkbox"/>	Citrobacter freundii strain BD genome	2344	2344	100%	0.0	92%	CP018810.1
<input type="checkbox"/>	Citrobacter braakii strain FDAARGOS_253, complete genome	2338	2338	100%	0.0	92%	CP020448.1
<input type="checkbox"/>	Citrobacter werkmani strain BF-6, complete genome	2106	2106	100%	0.0	90%	CP019986.1
<input type="checkbox"/>	Klebsiella oxytoca KQNH1, complete genome	2100	3217	100%	0.0	90%	CP008788.1

Figure H-22. BLAST analysis of wrongly annotated *DhaK*-coding sequence. BLAST search performed and figure created from <https://blast.ncbi.nlm.nih.gov/Blast.cgi>.

The sequence inserted into pET22b(+) probably encodes the *cfaS* gene from *C. freundii*:

NdeI putative *cfaS* →

CATATGTCTCAATTCTTTTTAACCAACGCACCCATCTTGTGAGCGACGTCATCGACGGGGCGATTATCGCCAGCCCATGGAATAACCTGGCGCGTCTGGAA
 AGCGATCCGGCCATTCGCATCGTGGTCCGTCGTGACCTTAATAAAAATAACGTAGCGGTCATTTCCGGCGCGGTTCCGGGACACGAAACCCGCGCACGTTG
 GGTTTATCGGTAAGGCATGCTAACCGCTGCGGTCTGCGGCGACGTTTTCCGCTCCCGAGCGTGGATGCTGACTGACCCGCGATTAGCGCGGTGACCCGG
 TGAGGCTGGCTGTTTGTGATTGTGAAAACTACACCGGTGACCGTCTTAATTCGCTCTCGCCGCGAGAAGGCGCGTGCCTTGGCTATAACGTTGAAA
 TGCTGATTGTCGGCGACGACATCTCCCTGCCGGATAACAAACACCCACGTGGCATTGCGGGAATATCCTGGTGCATAAAATCGCAGGCTATTTGCCGAA
 CGCGGCTATAACCTCGCCACCGTCTGCGTGAAGCGCAGTACGACGCCAGCAACACCTTTAGCCTGGGCGTAGCGCTTCCAGCTGTCTATCGCCGCAAGA
 AACCGACGCAACCCCTGCTATCATCCGGGTATGCGGAGCTGGGTATGGGAATCACGGCAACAGCGCATCGTTATCGACACCCAAAACAGTGCG
 CAAGTGGTAAACCTGATGTTGGATAAACTGCTGGCCGCCCTGCTGAAACCGGTCGCTGGCGGTGATGATTAATAATCTGGCGCGGTTCCGTGGCCG
 AAATGGCCATCATACCCCGCAACTCGCCAGCAGCCCGCTGCACTCGCGTATCGACTGGCTAATTGGCCCGCCTCGCTGGTACCGCGCTGGATATGAAA
 GGCTTCTCACTGACGGCCATCGTGTGGAAGAGAGCATCGAAAAGCACTGCTCACCGAAGTGGAAACAGCAACTGCGCGACGCGGTTCCACCCGCGTG
 AAATCACTGCGTAGTGTCTACGCTAGCAGCCCGCTGGAATTCAGCCTTCGCGAAACGCCCTGGTGGCCGGGATTGTGGAGCTGGTACCGCAACC
 CTTCCGATCTGGAGACTCATGTAATGCGCTGGACGCAAAAGTGGCGATGGCGATACCGGTTGACCTTTGCCGCGCGGCGCGTGAATTTGCCAGCCT
 GCTGCATCGCCAGCAGCTGCCGCTGAATAACCTTGCCACGCTGTTGCGCTGATTGGCGAACGCTGACCGTGGTGTGGGCGGTTCCAGCGGTGTGCTG
 ATGCAATCTTTTACC CGCCGCGGAGAACTGGAACAGGGCGCTAACGTTGTCGAAGCGCTAAATACGGGGCTGGCGCAGATGAAGTTCTACGGCG
 GCGCAGACGAAGCGATCGCACGATGATTGATGCGCTGCAACCGCCCTGACCTCGTCTCGCACAGCCGAAAAATCTGCAGGCCGATTCGACGCCG
 GCAAGCGGGAGCCGAACGAACTGTTTGTGAGCAAAGCAATGCGGGTGGCGATCGTATCTGAGCAGCGAAAGCCTGCTCGAAATATGACCCCGG
 CGCGCAGCCCTAGCGATGTTTAAAGCGTAGCGGAGAGTGAAGTGGGCTAACTCGAG

← putative *cfaS* *XhoI*

H I.1.10.3 *PDC_{Ap}*: pET22b(+)_{*pdC_{Ap}*}

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *pdC_{Ap}* gene was subcloned into pET22b(+) by *NdeI/XhoI*.

NdeI *PDC_{Ap}* →

CATATGACCTATACTGTTGGCATGTATCTTGAGAACGCCTTGTACAGATCGGGCTGAAGCATCACTTCGCCGTGGGCGGCGACTACAATCTCGTTCTTCTG
 GATCAGTTGCTCCTCAACAAGGACATGAAACAGATCTATTGCTGCAATGAGTTGAAGTGTGGCTTCAGCGCGAAGGCTACGCCGTTCTAACGGGGCTGC
 GGCAGCGGTTGTCACCTTCAGCGTTGGCGCAATTTCCGCCATGAACGCCCTCGGCGGCGCCTATGCCGAAACCTGCCGTTATCTGATTTCCGGCGCGC
 CCAACAGCAATGATCAGGGCACAGGTCATATCCTGCATCACACAATCGGCAAGACGGATTACAGCTACCAGCTTGAATGGCCGTCAGGTCACCTGTGCC
 GCCGAAAGCATTACCGACGCTCACTCCGCCCGGCAAGATTGACCACGTCATTCGCACGGCGCTGCGGAGCGTAAGCCGCCATCTGGACATCGCGT
 GCAACATTGCCTCCGAGCCCTGCGTGGCGCTGGCCCTGTCAGCAGCCTGCTGTCGAGCCTGAAATCGACCACACGAGCCTGAAGGCCGCGAGTGGACGC
 CACGGTTGCTTGTGAAAAATCGGCCAGCCCCGTCATGCTGCTGGGAGCAAGCTGCGGGCCGCAACGCACTGGCCGCAACCGAAACGCTGGCAGAC
 AAGCTGCAATGCGCGGTGACCATCATGGCGGCGCGAAAGGCTTTTTCCCGAAGACCACGCGGTTTCCGCGCCTGTACTGGGCGAAGTCTGAACC
 CCGCGTGCAGGAACTGGTGGAGACTCCGACGCACTGCTGTGCATCGCCCCGATTCAACGACTATTCAACAGTGGCTGGTGGGCGATGCCAAAGGG
 CCCAATGTGATTGCTGAGCCGACCGGTAACGGTGCATGGCCGCGCCTATGACGGCTTTACCCTGCGCGCCTTCTGCAGGCTCTGGCGGAAAAAG


```
GCTTTGCAAAACTGACTGAGCAGGTCAAAGAACAAGTTCAAAAATCGGCATTAGCTTCCCATGCCGGACGTAAGCTGAGCTTGAAGGAGATATCTT
CGCTCCATCGGCGATTGCGTTTGACGAACAATCAGCAACACCTGTAGAAATGTCAGCAGAAAAAGTAAAAGAAACGGTCCAGGAGTTCAAGCAAGCGGCT
GCCCGCGCAAAAAGAAGCCGGCTTTGATGTGATTGAAATTCATCGCGCCACGGATATTTAATTCATGAATTTTGTCTCCGCTTTCAACCATCGAACAGAT
GAATATGGCGGCTCACCTGAAAACCGCTATCGTTTCTGAGAGAGATCATTGATGAAGTCAAACAAGTATGGGACGGTCTTTATTTGTCGGTGTATCTGCT
TCTGACTACACTGATAAAGGCTTAGACATTGCCGATCACATCGTTTTGCAAAATGGATGAAGGAGCAGGGTGTGACTTAATTGACTGCAGCTCAGGGC
CCTTGTTCACGCAGACATTAACGTATCCCTGGCTATCAGGTCAGCTTCGCTGAGAAAATCCGTGAACAGGCGGACATGGCTACTGGTCCGCTCGGCATGA
TTACAGACGGTTCAATGGCTGAAGAAATCTGCAAAACGGACGTGCCGACCTCATTTTATCGGCAGAGAGCTTTGCGGGATCCATTTTTGCAAGAACT
GCTGCGAAACAGCTCAATACAGAGATTCCGGCCCTGTTCAATACGAAAGAGGCTGGTAA
← YqjM
```

H I.1.10.6 *YqjM*: pET22b(+)_{*YqjM*}

The plasmid map of the backbone is shown in **Figure H-4** (H I.1.1.4). The *yqjM* gene was subcloned into pET22b(+) by *NdeI/BamHI*.

NdeI YqjM →

```
CATATGCCCAGAAAATTATTACACCTATTACAATTAAGATATGACGTTAAAAACCGCATTGTCATGTCGCCAATGTGCATGTATTCTTCTCATGAAAAG
GACGGAAAATTAACACCGTTCCACATGGCACCATTACATATCGCGCGCAATCGGCCAGGTCGGACTGATTATTGTAGAGGCGTCAGCGGTTAACCCCAAGG
ACGAATCACTGACCAAGACTTAGGCATTTGGAGCGACGAGCATATTGAAGGCTTTGCAAAACTGACTGAGCAGGTCAAAGAACAAGGTTCAAAAATCGGC
ATTCAGCTTCCCATGCGGACGTAAGCTGAGCTTGAAGGAGATATCTTCGCTCCATCGCGGATTGCGTTTGACGAACAATCAGCAACACCTGTAGAAAT
GTCAGCAGAAAAAGTAAAAGAAACGGTCCAGGAGTTCAGCAAGCGGCTGCCCGCGCAAAAAGAAGCCGGCTTTGATGTGATTGAAATTCATCGCGCGCA
CGGATATTTAATTCATGAATTTTTGTCTCCGCTTTCAACCATCGAACAGATGAATATGGCGGCTCACCTGAAAACCGCTATCGTTTCTGAGAGAGATCATT
GATGAAGTCAAACAAGTATGGGACGGTCTTTATTTGTCCGTGTATCTGCTTCTGACTACACTGATAAAGGCTTAGACATTGCCGATCACATCGGTTTTGCA
AAATGGATGAAGGAGCAGGGTGTGACTTAATTGACTGCAGCTCAGGCCCTTTGTTACGCAGACATTAACGTATCCCTGGCTATCAGGTCAGCTTCGC
TGAGAAAATCCGTGAACAGGCGGACATGGCTACTGGTCCGCTCGGCATGATTACAGACGGTTCAATGGCTGAAGAAATCTGCAAAACGGACGTGCCGAC
CTCATCTTTATCGGCAGAGAGCTTTGCGGGATCCATTTTTGCAAGAACTGCTGCGAAACAGCTCAATACAGAGATTCCGGCCCTGTTCAATACGAAAGA
GGCTGGTAAGGATCC
← YqjM BamHI
```

H I.1.11 Cosmids

H I.1.11.1 pGEc47

The pGEc47 cosmid was donated by B. Bühler from the TU Dortmund, Germany and contained the *alkJ* gene amongst others. It was utilized as PCR template for target insert amplification (**Figure H-24**). The sequence coding the *alkBGHJKL* operon is given below. The backbone sequence can be retrieved from Julsing *et al.* (2012).^[289]

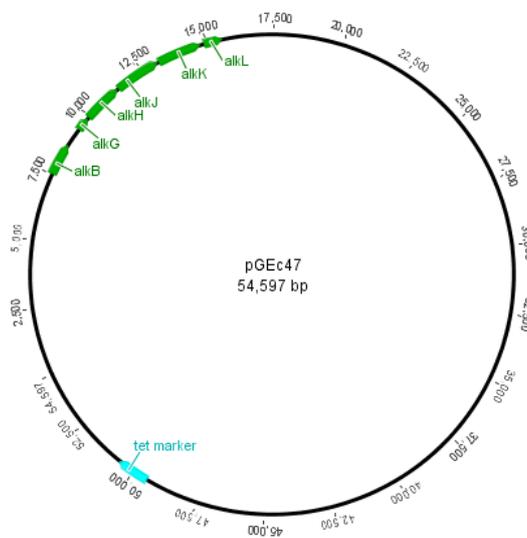


Figure H-24. pGEc47. The cosmid contains the *alkBGHJKL* operon and conveys resistance to Tet.^[289]

5	pKA1_fucA	FC	<i>Insert</i> FucA fwd: GGAGATACATATGATGGAACGAAATAAACTTGCTC FucA rev: GCAGCCGATCCTTACTTACTCTTCAATTCGTAACCC <i>Backbone</i> pKA1FucA fwd: CAAGTTTATTTTCGTTCCATCATATGTATATCTCCTTC pKA1FucA rev: GTTACGAATTGAAGAGTAAGTAAGGATCCGGCTGCTAAC	61.0 48.5	G VI.4.1.2
6	pET26b(+)_ phoN-Sf V78L _{fl}	Florida	PhoN-Se_fl_Ndel fwd: CGGCATATGAAAAGTCGTTATTTAGTATTTTTTC PhoN-Se_fl_XhoI rev: GTTTCTCGAGTCACTTATCATCGTCATCCTTG	65.0	G VI.5.1.1
7	pET26b(+)_ phoN-Sf V78L _{SP}	Florida	PhoN-Se_fl_Ndel fwd: CGGCATATGAAAAGTCGTTATTTAGTATTTTTTC PhoN-Se_w/o tag_XhoI rev: GGTCTCGAGTCAGCTCCTCACGG	65.0	
8	pET26b(+)_ phoN-Sf V78L _{3FLAG}	Florida	PhoN-Se_w/o SP_Ndel fwd: CTGCATATGAAATATACATCAGCAGAAACAGTGC PhoN-Se_fl_XhoI rev: GTTTCTCGAGTCACTTATCATCGTCATCCTTG	65.0	
9	pCOM_alkL	Florida	AlkL_Ndel fwd: GCGCATGAGTTTTTCTAATTATAAAG AlkL_BamHI rev: GCGGGATCCTTAGAAAACATATG	48.0	G VI.10.2
10	pOPE	SLICE	<i>Insert</i> OPE1A129SRBS fwd: GGATCCGGCTGCTAACTCTAGAAAATAATTTTGTTAAC OPE1A129S rev: CAGCAGCCAACCTCAGCTTTAAATCGACGTTCTGCCAAAC <i>Backbone</i> OPE1pAJ3A129S fwd: CTAGAGTTAGCAGCCGGATCCTTACATGCAGACAGCTATC OPE1pAJ3A129S rev: GGCAGAACGTCGATTTAAAGCTGAGTTGGCTGCTGCCAC	44.4 57.2	G VII.1.1
11	pPOP	FC	<i>Insert</i> PO1A129SPRM fwd: GTAAGGATCCGGCTGCTAACCGATCCCGCAAATTAATAC OPE1A129S rev: CAGCAGCCAACCTCAGCTTTAAATCGACGTTCTGCCAAAC <i>Backbone</i> PO1pAJ3A129S fwd: GTATTAATTCGCGGGATCGGTTAGCAGCCGGATCCTTAC OPE1pAJ3A129S rev: GGCAGAACGTCGATTTAAAGCTGAGTTGGCTGCTGCCAC	40.0 51.0	G VII.1.2
12	pMON1	Florida	B0011 fwd: CCGGGATCCAGAGAATATAAAAAAGCC B0011 rev: CGCGGATCCAAATAATAAAAAAGCCGG	57.5	G VII.1.3
13	pMON4	Florida	B0014 fwd: CCGGGATCCGGCTGCTAACTC B0014 rev: GGCGGATCCGGCTGCTAACAAATAATAAAAAAG	57.5	
14	pMON5	Florida	MON5_BglII fwd: GCGAGATCTCAGTGAATTCGAGCTCGGTAC MON5_BglII rev: GCGAGATCTGATTACGCCAAGCTTGCATGCAG	57.5	G VII.1.5.1
15	pMON6	Florida	MON6_BglII fwd: GCGAGATCTGATTAAGTTGGGTAACGCCAG MON6_BglII rev: GCGAGATCTGAATTGTGAGCGGATAACAATTC	57.5	
16	pKA1_alkJ::fucA	FC	<i>Insert</i> POPFUC1A fwd: GGAGATACCATGGAACATGGAACGAAATAAACTTG POPFUC1A rev: CAGCAGCCAACCTCAGCTTTATTACTCTTCAATTCGTAAC <i>Backbone</i> pPOPFUC1A fwd: GTTTATTTTCGTTCCATGTTCCATGGTATATCTCCTTC pPOPFUC1A rev: CGAATTGAAGAGTAATAAAGCTGAGTTGGCTGCTGCCAC	54.6 48.8	G VII.2.1
17	pKA1_fucA::alkJ	FC	<i>Insert</i> POPFUCALK fwd: CAAAGCCCGAAAGGATAATACGACTCACTATAGG POPFUCALK rev: CAGCCAACCTCAGCTTTACATGCAGACAGCTATCATG <i>Backbone</i> pPOPFUCALK fwd: CTGTCTGCATGTAAGCTGAGTTGGCTGCTGCCAC pPOPFUCALK rev: GTGAGTCGATTATCTTTTCGGGCTTTGTTAG	54.6 48.8	G VII.2.2

H III Publications resulting from this thesis

1. T. Wiesinger, * T. Bayer, * S. Milker, M.D. Mihovilovic, F. Rudroff
Cell Factory Design and Optimization for the Stereoselective Synthesis of Polyhydroxylated Compounds
ChemBioChem **2017**, accepted (DOI: 10.1002/cbic.201700464)
* with equal contributions
2. T. Bayer, S. Milker, T. Wiesinger, M. Winkler, M.D. Mihovilovic, F. Rudroff
In Vivo Synthesis of Polyhydroxylated Compounds from a “Hidden Reservoir” of Toxic Aldehyde Species
ChemCatChem **2017**, 9 (15), 2913–2919 (DOI: 10.1002/cctc.201700469)
3. T. Bayer, S. Milker, T. Wiesinger, M.D. Mihovilovic, F. Rudroff
Designer Microorganisms for Optimized Redox Cascade Reactions – Challenges and Future Perspectives
Adv Synth Catal **2015**, 357 (8), 1587–1618 (DOI: 10.1002/adsc.201500202)
4. N. Oberleitner, C. Peters, J. Muschiol, M. Kadow, S. Saß, T. Bayer, P. Schaaf, N. Iqbal, F. Rudroff, M.D. Mihovilovic, U.T. Bornscheuer
An Enzymatic Toolbox for Cascade Reactions: A Showcase for an In Vivo Redox Sequence in Asymmetric Synthesis
ChemCatChem **2013**, 5 (12), 3524–3528 (DOI: 10.1002/cctc.201300604)
5. P. Schaaf, V. Gojic, T. Bayer, F. Rudroff, M. Schnürch, M.D. Mihovilovic
Easy Access to Enantiopure (S)- and (R)-alcohols via Combination of Au(III)-catalyzed Alkyne Hydration and Enzymatic Reduction
Submitted **2017**

H IV Curriculum vitae



Born 6 December 1985 in Vienna, Austria

Education

Oct 2017	PhD thesis examination / <i>viva voce</i>
Oct 2012 – Sept 2017	Dissertation at the Institute of Applied Synthetic Chemistry, TU Wien under the supervision of Prof. Marko D. Mihovilovic and Dr. Florian Rudroff: <i>Design, Application, and Optimization of Synthetic Enzyme Cascades in Escherichia coli</i>
Sept 2014	Short internship with Prof. Uwe T. Bornscheuer at the Ernst-Moritz-Arndt University Greifswald, Germany
Jun 2012	Graduation with distinction in the Molecular Biology program at the University of Vienna, Austria
Feb – Jul 2010	Master thesis at the AFFiRiS AG under the supervision of Prof. Thomas Decker from the University of Vienna: <i>In Vitro and In Vivo Characterization of AFFITOPE®s to Treat Alzheimer's Disease</i>
Oct 2005 – Jun 2012	Master studies in Molecular Biology program at the University of Vienna Specializations: immunology, biochemistry, (bio)organic chemistry
Sept 2000 – Jun 2005	Graduation at the HBLVA für Chemische Industrie Rosensteingasse 79, 1170 Vienna, Austria Specializations: biochemistry, biotechnology, genetics Thesis in collaboration with the University of Vienna: <i>Antibody Production against Food Allergens</i>

Work Experience

Jul 2015 – Sept 2017	University assistant including teaching assignment and supervision of BSc and MSc theses at the Institute of Applied Synthetic Chemistry, TU Wien
Oct 2012 – Jun 2015	Project assistant at the Institute of Applied Synthetic Chemistry, TU Wien
Jul – Aug 2012	Internship at Baxter Innovations GmbH, Immunology Department Industriestraße 72, 1220 Vienna, Austria
Aug – Oct 2011	Internship at Baxter Innovations GmbH, Immunology Department Industriestraße 72, 1220 Vienna, Austria
Jun 2008 – Jun 2011	Research associate at AFFiRiS AG, Neurodegenerative Disease Department Karl-Farkas-Gasse 22, 1030 Vienna, Austria
Jul 2007	Short internship at Voith Paper Rolls GmbH & Co. KG Maretgasse 45, 2632 Wimpassing, Austria
Jul – Aug 2006	Internship at A.C.E. Chemical Engineers GmbH Loeschenkohl-gasse 8b/4-5, 1150 Vienna, Austria

Jul 2004	Short internship at A.C.E. Chemical Engineers GmbH Loeschenkohlgrasse 8b/4-5, 1150 Vienna, Austria
Jul 2003	Short internship at Beiersdorf GmbH, QC/QA Department Laxenburgerstraße 151, 1100 Vienna, Austria
Jul 2002	Short internship at Medical-diagnostic Laboratories Mühl-Speiser Clinical Chemistry and Hematology Department Praterstraße 22, 1020 Vienna, Austria

Conference Activities

Oral presentations

Oct 2016	Novel Enzymes 2016 (Groningen, The Netherlands) <i>Substrate Funneling: A Simple and Efficient Carbon Flux Optimization Tool for Synthetic Enzyme Cascades</i>
Jul 2016	Gordon Research Conference: Biocatalysis (Biddeford, ME, USA) Poster award session <i>Convergent Flux via Redox/Substrate Funneling for the In Vivo Preparation of Reactive Aldehyde Species</i>
Jul 2016	OxiZymes Meeting 2016 (Wageningen, The Netherlands) <i>Convergent Flux via Redox/Substrate Funneling for the In Vivo Preparation of Reactive Aldehyde Species</i>

Poster presentations

Jul 2017	Biotrans2017 (Budapest, Hungary) <i>Substrate Funneling Complements Flux Optimization Tools to Maximize Product Titters of In Vivo Cascades</i>
Oct 2016	Novel Enzymes 2016 (Groningen, The Netherlands) <i>Substrate Funneling as a Novel Flux Optimization Tool for Synthetic Enzyme Cascades In Vivo</i>
Jul 2016	Gordon Research Seminar & Conference: Biocatalysis (Biddeford, ME, USA) <i>Convergent Flux via Redox/Substrate Funneling for the Optimization of an Artificial In Vivo Cascade</i>
Jun 2015	Biotrans2015 (Vienna, Austria) <i>Optimization of an Artificial Mini-Pathway for the Production of Polyhydroxylated Compounds on a Genetic Level</i>

Awards

Jul 2016	Gordon Research Conference: Biocatalysis (Biddeford, ME, USA) Poster award
----------	--

Language Skills

German	mother tongue
English	native speaker level
Spanish	basic knowledge

Hobbies & Fields of Interest

Social activities with friends, enthusiastic 365-days-a-year cyclist, sports including volleyball, beach volleyball, swimming, cycling, and yoga

Voluntary volleyball trainer in the LGBTQ+ sports club *Aufschlag Wien* from 2011 to 2016

Since 2015, member of the *Nesterval. Ein Abenteuer in der Stadt* main cast with appearances in (selection):

Sep 2017	<i>Der Fall der Fanny Nesterval</i> (Infoscreen, Vienna, Austria)
Jul 2017	<i>Die Leiden der Gebrüder Grimm</i> (25 th Spectrum Theater Festival, Villach, Austria)
Jun 2017	<i>Zirkus der Schatten</i> (in cooperation with Circus Pikard, Lower Austria)
May 2017	<i>Die wundersame Entführung des Herrn Bruno</i> (Infoscreen, Innsbruck, Austria)
Dec 2016	<i>Die dunkle Weihnacht im Hause Grimm</i> (immersive theater, Vienna, Austria)
Nov 2016	<i>Vienna Secrets</i> (WienTourismus, European tour)
Jun 2016	<i>Die Leiden der Gebrüder Grimm</i> (immersive theater, Bruckneudorf, Austria)
May 2016	<i>Die 7 Sünden</i> (Vienna, Austria)
Mar 2016	<i>Nesterval. Der letzte Ball.</i> (Imagetanz 2016, Vienna, Austria)
May 2015	<i>Das Haus der Hildegard</i> (Vienna, Austria)

H V List of abbreviations

A	absorbance/absorption	Cl-Tet	chlorotetracyclin
ACN	acetonitrile	CoA	coenzyme A
ADAME	12-amino dodecanoic acid methyl ester	CO ₂	carbon dioxide
ADH	alcohol dehydrogenase	CPE	Thank you for your help.
<i>ADH-A</i>	secondary ADH from <i>R. ruber</i>	CRAS	containment of reactive aldehyde species (strain)
<i>ADH-ht</i>	ADH from <i>Geobacillus stearothermophilus</i>	<i>Cre</i>	<i>Cre</i> recombinase
<i>ADH_{Lk}</i>	secondary ADH from <i>L. kefir</i>	CRISPR	clustered regularly interspaced short palindromic repeats
<i>ADH_{Rr}</i>	secondary ADH from <i>R. ruber</i>	CWO	Carina Wokurek, BSc
<i>ADH_{Te}</i>	ADH from <i>Thermoanaerobacter ethanolicus</i>	C18 material	octadecyl carbon chain (C18)-bonded silica
ADK	adenosine kinase	<i>C. necator</i>	<i>Cupriavidus necator</i>
ADP	adenosine diphosphate	<i>C. freundii</i>	<i>Citrobacter freundii</i>
<i>AdS</i>	amorpho-4,11-diene synthetase	dH ₂ O	deionized/distilled water
AFucP	<i>AlkJ/FucA/PhoN-Sf</i> strain	DAME	dodecanoic acid methyl ester
AIM	autoinduction medium	DCPK	dicyclopropylketone
Ala	L-alanine	DCR	2,4-dienoyl-CoA reductase
AIDH	aldehyde dehydrogenase	DDAME	dodecanedioic acid methyl ester
<i>AlkJ</i>	ADH from <i>P. putida</i>	ds	double-stranded
<i>AlkJ_{trnc}</i>	truncated <i>AlkJ</i>	DSB	double strand break
AMP	adenosine monophosphate	DH	dehydrogenase
Amp	ampicillin	DHA	dihydroxyacetone
AOX	alcohol oxidase	DhaK	dihydroxyacetone kinase
Ara	arabinose	DHAP	dihydroxyacetone phosphate
ARE	banana	DMSO	dimethyl sulfoxide
<i>AtoB</i>	acetoacetyl-CoA thiolase	DMSO- <i>d</i> ₆	dimethyl sulfoxide, fully deuterated
ATP	adenosine triphosphate	DNA	deoxyribonucleic acid
AWA	Alex, intern III	<i>dr</i>	diastereomeric ratio
<i>A. pasteurianus</i>	<i>Acetobacter pasteurianus</i>	ECHN	ethyl-(<i>R</i>)-4-chloro-3-hydroxybutanoate
BCIP	5-bromo-4-chloro-3-indoyl phosphate	ECOB	ethyl-4-chlorooxobutanoate
BVMO	Baeyer-Villiger monooxygenase	EDTA	ethylene-diamineteraacetic acid
BWA	Birgit Walder, BSc	<i>ee</i>	enantiomeric excess
<i>B. subtilis</i>	<i>Bacillus subtilis</i>	ELISA	enzyme-linked immunosorbent assay
Cam	chloramphenicol	ERED	enoate reductase
CAR	carboxylic acid reductase	EtOAc	ethyl acetate
<i>CAR_{Mm}</i>	CAR from <i>M. marinum</i>	EtOH	ethanol
<i>CAR_{Ni}</i>	CAR from <i>N. iowensis</i>	<i>E. coli</i>	<i>Escherichia coli</i>
CASTing	combinatorial active-site saturation testing	FACS	fluorescens-activated cell sorting
<i>Cas9</i>	CRISPR-associated protein 9	FAD(H ₂)	flavin adenine dinucleotide (reduced)
CAT	chloramphenicol acetyltransferase	FAQ	frequently asked question
Cbz	carbobenzoyl	FBA	flux balance analysis
CFE	cell-free extract	FC	FastCloning
CFP	cyan fluorescent protein	FDH	formate dehydrogenase
ClChE	chemically inducible chromosomal evolution	FID	flame ionization detector
CHMO	cyclohexanone monooxygenase	<i>Flp</i>	flippase
CHMO _{Acineto}	cyclohexanone monooxygenase from <i>Acinetobacter</i> sp.	FMN	flavin mononucleotide
CH ₂ Cl ₂	dichloromethane	FRK	Franziska, intern II
		FRT	<i>Flp</i> recognition target

<i>FruA</i>	fructose 1,6-bisphosphate aldolase	KO	knock-out
<i>Fsa1</i>	D-fructose-6-phosphate aldolase	Lac	lactose
<i>Fsa1-A129S</i>	<i>Fsa1</i> A129S mutant	LB-0.8G	LB medium containing 0.8% (w/v) glucose
<i>FucA</i>	L-fucose 1-phosphate aldolase	LB medium	lysogeny broth medium
<i>FucP</i>	<i>FucA/PhoN-Sf</i> strain	LEP	late exponential phase
<i>FucY</i>	<i>FucA/YqaB</i> strain	Leti	smart and fabulous
<i>fwd</i>	forward (primer)	L-PAC	L-phenylacetyl carbinol
F-1,6-bisP	fructose 1,6-bisphosphate	<i>L. kefir</i>	<i>Lactobacillus kefir</i>
GAP	glyceraldehyde 3-phosphate	M	(enzyme) module
GBD	GTPase binding domain	MAE	Marcello loves cake too.
GC	gas chromatography	MAGE	multiplex automated genome engineering
GDH	glucose dehydrogenase	MAT	Met adenosyl transferase
<i>GDH_{2xBS}</i>	GDH double mutant from <i>B. subtilis</i>	MCS	multiple cloning site
<i>GDH_{7xBS}</i>	GDH seven-times mutant from <i>B. subtilis</i>	MeOH	methanol
GFP	green fluorescent protein	MeOH- <i>d</i> ₄	methanol, fully deuterated
Glu	glucose	Met	L-methionine
GlyGly	glycylglycine	MevK	mevalonate kinase
GK	glycerol kinase	MFA	metabolic flux analysis
gRNA	guide RNA	MMO	Madlen Mollik, <i>BSc</i>
G6P	D-glucose-6-phosphate	MON	monocistronic
G6P-DH	G6P dehydrogenase	mRNA	messenger RNA
G6Pase	glucose-6-phosphatase	MS	mass spectrometry
G-1,5-L6P	D-glucono-1,5-lactone-6-phosphate	MT	methyltransferase
H	homology	M9	minimal medium
HA	hydroxyacetone	M9-N*	nitrogen enriched minimal medium
HDAME	12-hydroxy dodecanoic acid methyl ester	<i>M. marinum</i>	<i>Mycobacterium marinum</i>
HE	homology extension	NAD ⁺ (NADH)	nicotinamide adenine dinucleotide (reduced)
HHDH	halohydrin dehalogenase	NADP ⁺ (NADPH)	nicotinamide adenine dinucleotide phosphate (reduced)
<i>HL-ADH</i>	horse liver ADH	NaOAc	sodium acetate
HMG	3-hydroxy-3-methylglutaryl	NAPRTase	nicotinic acid phosphoribosyltransferase
<i>HmgS</i>	HMG synthetase	<i>NemR</i>	N-ethylmaleimide reductase
HPLC	high performance liquid chromatography	N-free	nitrogen-free (without nitrogen source)
HS	heat shock	Niko	truly missed in the lab
H ₂ O ₂	hydrogen peroxide	NMR	nuclear magnetic resonance
<i>i</i>	<i>iso</i> -	NOMAD	nucleic acid ordered assembly with directionality
<i>IdK</i>	isopentenyl pyrophosphate isomerase	NOX	NADH oxidase
IEP	intron-encoded protein	NTPs	nucleoside triphosphates
iGI	immobilized glucose isomerase	n.a.	not applicable/not available
IMAC	ion metal affinity chromatography	n.c.	no conversion
IPTG	isopropyl β-D-1-thiogalactopyranoside	n.d.	not detected/not determined
IS	internal standard	<i>N. iowensis</i>	<i>Nocardia iowensis</i>
I<3C	I love cake.	ODAME	12-oxododecanoic acid methyl ester
JRE	Jakob Reiterlehner, <i>MSc</i>	OMG	Oh my god!
KAE	Katja, intern I	OMP	outer membrane protein
Kan	kanamycin	OPE	operon
KFI	Katharina, intern VI	ORF	open reading frame
KI	knock-in	ORI	origin of replication
KKL	Kathi, intern IV	PAGE	polyacrylamide gel electrophoresis
KLD	kinase-ligase- <i>Dpnl</i>		

PAM	protospacer adjacent motif	RBS	ribosome binding site
PAMO	phenylacetone monooxygenase	RC	resting cell
PCR	polymerase chain reaction	RCM	resting cell medium
PDC	pyruvate decarboxylase complex	λ <i>Red</i>	λ <i>Red</i> recombinase
<i>PDC_{Ap}</i>	PDC from <i>A. pasteurianus</i>	<i>rev</i>	reverse (primer)
<i>PDC_{Zm}</i>	PDC from <i>Zymomonas mobilis</i>	Rhm	rhamnose
PDZ	PSD-95/DlgA/ZO-1 domain	<i>RhuA</i>	rhamnulose 1-phosphate aldolase
PEP	phosphoenolpyruvate	RNA	ribonucleic acid
<i>PEST</i>	esterase from <i>Pyrobaculum calidifontis</i>	RNP	ribonucleoprotein
PhoN	nonspecific acidic phosphatase	ROFL	rolling on the floor laughing
<i>PhoN-Se</i>	phosphatase from	rRNA	ribosomal RNA
<i>PhoN-Sf</i>	phosphatase from <i>S. flexneri</i>	RT	room temperature
pP _i	polyphosphate	RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
PP _i	pyrophosphate	<i>R. ruber</i>	<i>Rhodococcus ruber</i>
PPK	polyphosphate kinase	SAM	S-adenosylmethionine
PLP	pyridoxal phosphate	SAH	S-adenosylhomocysteine
<i>PmD</i>	mevalonate pyrophosphate decarboxylase	SAHH	SAH hydrolase
<i>PmK</i>	phosphomevalonate kinase	SCK	Sophie Charlotte Knoll, <i>BSc</i>
pMON	plasmid containing the <i>alkJ</i> and the <i>fsa1-A129S</i> genes in MON configuration (<i>pKA1_alkJ::B001x::fsa1-A129S</i>)	SD	standard deviation
pNPA	p-nitrophenyl acetate	SD sequence	Shine-Dalgarno sequence
pNPP	p-nitrophenyl phosphate	SDS	sodium dodecyl sulfate
pOPE	plasmid containing the <i>alkJ</i> and the <i>fsa1-A129S</i> genes in OPE configuration (<i>pKA1_alkJ::fsa1-A129S</i>)	siRNA	small interfering RNA
POP	pseudo-operon	SH3	<i>Src</i> homology 3 (domain)
pPOP	plasmid containing the <i>alkJ</i> and the <i>fsa1-A129S</i> genes in POP configuration (<i>pKA1_alkJ::fsa1-A129S</i>)	SLIC	sequence- and ligation-independent cloning
PPP	penotse phosphate pathway	SLiCE	seamless and ligation-independent cloning extract
(p)ppGpp	guanosine tetraphosphate / (pentaphosphate)	SMI	Sofia: Katze.
PPtase	phosphopantetheinyl transferase	SP	signal peptide
<i>PPtase_{Ec}</i>	PPtase from <i>E. coli</i>	SPE	solid phase extraction
PrNH ₂	propyl amine	sp.	species
PrOH	propyl alcohol	sRNA	small regulatory RNA
proSAR	statistical analysis of protein sequence activity relationships	ss	single-stranded
PSC	<3	Str	streptomycin
PTDH	phosphite dehydrogenase	<i>S. carnosus</i>	<i>Staphylococcus carnosus</i>
P _{T7}	T7 promoter	<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
P _{tac}	<i>tac</i> promoter	<i>S. enterica</i>	<i>Salmonella enterica</i>
PWO	Patricia Wolf, intern V, made the best Schnitzel for lunch.	<i>S. flexneri</i>	<i>Shigella flexneri</i>
PYC	pyruvic acid carboxylase	t ₀ *	sample taken after mixing (t ≈ 0 h)
Pyr	pyruvate	<i>TagA</i>	tagatose-1,6-bisphosphate aldolase
P3HB	poly-3-hydroxy butyrate	TAR	Ttransformation-associated recombination
<i>P. fluorescens</i>	<i>Pseudomonas fluorescens</i>	TBA	the author of this thesis
<i>P. putida</i>	<i>Pseudomonas putida</i>	TB medium	terrific broth medium
RAMA	rabbit muscle aldolase	TCA	tricarboxic acid
RARE	reduced aromatic aldehyde reduction	TEV	tabacco etch virus
		TIGR	tuneable intergenic region
		T _m	melting temperature
		TMI	triosephosphate isomerase
		tRNA	transfer RNA
		T _{T1}	T1 terminator
		T _{T7}	T7 terminator
		T _{Syn}	synthetic terminator
		TWI	Wiesi: Babysteps.
		UTB	Thank you for your support.
		UTR	untranslated region

<i>VfIH6</i>	ω -TA from <i>Vibrio fluvialis</i>
YFP	yellow fluorescent protein
YT	yeast extract/tryptone
3xFLAG tag	3x FLAG epitope tag
3-TFAP	3-trifluoroacetamido propanal
6xHis tag	6x histidine tag

H VI References

- [1] E. Buchner, *Eur J Inorg Chem* **1897**, *30*, 1110-1113.
- [2] T. Schwann, *Ann Phys Chem* **1837**, *41*, 184-193.
- [3] J. Reiterlehner, *M.Sc. Thesis* **2017**, TU Wien / FH Wien, 1-75.
- [4] L. Rosenthaler, *Biochem Z* **1908**, *14*, 238-253.
- [5] N. Ladkau, A. Schmid, B. Bühler, *Curr Opin Biotechnol* **2014**, *30*, 178-189.
- [6] T. Bayer, S. Milker, T. Wiesinger, F. Rudroff, M. D. Mihovilovic, *Adv Synth Catal* **2015**, *357*, 1587-1618.
- [7] E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv Synth Catal* **2011**, *353*, 2239-2262.
- [8] J. Muschiol, C. Peters, N. Oberleitner, M. D. Mihovilovic, U. T. Bornscheuer, F. Rudroff, *Chem Commun* **2015**, *51*, 5798-5811.
- [9] J. H. Schrittwieser, J. Sattler, V. Resch, F. G. Mutti, W. Kroutil, *Curr Opin Chem Biol* **2011**, *15*, 249-256.
- [10] F. Lopez-Gallego, C. Schmidt-Dannert, *Curr Opin Chem Biol* **2010**, *14*, 174-183.
- [11] L. M. Blank, B. E. Ebert, K. Buehler, B. Bühler, *Antioxid Redox Signal* **2010**, *13*, 349-394.
- [12] H. Zhao, W. A. van der Donk, *Curr Opin Biotechnol* **2003**, *14*, 583-589.
- [13] D. L. Nelson, M. M. Cox, A. L. Lehninger, *Principles of biochemistry*, W.H. Freeman, New York, **2013**.
- [14] N. R. Mohamad, N. H. C. Marzuki, N. A. Buang, F. Huyop, R. A. Wahab, *Biotechnol Biotechnol Equip* **2015**, *29*, 205-220.
- [15] V. J. Jensen, S. Rugh, *Methods Enzymol* **1987**, *136*, 356-370.
- [16] S. K. Ma, J. Gruber, C. Davis, L. Newman, D. Gray, A. Wang, J. Grate, G. W. Huisman, R. A. Sheldon, *Green Chem* **2010**, *12*, 81-86.
- [17] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185-194.
- [18] D. A. Jackson, R. H. Symons, P. Berg, *PNAS* **1972**, *69*, 2904-2909.
- [19] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, R. B. Helling, *PNAS* **1973**, *70*, 3240-3244.
- [20] W. A. Greenberg, A. Varvak, S. R. Hanson, K. Wong, H. Huang, P. Chen, M. J. Burk, *PNAS* **2004**, *101*, 5788-5793.
- [21] D.-K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Y. Chang, S. T. Withers, Y. Shiba, R. Sarpong, J. D. Keasling, *Nature* **2006**, *440*, 940-943.
- [22] J. A. Castillo, J. Calveras, J. Casas, M. Mitjans, M. P. Vinardell, T. Parella, T. Inoue, G. A. Sprenger, J. Joglar, P. Clapés, *Org Lett* **2006**, *8*, 6067-6070.
- [23] C. E. Brocklehurst, K. Laumen, L. La Vecchia, D. Shaw, M. Vögtle, *Org Process Res Dev* **2011**, *15*, 294-300.
- [24] T. Furuya, M. Miura, M. Kuroiwa, K. Kino, *N Biotechnol* **2015**, *32*, 335-339.
- [25] J. A. Doudna, E. Charpentier, *Science* **2014**, *346*.
- [26] T. Schwander, L. Schada von Borzyskowski, S. Burgener, N. S. Cortina, T. J. Erb, *Science* **2016**, *354*, 900-904.
- [27] S. Yoshida, K. Hiraga, T. Takehana, I. Taniguchi, H. Yamaji, Y. Maeda, K. Toyohara, K. Miyamoto, Y. Kimura, K. Oda, *Science* **2016**, *351*, 1196-1199.
- [28] P. Lorenz, J. Eck, *Nat Rev Micro* **2005**, *3*, 510-516.
- [29] A. Liese, K. Seelbach, C. Wandrey, *Industrial biotransformations*, 2nd ed., Wiley-VCH, New York, **2006**.
- [30] A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, *Nature* **2001**, *409*, 258-268.
- [31] M. Höhne, S. Schätzle, H. Jochens, K. Robins, U. T. Bornscheuer, *Nat Chem Biol* **2010**, *6*, 807-813.
- [32] T. Ellis, T. Adie, G. S. Baldwin, *Integr Biol* **2011**, *3*, 109-118.
- [33] D. M. Fowler, C. L. Araya, S. J. Fleishman, E. H. Kellogg, J. J. Stephany, D. Baker, S. Fields, *Nat Meth* **2010**, *7*, 741-746.
- [34] J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi, S. Y. Lee, *Nat Chem Biol* **2012**, *8*, 536-546.
- [35] P. Anastas, N. Eghbali, *Chem Soc Rev* **2010**, *39*, 301-312.
- [36] C. W. Song, J. Lee, S. Y. Lee, *J Biotechnol* **2015**, *10*, 56-68.
- [37] T. Fehér, V. Burland, G. Pósfai, *J Biotechnol* **2012**, *160*, 72-79.
- [38] H. Wang, M. La Russa, L. S. Qi, *Annu Rev Biochem* **2016**, *85*, 227-264.
- [39] J. D. Sander, J. K. Joung, *Nat Biotech* **2014**, *32*, 347-355.
- [40] K. Selle, R. Barrangou, *Trends Microbiol* **2013**, *23*, 225-232.
- [41] W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, *Nat Biotech* **2013**, *31*, 233-239.
- [42] A. M. Kunjapur, Y. Tarasova, K. L. J. Prather, *J Am Chem Soc* **2014**, *136*, 11644-11654.
- [43] S. Wu, Y. Zhou, T. Wang, H.-P. Too, D. I. C. Wang, Z. Li, *Nat Commun* **2016**, *7*, 11917.
- [44] N. Oberleitner, C. Peters, J. Muschiol, M. Kadow, S. Saß, T. Bayer, P. Schaaf, N. Iqbal, F. Rudroff, M. D. Mihovilovic, U. T. Bornscheuer, *ChemCatChem* **2013**, *5*, 3524-3528.
- [45] E. Hao Yu, *Fuel Cells* **2016**, *16*, 517-521.
- [46] F. Baneyx, *Curr Opin Biotechnol* **1999**, *10*, 411-421.
- [47] H. P. Sørensen, K. K. Mortensen, *Microb Cell Fact* **2005**, *4*, 1.
- [48] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol Syst Biol* **2006**, *2*, n/a-n/a.
- [49] D. Nathans, H. O. Smith, *Annu Rev Biochem* **1975**, *44*, 273-293.
- [50] H. P. Sørensen, K. K. Mortensen, *J Biotechnol* **2005**, *115*, 113-128.
- [51] W. E. Bentley, D. S. Kompala, *Biotechnol Lett* **1990**, *12*, 329-334.
- [52] F. Hoffmann, U. Rinas, *Adv Biochem Eng Biotechnol* **2004**, *89*, 73-92.
- [53] D.-E. Chang, D. J. Smalley, T. Conway, *Mol Microbiol* **2002**, *45*, 289-306.
- [54] C. G. Kurland, H. Dong, *Mol Microbiol* **1996**, *21*, 1-4.
- [55] D. S.-W. Ow, D.-Y. Lee, H.-H. Tung, S. Lin-Chao, S. Y. Lee, *Systems Biology and Biotechnology of Escherichia coli*, Springer Netherlands, **2009**.
- [56] F. Silva, J. A. Queiroz, F. C. Domingues, *Biotechnol Adv* **2012**, *30*, 691-708.
- [57] P. K. Ajikumar, W.-H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, *Science* **2010**, *330*, 70-74.
- [58] B. Gasser, M. Saloheimo, U. Rinas, M. Dragosits, E. Rodríguez-Carmona, K. Baumann, M. Giuliani, E. Parrilli, P. Branduardi, C. Lang, D. Porro, P. Ferrer, M. L. Tutino, D. Mattanovich, A. Villaverde, *Microb Cell Fact* **2008**, *7*, 11.
- [59] S. Gräslund, M. Hallberg, P. Nordlund, J. Weigelt, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, *Nat Meth* **2008**, *5*, 135-146.
- [60] B. R. Glick, *Biotechnol Adv* **1995**, *13*, 247-261.
- [61] K. G. Thakur, R. K. Jaiswal, J. K. Shukla, T. Praveena, B. Gopal, *Protein Expr Purif* **2010**, *74*, 223-230.
- [62] M. K. Akhtar, N. J. Turner, P. R. Jones, *PNAS* **2013**, *110*, 87-92.
- [63] K. Napora-Wijata, G. A. Strohmeier, M. Winkler, *Biotechnol J* **2014**, *9*, 822-843.

- [64] W. Finnigan, A. Thomas, H. Cromar, B. Gough, R. Snajdrova, J. P. Adams, J. A. Littlechild, N. J. Harmer, *ChemCatChem* **2017**, *9*, 1005-1017.
- [65] N. H. Tolia, L. Joshua-Tor, *Nat Meth* **2006**, *3*, 55-64.
- [66] H.-C. Tseng, K. L. J. Prather, *PNAS* **2012**, *109*, 17925-17930.
- [67] M. J. Sheppard, A. M. Kunjapur, S. J. Wenck, K. L. J. Prather, *Nat Commun* **2014**, *5*, 5031.
- [68] N. J. Turner, E. O'Reilly, *Nat Chem Biol* **2013**, *9*, 285-288.
- [69] M. D. McMahon, K. L. J. Prather, *Appl Environ Microbiol* **2014**, *80*, 1042-1050.
- [70] J. E. Bailey, *Adv Biochem Eng Biotechnol* **1993**, *48*, 29-52.
- [71] J. D. Keasling, *Trends Biotechnol* **1999**, *17*, 452-460.
- [72] K. L. Jones, S.-W. Kim, J. D. Keasling, *Metab Eng* **2000**, *2*, 328-338.
- [73] T. Wang, X. Ma, G. Du, J. Chen, *Mol Biotechnol* **2012**, *52*, 300-308.
- [74] D. J. Pitera, C. J. Paddon, J. D. Newman, J. D. Keasling, *Metab Eng* **2007**, *9*, 193-207.
- [75] P. Xu, A. Vansiri, N. Bhan, M. A. G. Koffas, *ACS Synth Biol* **2012**, *1*, 256-266.
- [76] R. P. Shetty, D. Endy, T. F. Knight, *J Biol Eng* **2008**, *2*.
- [77] D. Rebatchouk, N. Daraselia, J. O. Narita, *PNAS* **1996**, *93*, 10891-10896.
- [78] T. S. Lee, R. A. Krupa, F. Zhang, M. Hajimorad, W. J. Holtz, N. Prasad, S. K. Lee, J. D. Keasling, *J Biol Eng* **2011**, *5*, 12.
- [79] J. C. Anderson, J. E. Dueber, M. Leguia, G. C. Wu, J. A. Goler, A. P. Arkin, J. D. Keasling, *J Biol Eng* **2010**, *4*, 1.
- [80] R. Carlson, *Nat Biotech* **2009**, *27*, 1091-1094.
- [81] S. Kosuri, G. M. Church, *Nat Meth* **2014**, *11*, 499-507.
- [82] H. König, D. Frank, R. Heil, C. Coenen, *Curr Genomics* **2013**, *14*, 11-24.
- [83] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat Methods* **2009**, *6*.
- [84] G. A. Benders, V. N. Noskov, E. A. Denisova, C. Lartigue, D. G. Gibson, N. Assad-Garcia, R.-Y. Chuang, W. Carrera, M. Moodie, M. A. Algire, Q. Phan, N. Alperovich, S. Vashee, C. Merryman, J. C. Venter, H. O. Smith, J. I. Glass, C. A. Hutchison, *Nucleic Acids Res* **2010**, *38*, 2558-2569.
- [85] J. Quan, J. Tian, *PLoS ONE* **2009**, *4*, e6441.
- [86] M. Z. Li, S. J. Elledge, *Nat Meth* **2007**, *4*, 251-256.
- [87] C. Li, A. Wen, B. Shen, J. Lu, Y. Huang, Y. Chang, *BMC Biotechnol* **2011**, *11*, 92-92.
- [88] S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, W. M. Rehrauer, *Microbiol Rev* **1994**, *58*, 401-465.
- [89] V. Larionov, N. Kouprina, J. Graves, X. N. Chen, J. R. Korenberg, M. A. Resnick, *PNAS* **1996**, *93*, 491-496.
- [90] D. G. Gibson, G. A. Benders, K. C. Axelrod, J. Zaveri, M. A. Algire, M. Moodie, M. G. Montague, J. C. Venter, H. O. Smith, C. A. Hutchison, *PNAS* **2008**, *105*, 20404-20409.
- [91] N. Traitcheva, H. Jenke-Kodama, J. He, E. Dittmann, C. Hertweck *ChemBioChem* **2007**, *8*, 1841-1849.
- [92] Y. S. Choi, T. W. Johannes, M. Simurdiak, Z. Shao, H. Lu, H. Zhao, *Mol BioSyst* **2010**, *6*, 336-338.
- [93] Z. Shao, Y. Luo, H. Zhao, *Mol BioSyst* **2011**, *7*, 1056-1059.
- [94] Z. Shao, H. Zhao, *Methods in enzymology, Vol. 517*, Academic Press, **2012**.
- [95] Z. Shao, H. Zhao, H. Zhao, *Nucleic Acids Res* **2009**, *37*.
- [96] D. G. Gibson, H. O. Smith, C. A. Hutchison, J. C. Venter, C. Merryman, *Nat Meth* **2010**, *7*, 901-903.
- [97] Y. Zhang, U. Werling, W. Edelmann, *Nucleic Acids Res* **2012**, *40*, e55.
- [98] K. S. Wilson, P. H. von Hippel, *PNAS* **1995**, *92*, 8793-8797.
- [99] P. J. Farnham, T. Platt, *Nucleic Acids Res* **1981**, *9*, 563-577.
- [100] S. D. Colloms, C. A. Merrick, F. J. Olorunniji, W. M. Stark, M. C. M. Smith, A. Osbourn, J. D. Keasling, S. J. Rosser, *Nucleic Acids Res* **2014**, *42*, e23.
- [101] N. G. A. Kuijpers, S. Chroumpi, T. Vos, D. Solis-Escalante, L. Bosman, J. T. Pronk, J.-M. Daran, P. Daran-Lapujade, *FEMS Yeast Res* **2013**, *13*, 769-781.
- [102] J. T. MacDonald, C. Barnes, R. I. Kitney, P. S. Freemont, G.-B. V. Stan, *Integr Biol* **2011**, *3*, 97-108.
- [103] K. E. J. Tyo, P. K. Ajikumar, G. Stephanopoulos, *Nat Biotech* **2009**, *27*, 760-765.
- [104] V. C. Dunkel, E. Zeiger, D. Brusick, E. McCoy, D. McGregor, K. Mortelmans, H. S. Rosenkranz, V. F. Simmon, *Environ Mutagen* **1985**, *7*, 1-19.
- [105] E. M. Witkin, *Bacteriol Rev* **1976**, *40*, 869-907.
- [106] R. Simon, U. Priefer, A. Puhler, *Nat Biotech* **1983**, *1*, 784-791.
- [107] E. Martínez-García, T. Aparicio, V. de Lorenzo, P. I. Nikel, *Front Bioeng Biotechnol* **2014**, *2*, 46.
- [108] S. C. Kowalczykowski, *Trends Biochem Sci*, *12*, 141-145.
- [109] D. Boyd, D. S. Weiss, J. C. Chen, J. Beckwith, *J Bacteriol* **2000**, *182*, 842-847.
- [110] Y.-Y. Cui, C. Ling, Y.-Y. Zhang, J. Huang, J.-Z. Liu, *Microb Cell Fact* **2014**, *13*, 21.
- [111] D. Koma, H. Yamanaka, K. Moriyoshi, T. Ohmoto, K. Sakai, *Appl Microbiol Biotechnol* **2012**, *93*, 815-829.
- [112] K. A. Datsenko, B. L. Wanner, *PNAS* **2000**, *97*, 6640-6645.
- [113] C. W. Song, S. Y. Lee, *Biotechnol J* **2013**, *8*, 776-784.
- [114] T. E. Kuhlman, E. C. Cox, *Nucleic Acids Res* **2010**, *38*, e92.
- [115] S. Sabri, J. A. Steen, M. Bongers, L. K. Nielsen, C. E. Vickers, *Microb Cell Fact* **2013**, *12*, 60.
- [116] C. D. Hook, V. V. Samsonov, A. A. Ublinskaya, T. M. Kuvaeva, E. V. Andreeva, L. Y. Gorbacheva, N. V. Stoyanova, *J Microbiol Methods* **2016**, *130*, 83-91.
- [117] M. Z. Li, S. J. Elledge, *Nat Genet* **2005**, *37*, 311-319.
- [118] A. Kuzminov, *Microbiol Mol Biol Rev* **1999**, *63*, 751-813.
- [119] Y. Ishino, H. Shinagawa, K. Makino, M. Amemura, A. Nakata, *J Bacteriol* **1987**, *169*, 5429-5433.
- [120] A. Bolotin, B. Quinquis, A. Sorokin, S. D. Ehrlich, *Microbiol* **2005**, *151*, 2551-2561.
- [121] C. Pourcel, G. Salvignol, G. Vergnaud, *Microbiol* **2005**, *151*, 653-663.
- [122] D. H. Haft, J. Selengut, E. F. Mongodin, K. E. Nelson, *PLoS Comput Biol* **2005**, *1*, e60.
- [123] R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, P. Horvath, *Science* **2007**, *315*, 1709-1712.
- [124] G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, *PNAS* **2012**, *109*, E2579-E2586.
- [125] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816-821.
- [126] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, *Nat Protoc* **2013**, *8*, 2281-2308.
- [127] M. L. Luo, R. T. Leenay, C. L. Beisel, *Biotechnol Bioeng* **2016**, *113*, 930-943.
- [128] I. Mougiakos, E. F. Bosma, W. M. de Vos, R. van Kranenburg, J. van der Oost, *Trends Biotechnol* **2016**, *34*, 575-587.
- [129] Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, *Appl Environ Microbiol* **2015**, *81*, 2506-2514.
- [130] M. E. Pyne, M. Moo-Young, D. A. Chung, C. P. Chou, *Appl Environ Microbiol* **2015**, *81*, 5103-5114.
- [131] M. C. Bassalo, A. D. Garst, A. L. Halweg-Edwards, W. C. Grau, D. W. Dommelle, V. K. Mutalik, A. P. Arkin, R. T. Gill, *ACS Synth Biol* **2016**, *5*, 561-568.
- [132] P. J. Enyeart, S. M. Chirieleison, M. N. Dao, J. Perutka, E. M. Quandt, J. Yao, J. T. Whitt, A. T. Keatinge-Clay, A. M. Lambowitz, A. D. Ellington, *Mol Syst Biol* **2013**, *9*, 685-685.
- [133] A. M. Lambowitz, S. Zimmerly, *Annu Rev Genet* **2004**, *38*, 1-35.
- [134] H. Guo, S. Zimmerly, P. S. Perlman, A. M. Lambowitz, *EMBO J* **1997**, *16*, 6835-6848.
- [135] H. Guo, M. Karberg, M. Long, J. P. Jones, B. Sullenger, A. M. Lambowitz, *Science* **2000**, *289*, 452-457.
- [136] K. Miura, Y. Tomioka, H. Suzuki, M. Yonezawa, T. Hishinuma, M. Mizugaki, *Biol Pharm Bull* **1997**, *20*.

- [137] R. E. Williams, N. C. Bruce, *Microbiol* **2002**, *148*, 1607-1614.
- [138] N. Oberleitner, *Ph.D. Thesis* **2016**, *TU Wien*, 1-148.
- [139] P. A. Hubbard, X. Liang, H. Schulz, J.-J. P. Kim, *J Biol Chem* **2003**, *278*, 37553-37560.
- [140] F. Baumgärtner, L. Seitz, G. A. Sprenger, C. Albermann, *Microb Cell Fact* **2013**, *12*, 40.
- [141] R. Agudo, M. T. Reetz, *Chem Commun* **2013**, *49*, 10914-10916.
- [142] L. O. Ingram, P. F. Gomez, X. Lai, M. Moniruzzaman, B. E. Wood, L. P. Yomano, S. W. York, *Biotechnol Bioeng* **1998**, *58*, 204-214.
- [143] S. Atsumi, J. C. Liao, *Curr Opin Biotechnol* **2008**, *19*, 414-419.
- [144] S. Atsumi, T. Hanai, J. C. Liao, *Nature* **2008**, *451*, 86-89.
- [145] W. Leuchtenberger, K. Huthmacher, K. Drauz, *Appl Microbiol Biotechnol* **2005**, *69*, 1-8.
- [146] R. M. Lennen, B. F. Pfleger, *Curr Opin Biotechnol* **2013**, *24*, 1044-1053.
- [147] E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. del Cardayre, J. D. Keasling, *Nature* **2010**, *463*, 559-562.
- [148] V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat Biotech* **2003**, *21*, 796-802.
- [149] R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach, M. Betlach, G. Ashley, *PNAS* **1999**, *96*, 1846-1851.
- [150] L. Z. Yuan, P. E. Rouvière, R. A. LaRossa, W. Suh, *Metab Eng* **2006**, *8*, 79-90.
- [151] G. DeSantis, K. Wong, B. Farwell, K. Chatman, Z. Zhu, G. Tomlinson, H. Huang, X. Tan, L. Bibbs, P. Chen, K. Kretz, M. J. Burk, *J Am Chem Soc* **2003**, *125*, 11476-11477.
- [152] E. Fernández-Álvarez, R. Snajdrova, H. Jochens, T. Davids, D. Böttcher, U. T. Bornscheuer, *Angew Chem Int Ed* **2011**, *50*, 8584-8587.
- [153] E. Whittle, J. Shanklin, *J Biol Chem* **2001**, *276*, 21500-21505.
- [154] S. Milker, *Ph.D. Thesis* **2017**, *TU Wien*, 1-236.
- [155] S. Milker, M. J. Fink, N. Oberleitner, A. K. Ressmann, U. T. Bornscheuer, M. D. Mihovilovic, F. Rudroff, *ChemCatChem* **2017**, n/a-n/a.
- [156] D. Koma, H. Yamanaka, K. Moriyoshi, T. Ohmoto, K. Sakai, *Appl Environ Microbiol* **2012**, *78*, 6203-6216.
- [157] D. Omotajo, T. Tate, H. Cho, M. Choudhary, *BMC Genomics* **2015**, *16*, 604.
- [158] H. M. Salis, E. A. Mirsky, C. A. Voigt, *Nat Biotech* **2009**, *27*, 946-950.
- [159] B. S. Laursen, H. P. Sørensen, K. K. Mortensen, H. U. Sperling-Petersen, *Microbiol Mol Biol Rev* **2005**, *69*, 101-123.
- [160] A. G. Vitreschak, D. A. Rodionov, A. A. Mironov, M. S. Gelfand, *Trends Genet*, *20*, 44-50.
- [161] B. J. Tucker, R. R. Breaker, *Curr Opin Struct Biol* **2005**, *15*, 342-348.
- [162] D. Na, S. M. Yoo, H. Chung, H. Park, J. H. Park, S. Y. Lee, *Nat Biotech* **2013**, *31*, 170-174.
- [163] S. M. Yoo, D. Na, S. Y. Lee, *Nat Protoc* **2013**, *8*, 1694-1707.
- [164] H. Chen, M. Bjerknes, R. Kumar, E. Jay, *Nucleic Acids Res* **1994**, *22*, 4953-4957.
- [165] A. D. Garst, A. L. Edwards, R. T. Batey, *Cold Spring Harb Perspect Biol* **2011**, *3*, a003533.
- [166] A. Srivatsan, J. D. Wang, *Curr Opin Microbiol* **2008**, *11*, 100-105.
- [167] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular biology of the cell*, Garland Science, New York, **2002**.
- [168] W. A. Haseltine, R. Block, *PNAS* **1973**, *70*, 1564-1568.
- [169] A. Battesti, E. Bouveret, *Mol Microbiol* **2006**, *62*, 1048-1063.
- [170] J. Gallant, L. Palmer, C. C. Pao, *Cell* **1977**, *11*, 181-185.
- [171] M. Jishage, K. Kvint, V. Shingler, T. Nyström, *Genes Dev* **2002**, *16*, 1260-1270.
- [172] M. F. Traxler, S. M. Summers, H.-T. Nguyen, V. M. Zacharia, G. A. Hightower, J. T. Smith, T. Conway, *Mol Microbiol* **2008**, *68*, 1128-1148.
- [173] G. Hannig, S. C. Makrides, *Trends Biotechnol*, *16*, 54-60.
- [174] F. Jacob, J. Monod, *J Mol Biol* **1961**, *3*, 318-356.
- [175] W. Gilbert, B. Müller-Hill, *PNAS* **1966**, *56*, 1891-1898.
- [176] B. Müller-Hill, *The Lac Operon: a short history of a genetic paradigm*, de Gruyter, Berlin, **1996**.
- [177] J. W. Dubendorff, F. W. Studier, *J Mol Biol* **1991**, *219*, 61-68.
- [178] G. J. Balzer, C. Thakker, G. N. Bennett, K.-Y. San, *Metab Eng* **2013**, *20*, 1-8.
- [179] M. Giacalone, A. Gentile, B. Lovitt, N. Berkley, C. Gunderson, M. Surber, *BioTechniques* **2006**, *40*, 355-364.
- [180] Y. J. Choi, L. Morel, T. Le François, D. Bourque, L. Bourget, D. Groleau, B. Massie, C. B. Míguez, *Appl Environ Microbiol* **2010**, *76*, 5058-5066.
- [181] T. Brautaset, R. Lale, S. Valla, *Microb Biotechnol* **2009**, *2*, 15-30.
- [182] X.-X. Wei, Z.-Y. Shi, M.-Q. Yuan, G.-Q. Chen, *Appl Microbiol Biotechnol* **2008**, *82*, 703.
- [183] S. Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, *Nat Biotech* **2002**, *20*, 1041-1044.
- [184] D. Camsund, P. Lindblad, A. Jaramillo, *Biotechnol J* **2011**, *6*, 826-836.
- [185] U. Krauss, T. Drepper, K.-E. Jaeger, *Chem Eur J* **2011**, *17*, 2552-2560.
- [186] R. H. Dahl, F. Zhang, J. Alonso-Gutierrez, E. Baidoo, T. S. Batth, A. M. Redding-Johanson, C. J. Petzold, A. Mukhopadhyay, T. S. Lee, P. D. Adams, J. D. Keasling, *Nat Biotech* **2013**, *31*, 1039-1046.
- [187] M. B. M. and, B. L. Bassler, *Annu Rev Microbiol* **2001**, *55*, 165-199.
- [188] B. M. M. Ahmer, *Mol Microbiol* **2004**, *52*, 933-945.
- [189] C.-Y. Tsao, L. Wang, Y. Hashimoto, H. Yi, J. C. March, M. P. DeLisa, T. K. Wood, J. J. Valdes, W. E. Bentley, *Appl Environ Microbiol* **2011**, *77*, 2141-2152.
- [190] C. M. Waters, B. L. Bassler, *Annu Rev Cell Dev Biol* **2005**, *21*, 319-346.
- [191] W.-L. Ng, B. L. Bassler, *Annu Rev Genet* **2009**, *43*, 197-222.
- [192] R. Popat, D. M. Cornforth, L. McNally, S. P. Brown, *J R Soc Interface* **2015**, *12*.
- [193] H. Alper, C. Fischer, E. Nevoigt, G. Stephanopoulos, *PNAS* **2005**, *102*, 12678-12683.
- [194] J. R. Anthony, L. C. Anthony, F. Nowroozi, G. Kwon, J. D. Newman, J. D. Keasling, *Metab Eng* **2009**, *11*.
- [195] M. De Mey, J. Maertens, S. Boogmans, W. K. Soetaert, E. J. Vandamme, R. Cunin, M. R. Foulquié-Moreno, *BMC Biotechnol* **2010**, *10*, 26-26.
- [196] J. Blazeck, R. Garg, B. Reed, H. S. Alper, *Biotechnol Bioeng* **2012**, *109*, 2884-2895.
- [197] K.-H. Siu, R. P. Chen, Q. Sun, L. Chen, S.-L. Tsai, W. Chen, *Curr Opin Biotechnol* **2015**, *36*, 98-106.
- [198] H. Aiba, A. Hanamura, H. Yamano, *J Biol Chem* **1991**, *266*, 1721-1727.
- [199] M. Kozak, *Gene* **1999**, *234*, 187-208.
- [200] M. Kozak, *Nucleic Acids Res* **2005**, *33*, 6593-6602.
- [201] M. H. de Smit, J. van Duin, *PNAS* **1990**, *87*, 7668-7672.
- [202] D. Na, D. Lee, *Bioinformatics* **2010**, *26*, 2633-2634.
- [203] D. Na, S. Lee, D. Lee, *BMC Syst Biol* **2010**, *4*, 71-71.
- [204] H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest, G. M. Church, *Nature* **2009**, *460*, 894-898.
- [205] P. J. Westfall, D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, C. J. Paddon, *PNAS* **2012**, *109*, E111-E118.
- [206] F. F. Nowroozi, E. E. K. Baidoo, S. Ermakov, A. M. Redding-Johanson, T. S. Batth, C. J. Petzold, J. D. Keasling, *Appl Microbiol Biotechnol* **2014**, *98*, 1567-1581.

- [207] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790-1792.
- [208] G. Cannarozzi, N. N. Schraudolph, M. Faty, P. von Rohr, M. T. Friberg, A. C. Roth, P. Gonnet, G. Gonnet, Y. Barral, *Cell*, *141*, 728.
- [209] K. Fredrick, M. Ibba, *Cell* **2010**, *141*, 227-229.
- [210] T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, Y. Pilpel, *Cell*, *141*, 344-354.
- [211] S. R. Eddy, *Nat Rev Genet* **2001**, *2*, 919-929.
- [212] a. E G H Wagner, R. W. Simons, *Annu Rev Microbiol* **1994**, *48*, 713-742.
- [213] P. Stougaard, S. Molin, K. Nordström, *PNAS* **1981**, *78*, 6008-6012.
- [214] A. S. Mironov, I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, E. Nudler, *Cell*, *111*, 747-756.
- [215] M. Morita, M. Kanemori, H. Yanagi, T. Yura, *J Bacteriol* **1999**, *181*, 401-410.
- [216] C.-H. Hoe, C. A. Raabe, T. S. Rozhdestvensky, T.-H. Tang, *Int J Med Microbiol* **2013**, *303*, 217-229.
- [217] N. Majdalani, D. Hernandez, S. Gottesman, *Mol Microbiol* **2002**, *46*, 813-826.
- [218] S. Gottesman, *Trends Genet* **2005**, *21*, 399-404.
- [219] T. Soper, P. Mandin, N. Majdalani, S. Gottesman, S. A. Woodson, *PNAS* **2010**, *107*, 9602-9607.
- [220] P. Yi, T. J. Soper, S. A. Woodson, *Methods Mol Biol* **2012**, *905*, 213-224.
- [221] F. J. Isaacs, D. J. Dwyer, C. Ding, D. D. Pervouchine, C. R. Cantor, J. J. Collins, *Nat Biotech* **2004**, *22*, 841-847.
- [222] S. A. Lynch, S. K. Desai, H. K. Sajja, J. P. Gallivan, *Chem Biol* **2007**, *14*, 173-184.
- [223] G. Rodrigo, T. E. Landrain, A. Jaramillo, *PNAS* **2012**, *109*, 15271-15276.
- [224] Z. Kang, X. Wang, Y. Li, Q. Wang, Q. Qi, *Biotechnol Lett* **2012**, *34*, 527-531.
- [225] Z. Kang, Q. Wang, H. Zhang, Q. Qi, *Appl Microbiol Biotechnol* **2008**, *79*, 203-208.
- [226] D. Juminaga, E. E. K. Baidoo, A. M. Redding-Johanson, T. S. Batth, H. Burd, A. Mukhopadhyay, C. J. Petzold, J. D. Keasling, *Appl Environ Microbiol* **2012**, *78*, 89-98.
- [227] C. D. Smolke, T. A. Carrier, J. D. Keasling, *Appl Environ Microbiol* **2000**, *66*, 5399-5405.
- [228] L. Berg, R. Lale, I. Bakke, N. Burroughs, S. Valla, *Microb Biotechnol* **2009**, *2*, 379-389.
- [229] B. F. Pfeleger, D. J. Pitera, C. D. Smolke, J. D. Keasling, *Nat Biotech* **2006**, *24*, 1027-1032.
- [230] J.-W. Song, J.-M. Woo, G. Y. Jung, U. T. Bornscheuer, J.-B. Park, *Sci Rep* **2016**, *6*, 29406.
- [231] C. Cho, S. Y. Choi, Z. W. Luo, S. Y. Lee, *Biotechnol Adv* **2015**, *33*, 1455-1466.
- [232] L. Lv, Y.-L. Ren, J.-C. Chen, Q. Wu, G.-Q. Chen, *Metab Eng* **2015**, *29*, 160-168.
- [233] I. Massaiu, L. Pasotti, M. Casanova, N. Politi, S. Zucca, M. G. Cusella De Angelis, P. Magni, *Syst Synth Biol* **2015**, *9*, 107-123.
- [234] M. T. Reetz, *J Am Chem Soc* **2013**, *135*, 12480-12496.
- [235] T.-W. Wang, H. Zhu, X.-Y. Ma, T. Zhang, Y.-S. Ma, D.-Z. Wei, *Mol Biotechnol* **2006**, *34*, 55-68.
- [236] C. Jäckel, P. Kast, D. Hilvert, *Annu Rev Biophys* **2008**, *37*, 153-173.
- [237] C. A. Tracewell, F. H. Arnold, *Curr Opin Chem Biol* **2009**, *13*, 3-9.
- [238] G. L. Butterfoss, B. Kuhlman, *Annu Rev Biophys Biomol Struct* **2006**, *35*, 49-65.
- [239] A. M. Poole, R. Ranganathan, *Curr Opin Struct Biol* **2006**, *16*, 508-513.
- [240] S.-g. Kang, J. G. Saven, *Curr Opin Chem Biol* **2007**, *11*, 329-334.
- [241] A. Davidson, *Protein design*, Vol. 340, Springer, New York, **2006**.
- [242] K. B. Otte, B. Hauer, *Curr Opin Biotechnol* **2015**, *35*, 16-22.
- [243] R. J. Fox, S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J. C. Whitman, R. A. Sheldon, G. W. Huisman, *Nat Biotech* **2007**, *25*, 338-344.
- [244] W. P. C. Stemmer, *Nature* **1994**, *370*, 389-391.
- [245] J. E. Ness, M. Welch, L. Giver, M. Bueno, J. R. Cherry, T. V. Borchert, W. P. C. Stemmer, J. Minshull, *Nat Biotech* **1999**, *17*, 893-896.
- [246] E. O'Reilly, N. J. Turner, *Perspect Science* **2015**, *4*, 55-61.
- [247] M. Höhne, U. T. Bornscheuer, *ChemCatChem* **2009**, *1*, 42-51.
- [248] B. Gallwitz, *Vas Health Risk Manag* **2007**, *3*, 203-210.
- [249] K. B. Hansen, J. Balsells, S. Dreher, Y. Hsiao, M. Kubryk, M. Palucki, N. Rivera, D. Steinhuebel, J. D. Armstrong, D. Askin, E. J. J. Grabowski, *Org Process Res Dev* **2005**, *9*, 634-639.
- [250] K. B. Hansen, Y. Hsiao, F. Xu, N. Rivera, A. Clausen, M. Kubryk, S. Krska, T. Rosner, B. Simmons, J. Balsells, N. Ikemoto, Y. Sun, F. Spindler, C. Malan, E. J. J. Grabowski, J. D. Armstrong, *J Am Chem Soc* **2009**, *131*, 8798-8804.
- [251] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305-309.
- [252] Z.-G. Zhang, R. Lonsdale, J. Sanchis, M. T. Reetz, *J Am Chem Soc* **2014**, *136*, 17262-17272.
- [253] N. G. Schmidt, E. Eger, W. Kroutil, *ACS Catal* **2016**, *6*, 4286-4311.
- [254] A. K. Samland, M. Rale, G. A. Sprenger, W.-D. Fessner, *ChemBioChem* **2011**, *12*, 1454-1474.
- [255] X. Garrabou, J. A. Castillo, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, P. Clapés, *Angew Chem Int Ed* **2009**, *48*, 5521-5525.
- [256] P. Clapés, X. Garrabou, *Adv Synth Catal* **2011**, *353*, 2263-2283.
- [257] D. Güclü, A. Szekrenyi, X. Garrabou, M. Kickstein, S. Junker, P. Clapés, W.-D. Fessner, *ACS Catal* **2016**, *6*, 1848-1852.
- [258] M. Gutierrez, T. Parella, J. Joglar, J. Bujons, P. Clapes, *Chem Commun* **2011**, *47*, 5762-5764.
- [259] A. Szekrenyi, A. Soler, X. Garrabou, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, J. Bujons, P. Clapés, *Chem Eur J* **2014**, *20*, 12572-12583.
- [260] A. Szekrenyi, X. Garrabou, T. Parella, J. Joglar, J. Bujons, P. Clapés, *Nat Chem* **2015**, *7*, 724-729.
- [261] P. S. Coelho, E. M. Brustad, A. Kannan, F. H. Arnold, *Science* **2013**, *339*, 307-310.
- [262] H. Renata, Z. J. Wang, R. Z. Kitto, F. H. Arnold, *Catal Sci Technol* **2014**, *4*, 3640-3643.
- [263] J. Alliot, E. Gravel, F. Pillon, D.-A. Buisson, M. Nicolas, E. Doris, *Chem Commun* **2012**, *48*, 8111-8113.
- [264] Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho, F. H. Arnold, *Angew Chem Int Ed* **2014**, *53*, 6810-6813.
- [265] T. Heel, J. A. McIntosh, S. C. Dodani, J. T. Meyerowitz, F. H. Arnold, *ChemBioChem* **2014**, *15*, 2556-2562.
- [266] C. K. Savile, J. J. Lalonde, *Curr Opin Biotechnol* **2011**, *22*, 818-823.
- [267] J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St.Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael, D. Baker, *Science* **2010**, *329*, 309-313.
- [268] H. Jochens, U. T. Bornscheuer, *ChemBioChem* **2010**, *11*, 1861-1866.
- [269] K. Bernath, M. Hai, E. Mastrobattista, A. D. Griffiths, S. Magdassi, D. S. Tawfik, *Anal Biochem* **2004**, *325*, 151-157.
- [270] S. Becker, H. Höbenreich, A. Vogel, J. Knorr, S. Wilhelm, F. Rosenau, K.-E. Jaeger, M. T. Reetz, H. Kolmar, *Angew Chem Int Ed* **2008**, *47*, 5085-5088.
- [271] M. Wójcik, A. Telzerow, W. J. Quax, Y. L. Boersma, *Int J Mol Sci* **2015**, *16*, 24918-24945.

- [272] B. Chen, S. Lim, A. Kannan, S. C. Alford, F. Sunden, D. Herschlag, I. K. Dimov, T. M. Baer, J. R. Cochran, *Nat Chem Biol* **2016**, *12*, 76-81.
- [273] M. S. Packer, D. R. Liu, *Nat Rev Genet* **2015**, *16*, 379-394.
- [274] A. M. Weeks, M. C. Y. Chang, *Biochem* **2011**, *50*, 5404-5418.
- [275] S. Osswald, W. Wienand, U. Becker, K. Doderer, S. Verseck, *Pharm Technol* **2008**, *4*, 1-6.
- [276] R. Kratzer, J. M. Woodley, B. Nidetzky, *Biotechnol Adv* **2015**, *33*, 1641-1652.
- [277] S. P. France, L. J. Hepworth, N. J. Turner, S. L. Flitsch, *ACS Catal* **2017**, *7*, 710-724.
- [278] R. R. Chen, *Appl Microbiol Biotechnol* **2007**, *74*, 730-738.
- [279] T. Schweder, H. Lin, B. Jürgen, A. Breitenstein, S. Riemschneider, V. Khalameyzer, A. Gupta, K. Büttner, P. Neubauer, *Appl Microbiol Biotechnol* **2002**, *58*, 330-337.
- [280] A. M. Kunjapur, K. L. J. Prather, *Appl Environ Microbiol* **2015**, *81*, 1892-1901.
- [281] S. Singh, C. Brocker, V. Koppaka, C. Ying, B. Jackson, A. Matsumoto, D. C. Thompson, V. Vasiliou, *Free Radic Biol Med* **2013**, *56*, 89-101.
- [282] M. C. Y. Chang, R. A. Eachus, W. Trieu, D.-K. Ro, J. D. Keasling, *Nat Chem Biol* **2007**, *3*, 274-277.
- [283] J. E. Dueber, G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. J. Prather, J. D. Keasling, *Nat Biotech* **2009**, *27*, 753-759.
- [284] Y.-C. He, F. Liu, D.-P. Zhang, S. Gao, Z.-Q. Li, Z.-C. Tao, C.-L. Ma, *Appl Biochem Biotechnol* **2014**, *3*, 1-12.
- [285] D. Gobinath, S. G. Prapulla, *Biotechnol Lett* **2014**, *36*, 153-157.
- [286] H. J. Park, J. Jung, H. Choi, K. N. Uhm, H. K. Kim, *J Microbiol Biotechnol* **2010**, *20*, 1300-1306.
- [287] G. Thedei, D. P. S. Leitão, M. Bolean, T. P. Paulino, A. C. C. Spadaro, P. Ciancaglini, *Braz J Med Biol Res* **2008**, *41*, 1047-1053.
- [288] C. Grant, D. Deszcz, Y.-C. Wei, R. J. Martínez-Torres, P. Morris, T. Folliard, R. Sreenivasan, J. Ward, P. Dalby, J. M. Woodley, F. Baganz, *Sci Rep* **2014**, *4*, 5844.
- [289] M. K. Julsing, M. Schrewe, S. Cornelissen, I. Hermann, A. Schmid, B. Bühler, *Appl Environ Microbiol* **2012**, *78*, 5724-5733.
- [290] J. L. Coll, M. Heyde, R. Portalier, *Mol Microbiol* **1994**, *12*, 83-93.
- [291] H. Nikaido, *Microbiol Mol Biol Rev* **2003**, *67*, 593-656.
- [292] S. Kojima, H. Nikaido, *J Biol Chem* **2014**, *289*, 26464-26473.
- [293] M. Wei, Z. Li, T. Li, B. Wu, Y. Liu, J. Qu, X. Li, L. Li, L. Cai, P. G. Wang, *ACS Catal* **2015**, *5*, 4060-4065.
- [294] K. M. Dombek, L. O. Ingram, *Appl Environ Microbiol* **1987**, *53*, 1286-1291.
- [295] M. Sauer, D. Porro, D. Mattanovich, P. Branduardi, *Trends Biotechnol* **2008**, *26*, 100-108.
- [296] X. Chen, M. Li, L. Zhou, W. Shen, G. Algasan, Y. Fan, Z. Wang, *Bioresour Technol* **2014**, *166*, 64-71.
- [297] L. C. Thomason, N. Costantino, D. L. Court, in *Curr Protoc Mol Biol*, John Wiley & Sons, Inc., **2001**.
- [298] D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z.-Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, H. O. Smith, J. C. Venter, *Science* **2010**, *329*, 52-56.
- [299] C. A. Hutchison, R.-Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z.-Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson, J. C. Venter, *Science* **2016**, *351*.
- [300] B. Kim, W. J. Kim, D. I. Kim, S. Y. Lee, *J Ind Microbiol Biotechnol* **2015**, *42*, 339-348.
- [301] J. H. Ahn, Y.-S. Jang, S. Y. Lee, *Curr Opin Biotechnol* **2016**, *42*, 54-66.
- [302] S. Y. Lee, D.-Y. Lee, T. Y. Kim, *Trends Biotechnol*, *23*, 349-358.
- [303] J. D. Orth, I. Thiele, B. O. Palsson, *Nat Biotech* **2010**, *28*, 245-248.
- [304] X. Chen, A. P. Alonso, D. K. Allen, J. L. Reed, Y. Shachar-Hill, *Metab Eng* **2011**, *13*, 38-48.
- [305] J. Schellenberger, J. O. Park, T. M. Conrad, B. Ø. Palsson, *BMC Bioinf* **2010**, *11*, 213-213.
- [306] C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay, R. L. Stevens, *Nat Biotech* **2010**, *28*, 977-982.
- [307] J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke, B. O. Palsson, *Nat Protoc* **2011**, *6*, 1290-1307.
- [308] M. Lakshmanan, T. Y. Kim, B. K. S. Chung, S. Y. Lee, D.-Y. Lee, *BMC Syst Biol* **2015**, *9*, 73.
- [309] S. G. Oliver, M. K. Winson, D. B. Kell, F. Baganz, *Trends Biotechnol* **1998**, *16*, 373-378.
- [310] K. Tokuyama, S. Ohno, K. Yoshikawa, T. Hirasawa, S. Tanaka, C. Furusawa, H. Shimizu, *Microb Cell Fact* **2014**, *13*, 64.
- [311] R. R. Bommareddy, Z. Chen, S. Rappert, A.-P. Zeng, *Metab Eng* **2014**, *25*, 30-37.
- [312] K. Schroer, B. Zelic, M. Oldiges, S. Lütz, *Biotechnol Bioeng* **2009**, *104*, 251-260.
- [313] I. M. B. Reizman, A. R. Stenger, C. R. Reisch, A. Gupta, N. C. Connors, K. L. J. Prather, *Metab Eng Commun* **2015**, *2*, 109-116.
- [314] J. Koussa, A. Chaiboonchoe, K. Salehi-Ashtiani, *BioMed Res Int* **2014**, *2014*, 649453.
- [315] K. Faber, *Bioreformations in organic chemistry*, 5th ed., Springer, Berlin Heidelberg, **2004**.
- [316] R. D. Woodyer, T. W. Johannes, H. Zhao, *Enzyme technology*, 2nd ed., Springer, New York, **2008**.
- [317] H. K. Chenault, E. S. Simon, G. M. Whitesides, *Biotechnol Genet Eng Rev* **1988**, *6*, 221-270.
- [318] Z. Xiao, C. Lv, C. Gao, J. Qin, C. Ma, Z. Liu, P. Liu, L. Li, P. Xu, *PLoS ONE* **2010**, *5*, e8860.
- [319] E. García-Junceda, I. Lavandera, D. Rother, J. H. Schrittwieser, *Journal of Molecular Catalysis B: Enzymatic* **2015**, *114*, 1-6.
- [320] W. Hummel, H. Gröger, *Journal of Biotechnology* **2014**, *191*, 22-31.
- [321] A. Bornadel, R. Hatti-Kaul, F. Hollmann, S. Kara, *ChemCatChem* **2015**, *7*, 2442-2445.
- [322] N. Oberleitner, C. Peters, F. Rudroff, U. T. Bornscheuer, M. D. Mihovilovic, *Biotechnol J* **2014**, *192*, 393-399.
- [323] A. Bornadel, R. Hatti-Kaul, F. Hollmann, S. Kara, *Tetrahedron* **2016**, *72*, 7222-7228.
- [324] V. I. Tishkov, A. G. Galkin, V. V. Fedorchuk, P. A. Savitsky, A. M. Rojkova, H. Gieren, M.-R. Kula, *Biotechnol Bioeng* **1999**, *64*, 187-193.
- [325] A. M. G. Costas, A. K. White, W. W. Metcalf, *J Biol Chem* **2001**, *276*, 17429-17436.
- [326] J. M. Vrtis, A. K. White, W. W. Metcalf, W. A. van der Donk, *J Am Chem Soc* **2001**, *123*, 2672-2673.
- [327] R. Woodyer, W. A. van der Donk, H. Zhao, *Biochem* **2003**, *42*, 11604-11614.
- [328] D. E. Torres Pazmiño, A. Riebel, J. de Lange, F. Rudroff, M. D. Mihovilovic, M. W. Fraaije, *ChemBioChem* **2009**, *10*, 2595-2598.
- [329] A. Beier, S. Bordewick, M. Genz, S. Schmidt, T. van den Bergh, C. Peters, H.-J. Joosten, U. T. Bornscheuer, *ChemBioChem* **2016**, *17*, 2312-2315.
- [330] Y. Wang, K.-Y. San, G. N. Bennett, *Curr Opin Biotechnol* **2013**, *24*, 994-999.
- [331] S. Klatte, V. F. Wendisch, *Bioorg Med Chem* **2014**, *22*, 5578-5585.
- [332] E. Vázquez-Figueroa, J. Chaparro-Riggers, A. S. Bommarius, *ChemBioChem* **2007**, *8*, 2295-2301.
- [333] S. J. Berríos-Rivera, G. N. Bennett, K.-Y. San, *Metab Eng* **2002**, *4*, 217-229.

- [334] Y. Wang, L. Li, C. Ma, C. Gao, F. Tao, P. Xu, *Sci Rep* **2013**, *3*, 2643.
- [335] Y. J. Zhou, W. Yang, L. Wang, Z. Zhu, S. Zhang, Z. K. Zhao, *Microb Cell Fact* **2013**, *12*, 103.
- [336] R. Liu, L. Liang, M. Wu, K. Chen, M. Jiang, J. Ma, P. Wei, P. Ouyang, *Biochem Eng J* **2013**, *79*, 77-83.
- [337] A. Singh, M. D. Lynch, R. T. Gill, *Metab Eng* **2009**, *11*, 347-354.
- [338] S. K. Spaans, R. A. Weusthuis, J. van der Oost, S. W. M. Kengen, *Front Microbiol* **2015**, *6*, 742.
- [339] J. W. Chin, P. C. Cirino, *Biotechnol Prog* **2011**, *27*, 333-341.
- [340] S. Siedler, S. Bringer, L. M. Blank, M. Bott, *Appl Microbiol Biotechnol* **2012**, *93*, 1459-1467.
- [341] A. Shi, X. Zhu, J. Lu, X. Zhang, Y. Ma, *Metab Eng* **2013**, *16*, 1-10.
- [342] A. Weckbecker, H. Gröger, W. Hummel, in *Biosystems engineering I: Creating superior biocatalysts* (Eds.: C. Wittmann, R. Krull), Springer Berlin Heidelberg, Berlin, Heidelberg, **2010**, pp. 195-242.
- [343] R. Wichmann, D. Vasic-Racki, in *Technology transfer in biotechnology: From lab to industry to production* (Ed.: U. Kragl), Springer Berlin Heidelberg, Berlin, Heidelberg, **2005**, pp. 225-260.
- [344] D. Holtmann, S. Kochius, in *Encyclopedia of applied electrochemistry* (Eds.: G. Kreysa, K.-i. Ota, R. F. Savinell), Springer New York, New York, NY, **2014**, pp. 214-221.
- [345] J. N. Andexer, M. Richter, *ChemBioChem* **2015**, *16*, 380-386.
- [346] S. Mordhorst, J. Siegrist, M. Müller, M. Richter, J. N. Andexer, *Angew Chem Int Ed* **2017**, *56*, 4037-4041.
- [347] Y. Satoh, K. Tajima, H. Tannai, M. Munekata, *J Biosci Bioeng* **2003**, *95*, 335-341.
- [348] T. Bayer, S. Milker, T. Wiesinger, M. Winkler, M. D. Mihovilovic, F. Rudroff, *ChemCatChem* **2017**, *9*, 2919-2923.
- [349] I. Wheeldon, S. D. Minter, S. Banta, S. C. Barton, P. Atanassov, M. Sigman, *Nat Chem* **2016**, *8*, 299-309.
- [350] H. O. Spivey, J. Ovádi, *Methods* **1999**, *19*, 306-321.
- [351] E. W. Miles, S. Rhee, D. R. Davies, *J Biol Chem* **1999**, *274*, 12193-12196.
- [352] H. Lee, W. C. DeLoache, J. E. Dueber, *Metab Eng* **2012**, *14*, 242-251.
- [353] C. M. Agapakis, P. M. Boyle, P. A. Silver, *Nat Chem Biol* **2012**, *8*, 527-535.
- [354] C. C. Hyde, S. A. Ahmed, E. A. Padlan, E. W. Miles, D. R. Davies, *J Biol Chem* **1988**, *263*, 17857-17871.
- [355] K. S. Anderson, E. W. Miles, K. A. Johnson, *J Biol Chem* **1991**, *266*, 8020-8033.
- [356] F. Wu, S. Minter, *Angew Chem Int Ed* **2015**, *54*, 1851-1854.
- [357] J. Wang, N. S. Nemeria, K. Chandrasekhar, S. Kumaran, P. Arjunan, S. Reynolds, G. Calero, R. Brukh, L. Kakalis, W. Furey, F. Jordan, *J Biol Chem* **2014**, *289*, 15215-15230.
- [358] R. N. Perham, *Annu Rev Biochem* **2000**, *69*, 961-1004.
- [359] M. Leibundgut, T. Maier, S. Jenni, N. Ban, *Curr Opin Struct Biol* **2008**, *18*, 714-725.
- [360] M. Ishikawa, D. Tsuchiya, T. Oyama, Y. Tsunaka, K. Morikawa, *EMBO J* **2004**, *23*, 2745-2754.
- [361] Y. H. P. Zhang, *Biotechnol Adv* **2011**, *29*, 715-725.
- [362] K. Yu, C. Liu, B.-G. Kim, D.-Y. Lee, *Biotechnol Adv* **2015**, *33*, 155-164.
- [363] L. Iturrate, I. Sánchez-Moreno, I. Oroz-Guinea, J. Pérez-Gil, E. García-Junceda, *Chem Eur J* **2010**, *16*, 4018-4030.
- [364] L. Kizer, D. J. Pitera, B. F. Pfleger, J. D. Keasling, *Appl Environ Microbiol* **2008**, *74*, 3229-3241.
- [365] R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnšek, N. Tomšič, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Benčina, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala, M. P. DeLisa, *Nucleic Acids Res* **2012**, *40*, 1879-1889.
- [366] C. J. Delebecque, A. B. Lindner, P. A. Silver, F. A. Aldaye, *Science* **2011**, *333*, 470-474.
- [367] J. Fu, M. Liu, Y. Liu, N. W. Woodbury, H. Yan, *J Am Chem Soc* **2012**, *134*, 5516-5519.
- [368] O. I. Wilner, S. Shimron, Y. Weizmann, Z.-G. Wang, I. Willner, *Nano Lett* **2009**, *9*, 2040-2043.
- [369] J. Müller, C. M. Niemeyer, *Biochem Biophys Res Commun* **2008**, *377*, 62-67.
- [370] D. M. Vriezema, P. M. L. Garcia, N. Sancho Oltra, N. S. Hatzakis, S. M. Kuiper, R. J. M. Nolte, A. E. Rowan, J. C. M. van Hest, *Angew Chem Int Ed* **2007**, *46*, 7378-7382.
- [371] X. Wang, Z. Li, J. Shi, H. Wu, Z. Jiang, W. Zhang, X. Song, Q. Ai, *ACS Catal* **2014**, *4*, 962-972.
- [372] C. Schmidt-Dannert, F. Lopez-Gallego, *Microbial Biotechnol* **2016**, *9*, 601-609.
- [373] A. H. Chen, P. A. Silver, *Trends Cell Biol* **2012**, *22*, 662-670.
- [374] C. M. Metallo, M. G. V. Heiden, *Mol Cell* **2013**, *49*, 388-398.
- [375] D. P. Patterson, B. Schwarz, R. S. Waters, T. Gedeon, T. Douglas, *ACS Chem Biol* **2014**, *9*, 359-365.
- [376] J. E. Glasgow, M. A. Asensio, C. M. Jakobson, M. B. Francis, D. Tullman-Ercek, *ACS Synth Biol* **2015**, *4*, 1011-1019.
- [377] M. Held, A. Kolb, S. Perdue, S.-Y. Hsu, S. E. Bloch, M. B. Quin, C. Schmidt-Dannert, *Sci Rep* **2016**, *6*, 24359.
- [378] C. Aussignargues, B. C. Paasch, R. Gonzalez-Esquer, O. Erbilgin, C. A. Kerfeld, *Commun Integr Biol* **2015**, *8*, e1039755.
- [379] A. D. Lawrence, S. Frank, S. Newnham, M. J. Lee, I. R. Brown, W.-F. Xue, M. L. Rowe, D. P. Mulvihill, M. B. Prentice, M. J. Howard, M. J. Warren, *ACS Synth Biol* **2014**, *3*, 454-465.
- [380] M. Held, M. B. Quin, C. Schmidt-Dannert, *J Mol Microbiol Biotechnol* **2013**, *23*, 308-320.
- [381] T. A. Bobik, B. P. Lehman, T. O. Yeates, *Mol Microbiol* **2015**, *98*, 193-207.
- [382] C. A. Kerfeld, M. R. Melnicki, *Curr Opin Plant Biol* **2016**, *31*, 66-75.
- [383] F. Caschera, V. Noireaux, *Curr Opin Chem Biol* **2014**, *22*, 85-91.
- [384] Y. Elani, R. V. Law, O. Ces, *Nat Commun* **2014**, *5*, 5305.
- [385] A. Kuchler, M. Yoshimoto, S. Luginbuhl, F. Mavelli, P. Walde, *Nat Nano* **2016**, *11*, 409-420.
- [386] J. L. Snoep, H. V. Westerhoff, in *Systems Biology: Definitions and Perspectives* (Eds.: L. Alberghina, H. V. Westerhoff), Springer Berlin Heidelberg, Berlin, Heidelberg, **2005**, pp. 13-30.
- [387] B. N. Kholodenko, F. J. Bruggeman, H. M. Sauro, in *Systems Biology: Definitions and Perspectives* (Eds.: L. Alberghina, H. V. Westerhoff), Springer Berlin Heidelberg, Berlin, Heidelberg, **2005**, pp. 143-159.
- [388] D. di Bernardo, M. J. Thompson, T. S. Gardner, S. E. Chobot, E. L. Eastwood, A. P. Wojtovich, S. J. Elliott, S. E. Schaus, J. J. Collins, *Nat Biotech* **2005**, *23*, 377-383.
- [389] M. Schrewe, N. Ladkau, B. Bühler, A. Schmid, *Adv Synth Catal* **2013**, *355*, 1693-1697.
- [390] G. M. Rodriguez, S. Atsumi, *Microb Cell Fact* **2012**, *11*, 1-11.
- [391] M. Schrewe, M. K. Julsing, K. Lange, E. Czarnotta, A. Schmid, B. Bühler, *Biotechnol Bioeng* **2014**, *111*, 1820-1830.
- [392] A. S. Bommarius, B. R. Riebel-Bommarius, *Biocatalysis: fundamentals and applications*, Wiley-VCH, New York, **2004**.
- [393] B. M. Nestl, B. Hauer, *Nat Chem Biol* **2013**, *9*, 470-471.
- [394] F. H. Arnold, *Nature* **2001**, *409*, 253-257.
- [395] S. Y. Lee, H. U. Kim, *Nat Biotech* **2015**, *33*, 1061-1072.
- [396] K.-G. Fahlbusch, F.-J. Hammerschmidt, J. Panten, W. Pickenhagen, D. Schatkowski, K. Bauer, D. Garbe, H. Surburg, in *Ullmann's encyclopedia of industrial chemistry*, Wiley-VCH, New York, **2000**.

- [397] H. Yun, B.-G. Kim, *Biotechnol Bioprocess Eng* **2008**, *13*, 372.
- [398] D. Meyer, L. Walter, G. Kolter, M. Pohl, M. Müller, K. Tittmann, *J Am Chem Soc* **2011**, *133*, 3609-3616.
- [399] E. C. Hayden, *Nature* **2014**, *505*, 598.
- [400] E. Cabisco, J. Tamarit, J. Ros, *Int Microbial* **2000**, *3*, 3-8.
- [401] N. Kato, H. Konishi, K. Uda, M. Shimao, C. Sakazawa, *Agric Biol Chem* **1988**, *52*, 1885-1886.
- [402] J. Casey, R. Dobb, *Enzyme Microb Technol* **1992**, *14*, 739-747.
- [403] T. Li, J. P. Rosazza, *J Bacteriol* **1997**, *179*, 3482-3487.
- [404] A. He, T. Li, L. Daniels, I. Fotheringham, J. P. N. Rosazza, *Appl Environ Microbiol* **2004**, *70*, 1874-1881.
- [405] P. Venkatasubramanian, L. Daniels, J. P. N. Rosazza, *J Biol Chem* **2007**, *282*, 478-485.
- [406] G. M. Rodriguez, S. Atsumi, *Metab Eng* **2014**, *25*, 227-237.
- [407] D. Romano, R. Villa, F. Molinari, *ChemCatChem* **2012**, *4*, 739-749.
- [408] R. Gandolfi, R. Villa, F. Molinari, *Tetrahedron Lett* **2001**, *42*, 513-514.
- [409] V. C. Corberán, M. E. González-Pérez, S. Martínez-González, A. Gómez-Avilés, *Appl Catal A* **2014**, *474*, 211-223.
- [410] S. Shi, M. Liu, L. Zhao, M. Wang, C. Chen, J. Gao, J. Xu, *Chem Asian J* **2017**, *1*, 1-8.
- [411] S. J. B. Duff, W. D. Murray, *Biotechnol Bioeng* **1989**, *34*, 153-159.
- [412] P. Zambelli, A. Pinto, D. Romano, E. Crotti, P. Conti, L. Tamborini, R. Villa, F. Molinari, *Green Chem* **2012**, *14*, 2158-2161.
- [413] J. B. Jones, I. J. Jakovac, *Can. J. Chem.* **1982**, *60*, 19-28.
- [414] T. Orbeogo, I. Lavandera, W. M. F. Fabian, B. Mautner, J. G. de Vries, W. Kroutil, *Tetrahedron* **2009**, *65*, 6805-6809.
- [415] W.-D. Fessner, V. Helaine, *Curr Opin Biotechnol* **2001**, *12*, 574-586.
- [416] C. L. Windle, M. Müller, A. Nelson, A. Berry, *Curr Opin Chem Biol* **2014**, *19*, 25-33.
- [417] K. Napora-Wijata, K. Robins, A. Osorio-Lozada, M. Winkler, *ChemCatChem* **2014**, *6*, 1089-1095.
- [418] M. Fuchs, K. Tauber, J. Sattler, H. Lechner, J. Pfeffer, W. Kroutil, K. Faber, *RSC Adv* **2012**, *2*, 6262-6265.
- [419] E.-S. Park, M. Kim, J.-S. Shin, *Appl Microbiol Biotechnol* **2012**, *93*, 2425-2435.
- [420] M. J. H. Moonen, A. H. Westphal, I. M. C. M. Rietjens, W. J. H. van Berkel, *Adv Synth Catal* **2005**, *347*, 1027-1034.
- [421] K. E. Scholz, D. Okrob, B. Kopka, A. Grünberger, M. Pohl, K.-E. Jaeger, U. Krauss, *Appl Environ Microbiol* **2012**, *78*, 5025-5027.
- [422] C. M. Tripathi, S. C. Agarwal, S. K. Basu, *J Ferment Bioeng* **1997**, *84*, 487-492.
- [423] V. B. Shukla, P. R. Kulkarni, *World J Microbiol Biotechnol* **2000**, *16*, 499-506.
- [424] M. Pohl, B. Lingen, M. Müller, *Chem Eur J* **2002**, *8*, 5288-5295.
- [425] G. Brahmachari, *RSC Adv* **2016**, *6*, 64676-64725.
- [426] T. D. Machajewski, C.-H. Wong, *Angew Chem Int Ed* **2000**, *39*, 1352-1375.
- [427] M. M. Heravi, S. Asadi, *Tetrahedron: Asymmetry* **2012**, *23*, 1431-1465.
- [428] M. Breuer, B. Hauer, *Curr Opin Biotechnol* **2003**, *14*, 570-576.
- [429] S. M. Zgiby, G. J. Thomson, S. Qamar, A. Berry, *Eur J Biochem* **2000**, *267*, 1858-1868.
- [430] M. D. Bednarski, E. S. Simon, N. Bischofberger, W. D. Fessner, M. J. Kim, W. Lees, T. Saito, H. Waldmann, G. M. Whitesides, *J Am Chem Soc* **1989**, *111*, 627-635.
- [431] M. Schürmann, G. A. Sprenger, *J Biol Chem* **2001**, *276*, 11055-11061.
- [432] J. A. Castillo, C. Guérard-Hélaine, M. Gutiérrez, X. Garrabou, M. Sancelme, M. Schürmann, T. Inoue, V. Hélaine, F. Charmantray, T. Gefflaut, L. Hecquet, J. Joglar, P. Clapés, G. A. Sprenger, M. Lemaire, *Adv Synth Catal* **2010**, *352*, 1039-1046.
- [433] R. Roldán, I. Sanchez-Moreno, T. Scheidt, V. Hélaine, M. Lemaire, T. Parella, P. Clapés, W.-D. Fessner, C. Guérard-Hélaine, *Chem Eur J* **2017**, *23*, 5005-5009.
- [434] A. Soler, M. L. Gutiérrez, J. Bujons, T. Parella, C. Minguillon, J. Joglar, P. Clapés, *Adv Synth Catal* **2015**, *357*, 1787-1807.
- [435] T. Wiesinger, T. Bayer, S. Milker, M. D. Mihovilovic, F. Rudroff, *ChemBioChem* **2017**, n/a.
- [436] M. Sugiyama, Z. Hong, P.-H. Liang, S. M. Dean, L. J. Whalen, W. A. Greenberg, C.-H. Wong, *J Am Chem Soc* **2007**, *129*, 14811-14817.
- [437] J. Hao, A. Berry, *Prot Eng Des Sel* **2004**, *17*, 689-697.
- [438] M. Sugiyama, Z. Hong, W. A. Greenberg, C.-H. Wong, *Bioorg Med Chem* **2007**, *15*, 5905-5911.
- [439] X. Garrabou, J. Joglar, T. Parella, R. Crehuet, J. Bujons, P. Clapés, *Adv Synth Catal* **2011**, *353*, 89-99.
- [440] A. Bolt, A. Berry, A. Nelson, *Arch Biochem Biophys* **2008**, *474*, 318-330.
- [441] G. J. Williams, S. Domann, A. Nelson, A. Berry, *PNAS* **2003**, *100*, 3143-3148.
- [442] E. A. Althoff, L. Wang, L. Jiang, L. Giger, J. K. Lassila, Z. Wang, M. Smith, S. Hari, P. Kast, D. Herschlag, D. Hilvert, D. Baker, *Protein Sci* **2012**, *21*, 717-726.
- [443] R. Obexer, A. Godina, X. Garrabou, P. R. E. Mittl, D. Baker, A. D. Griffiths, D. Hilvert, *Nat Chem* **2017**, *9*, 50-56.
- [444] A. L. Concia, C. Lozano, J. A. Castillo, T. Parella, J. Joglar, P. Clapés, *Chem Eur J* **2009**, *15*, 3808-3816.
- [445] V. Hélaine, R. Mahdi, G. V. Sudhir Babu, V. de Berardinis, R. Wohlgemuth, M. Lemaire, C. Guérard-Hélaine, *Adv Synth Catal* **2015**, *357*, 1703-1708.
- [446] M. Schümperli, R. Pellaux, S. Panke, *Appl Microbiol Biotechnol* **2007**, *75*, 33.
- [447] R. Daniel, K. Stuert, G. Gottschalk, *J Bacteriol* **1995**, *177*, 4392-4401.
- [448] W. Kroutil, H. Mang, K. Edegger, K. Faber, *Adv Synth Catal* **2004**, *346*, 125-142.
- [449] A. H. Romano, T. Conway, *Res Microbiol* **1996**, *147*, 448-455.
- [450] T. K. Harris, R. N. Cole, F. I. Comer, A. S. Mildvan, *Biochemistry* **1998**, *37*, 16828-16838.
- [451] S. Wang, S. Vidal, in *Carbohydrate chemistry, Vol. 39*, The Royal Society of Chemistry, London, **2013**, pp. 78-101.
- [452] P. Compain, O. R. Martin, *Bioorg Med Chem* **2001**, *9*, 3077-3092.
- [453] K. Furukawa, Y. Ohkawa, Y. Yamauchi, K. Hamamura, Y. Ohmi, K. Furukawa, *Journal of Biochemistry* **2012**, *151*, 573-578.
- [454] P. A. Santacoloma, G. Sin, K. V. Gernaey, J. M. Woodley, *Organic Process Research & Development* **2011**, *15*, 203-212.
- [455] E. J. Corey, *Angew Chem Int Ed* **1991**, *30*, 455-465.
- [456] D. Romano, F. Bonomi, M. C. de Mattos, T. de Sousa Fonseca, M. d. C. F. de Oliveira, F. Molinari, *Biotechnol Adv* **2015**, *33*, 547-565.
- [457] U. T. Bornscheuer, *FEMS Microbiol Rev* **2002**, *26*, 73-81.
- [458] P. Stallforth, B. Lepenies, A. Adibekian, P. H. Seeberger, *Journal of Medicinal Chemistry* **2009**, *52*, 5561-5577.
- [459] I.-I. C. o. B. Nomenclature, *Eur J Biochem* **1999**, *264*, 610-650.
- [460] W. J. Quax, C. P. Broekhuizen, *Appl Microbiol Biotechnol* **1994**, *41*, 425-431.
- [461] C. José, M. V. Toledo, L. E. Briand, *Crit Rev Biotechnol* **2016**, *36*, 891-903.
- [462] S. Sood, A. Sharma, N. Sharma, S. Singh Kanwar, *Insights Enzyme Res* **2016**, *1*, 1-11.
- [463] M. Schmidt, E. Henke, B. Heinze, R. Kourist, A. Hidalgo, U. T. Bornscheuer, *Biotechnol J* **2007**, *2*, 249-253.
- [464] N. Krebsfänger, K. Schierholz, U. T. Bornscheuer, *J Biotechnol* **1998**, *60*, 105-111.

- [465] N. Krebsfänger, F. Zocher, J. Altenbuchner, U. T. Bornscheuer, *Enzyme Microb Technol* **1998**, *22*, 641-646.
- [466] G. J. Palm, E. Fernández-Álvarez, X. Bogdanović, S. Bartsch, J. Sczodrok, R. K. Singh, D. Böttcher, H. Atomi, U. T. Bornscheuer, W. Hinrichs, *Appl Microbiol Biotechnol* **2011**, *91*, 1061-1072.
- [467] J. Wu, M. H. Li, J. P. Lin, D. Z. Wei, *Curr Microbiol* **2011**, *62*, 1123-1127.
- [468] B. González-Flecha, B. Demple, *J Biol Chem* **1995**, *270*, 13681-13687.
- [469] A. T. Jacobs, L. J. Marnett, *Acc Chem Res* **2010**, *43*, 673-683.
- [470] S. B. Farr, T. Kogoma, *Microbiol Rev* **1991**, *55*, 561-585.
- [471] E. Hidalgo, B. Demple, *Adaptive responses to oxidative stress: The soxRS and oxyR regulons*, Springer, Boston, MA, USA, **1996**.
- [472] V. Vasilioiu, A. Pappa, T. Estey, *Drug Metab Rev* **2004**, *36*, 279-299.
- [473] C. Michalski, H. Mohagheghi, M. Nimtz, J. Pasteels, D. Ober, *J Biol Chem* **2008**, *283*, 19219-19228.
- [474] A. Weckbecker, W. Hummel, *Biocatal Biotrans* **2006**, *24*, 380-389.
- [475] C. Wuensch, H. Lechner, S. M. Glueck, K. Zangger, M. Hall, K. Faber, *ChemCatChem* **2013**, *5*, 1744-1748.
- [476] P. Schaaf, V. Gojic, T. Bayer, F. Rudroff, M. Schnürch, M. D. Mihovilovic, *Chem Commun* **2017**, *n/a*, n/a-n/a.
- [477] J. H. Sattler, M. Fuchs, K. Tauber, F. G. Mutti, K. Faber, J. Pfeiffer, T. Haas, W. Kroutil, *Angew Chem Int Ed* **2012**, *51*, 9156-9159.
- [478] S. Chong, F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus, M.-Q. Xu, *Gene* **1997**, *192*, 271-281.
- [479] T. Watanabe, Y. Ito, T. Yamada, M. Hashimoto, S. Sekine, H. Tanaka, *J Bacteriol* **1994**, *176*, 4465-4472.
- [480] S. Chong, G. E. Montello, A. Zhang, E. J. Cantor, W. Liao, M. Q. Xu, J. Benner, *Nucleic Acids Res* **1998**, *26*, 5109-5115.
- [481] L. Kirmair, A. Skerra, *Appl Environ Microbiol* **2014**, *80*, 2468-2477.
- [482] D. Schwendenwein, G. Fiume, H. Weber, F. Rudroff, M. Winkler, *Advanced Synthesis & Catalysis* **2016**, *358*, 3414-3421.
- [483] M. Winkler, C. K. Winkler, *Chem Monthly* **2016**, *147*, 575-578.
- [484] T. van Herk, A. F. Hartog, L. Babich, H. E. Schoemaker, R. Wever, *ChemBioChem* **2009**, *10*, 2230-2235.
- [485] G. Tasnádi, M. Lukesch, M. Zechner, W. Jud, M. Hall, K. Ditrach, K. Baldenius, A. F. Hartog, R. Wever, K. Faber, *Eur J Org Chem* **2016**, *2016*, 45-50.
- [486] M. Kickstein, correspondence with T. Bayer, TU Darmstadt / TU Wien, **2013**.
- [487] T. Wiesinger, *Ph.D. Thesis* **2017**, TU Wien, 1-200.
- [488] W. B. Freedberg, W. S. Kistler, E. C. C. Lin, *J Bacteriol* **1971**, *108*, 137-144.
- [489] S. Qamar, K. Marsh, A. Berry, *Protein Sci* **1996**, *5*, 154-161.
- [490] S. Cardinale, A. P. Arkin, *Biotechnol J* **2012**, *7*, 856-866.
- [491] U. Deuschle, W. Kammerer, R. Gentz, H. Bujard, *EMBO J* **1986**, *5*, 2987-2994.
- [492] G. L. Rosano, E. A. Ceccarelli, *Front Microbiol* **2014**, *5*, 172.
- [493] N. U. Gandhi, S. B. Chandra, *Acta Inform Med* **2012**, *20*, 167-173.
- [494] J. Lotze, U. Reinhardt, O. Seitz, A. G. Beck-Sickinger, *Mol Biosyst* **2016**, *12*, 1731-1745.
- [495] M. E. Kimple, A. L. Brill, R. L. Pasker, *Curr Protoc Protein Sci* **2013**, *73*, Unit-9.9.
- [496] T. P. Hopp, K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. Pat Cerretti, D. L. Urdal, P. J. Conlon, *Nat Biotech* **1988**, *6*, 1204-1210.
- [497] N. Tanaka, V. Dumay, Q. Liao, A. J. Lange, R. Wever, *Eur J Biochem* **2002**, *269*, 2162-2167.
- [498] J. P. Horwitz, J. Chua, M. Noel, J. T. Donatti, J. Freisler, *J Med Chem* **1966**, *9*, 447.
- [499] M. R. Sarikhani, M. A. Malboobi, N. Aliasgharzad, R. Greiner, B. Yakhchali, *Iran J Biotechnol* **2010**, *8*, 275-279.
- [500] R. Wever, correspondence with F. Rudroff, University of Amsterdam / TU Wien, **2013**.
- [501] E. Kuznetsova, M. Proudfoot, C. F. Gonzalez, G. Brown, M. V. Omelchenko, I. Borozan, L. Carmel, Y. I. Wolf, H. Mori, A. V. Savchenko, C. H. Arrowsmith, E. V. Koonin, A. M. Edwards, A. F. Yakunin, *J Biol Chem* **2006**, *281*, 36149-36161.
- [502] S. Wagner, L. Baars, A. J. Ytterberg, A. Klussmeier, C. S. Wagner, O. Nord, P.-Å. Nygren, K. J. van Wijk, J.-W. de Gier, *Mol Cell Proteomics* **2007**, *6*, 1527-1550.
- [503] P. Fu, S. Panke, *Systems and synthetic biology*, John Wiley & Sons, New York, **2009**.
- [504] R. Shetty, correspondence with T. Bayer, UC Berkeley / TU Wien, **2014**.
- [505] J. Mairhofer, A. Wittwer, M. Cserjan-Puschmann, G. Striedner, *ACS Synth Biol* **2015**, *4*, 265-273.
- [506] S. H. Yoon, M.-J. Han, H. Jeong, C. H. Lee, X.-X. Xia, D.-H. Lee, J. H. Shim, S. Y. Lee, T. K. Oh, J. F. Kim, *Genome Biol* **2012**, *13*, R37.
- [507] R. Z. Jin, E. C. Lin, *J Gen Microbiol* **1984**, *130*, 83-88.
- [508] M. Sudar, Z. Findrik, D. Vasic-Racki, A. Soler, P. Clapes, *RSC Advances* **2015**, *5*, 69819-69828.
- [509] B. Walder, *B.Sc. Thesis* **2016**, TU Wien / University of Vienna, 1-25.
- [510] R. Singh, R. J. Mailloux, S. Puiseux-Dao, V. D. Appanna, *J Bacteriol* **2007**, *189*, 6665-6675.
- [511] N. Tanaka, Z. Hasan, A. F. Hartog, T. van Herk, R. Wever, R. J. Sanders, *Org Biomol Chem* **2003**, *1*, 3470.
- [512] O. H. Lowry, J. Carter, J. B. Ward, L. Glaser, *J Biol Chem* **1971**, *246*, 6511-6521.
- [513] Y. Takatsuka, C. Chen, H. Nikaido, *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 6559-6565.
- [514] J. W. Lee, T. Y. Kim, Y.-S. Jang, S. Choi, S. Y. Lee, *Trends Biotechnol* **2011**, *29*, 370-378.
- [515] E. Nudler, M. E. Gottesman, *Genes Cells* **2002**, *7*, 755-768.
- [516] I. Kuzmine, P. A. Gottlieb, C. T. Martin, *Nucleic Acids Res* **2001**, *29*, 2601-2606.
- [517] L. E. Macdonald, Y. Zhou, W. T. McAllister, *J Mol Biol* **1993**, *232*, 1030-1047.
- [518] Estienne C. Swart, V. Serra, G. Petroni, M. Nowacki, *Cell* **2016**, *166*, 691-702.
- [519] C. Wokurek, *B.Sc. Thesis* **2017**, TU Wien, 1-37.
- [520] B. Canton, A. Labno, D. Endy, *Nat Biotechnol* **2008**, *26*.
- [521] W. Hilt, G. Pfeleiderer, P. Fortnagel, *Biochim Biophys Acta* **1991**, *1076*, 298-304.
- [522] P. Wolf, *Internal report* **2017**, TU Wien, 1-38.
- [523] R. F. Ramaley, N. Vasantha, *J Biol Chem* **1983**, *258*, 12558-12565.
- [524] F.-G. Meng, Y.-K. Hong, H.-W. He, A. E. Lyubarev, B. I. Kurganov, Y.-B. Yan, H.-M. Zhou, *Biophys J* **2004**, *87*, 2247-2254.
- [525] T. B. Fitzpatrick, N. Amrhein, P. Macheroux, *J Biol Chem* **2003**, *278*, 19891-19897.
- [526] T. Classen, J. Pietruszka, S. M. Schuback, *ChemCatChem* **2013**, *5*, 711-713.
- [527] C. K. Winkler, G. Tasnádi, D. Clay, M. Hall, K. Faber, *J Biotechnol* **2012**, *162*, 381-389.
- [528] M. Hall, C. Stueckler, H. Ehammer, E. Pointner, G. Oberdorfer, K. Gruber, B. Hauer, R. Stuermer, W. Kroutil, P. Macheroux, K. Faber, *Adv Synth Catal* **2008**, *350*, 411-418.
- [529] K. Tauber, M. Hall, W. Kroutil, W. M. F. Fabian, K. Faber, S. M. Glueck, *Biotechnol Bioeng* **2011**, *108*, 1462-1467.
- [530] S. K. Padhi, D. J. Bougioukou, J. D. Stewart, *J Am Chem Soc* **2009**, *131*, 3271-3280.

- [531] T. Bickle, *Nucleases, Vol. 26*, CHS, New York, **1993**.
- [532] E. A. Raleigh, G. Wilson, *PNAS* **1986**, *83*, 9070-9074.
- [533] E. A. Raleigh, *Mol Microbiol* **1992**, *6*, 1079-1086.
- [534] P. A. Waite-Rees, C. J. Keating, L. S. Moran, B. E. Slatko, L. J. Hornstra, J. S. Benner, *J Bacteriol* **1991**, *173*, 5207-5219.
- [535] J. E. Kelleher, E. A. Raleigh, *J Bacteriol* **1991**, *173*, 5220-5223.
- [536] A. J. Humphrey, N. J. Turner, R. McCague, S. J. C. Taylor, *Journal of the Chemical Society, Chemical Communications* **1995**, 2475-2476.
- [537] X. Garrabou, J. Calveras, J. Joglar, T. Parella, J. Bujons, P. Clapes, *Org Biomol Chem* **2011**, *9*, 8430-8436.
- [538] K. R. Albe, M. H. Butler, B. E. Wright, *J Theor Biol* **1990**, *143*, 163-195.
- [539] R. Daniel, K. Stuert, G. Gottschalk, *J Bacteriol* **1995**, *177*, 4392-4401.
- [540] D. E. Atkinson, *Biochem* **1968**, *7*, 4030-4034.
- [541] R. O. M. A. de Souza, L. S. M. Miranda, U. T. Bornscheuer, *Chem Eur J* **2017**, *23*, 12040-12063.
- [542] Y.-J. Chen, P. Liu, A. A. K. Nielsen, J. A. N. Brophy, K. Clancy, T. Peterson, C. A. Voigt, *Nat Meth* **2013**, *10*, 659-664.
- [543] F. W. Studier, *Protein Expr Purif* **2005**, *41*, 207-234.
- [544] T. C. Evans, J. Benner, M.-Q. Xu, *Protein Sci* **1998**, *7*, 2256-2264.
- [545] S. Chong, K. S. Williams, C. Wotkowicz, M.-Q. Xu, *J Biol Chem* **1998**, *273*, 10567-10577.
- [546] G. Volkmann, W. Sun, X.-Q. Liu, *Protein Sci* **2009**, *18*, 2393-2402.
- [547] A. Guagliardi, M. Martino, I. Iaccarino, M. D. Rosa, M. Rossi, S. Bartolucci, *Int J Biochem Cell Biol* **1996**, *28*, 239-246.
- [548] K. Edegger, C. C. Gruber, T. M. Poessl, S. R. Wallner, I. Lavandera, K. Faber, F. Niehaus, J. Eck, R. Oehrlein, A. Hafner, W. Kroutil, *Chem Commun* **2006**, 2402-2404.
- [549] P. Schaaf, *Ph.D. Thesis* **2017**, TU Wien, 1-181.
- [550] M. Winkler, correspondence with T. Bayer, ACIB Graz / TU Wien, **2016**.
- [551] M. Sauer, *Methods in microbiology*, W.H. Freeman, New York, **2012**.
- [552] M. J. Fink, *Ph.D. Thesis* **2015**, TU Wien, 1-344.
- [553] M. D. Mihovilovic, correspondence with T. Bayer, TU Wien, **2017**.
- [554] S. Feroz, *Ph.D. Thesis* **2013**, TU Wien, 1-270.
- [555] F. Steffen-Munsberg, C. Vickers, A. Thontowi, S. Schätzle, T. Meinhardt, M. Svedendahl Humble, H. Land, P. Berglund, U. T. Bornscheuer, M. Höhne, *ChemCatChem* **2013**, *5*, 154-157.
- [556] D. Rother, G. Kolter, T. Gerhards, C. L. Berthold, E. Gauchenova, M. Knoll, J. Pleiss, M. Müller, G. Schneider, M. Pohl, *ChemCatChem* **2011**, *3*, 1587-1596.
- [557] E.-Y. Jeon, J.-H. Seo, W.-R. Kang, M.-J. Kim, J.-H. Lee, D.-K. Oh, J.-B. Park, *ACS Catal* **2016**, *6*, 7547-7553.
- [558] T. H. M. Smits, M. A. Seeger, B. Witholt, J. B. van Beilen, *Plasmid* **2001**, *46*, 16-24.