



THESIS

3

2019

LIBRARY
Michigan State
University

This is to certify that the
dissertation entitled

Syntheses and Downstream Purification of 1,2,4-Butanetriol

presented by

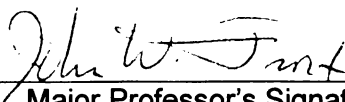
Man Kit Lau

has been accepted towards fulfillment
of the requirements for the

Doctoral

degree in

Chemistry



Major Professor's Signature

7/13/07

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

SYNTHESES AND DOWNSTREAM PURIFICATION OF 1,2,4-BUTANETRIOL

By

Man Kit Lau

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

2007

ABSTRACT

SYNTHESES AND DOWNSTREAM PURIFICATION OF 1,2,4-BUTANETRIOL

By

Man Kit Lau

An extensive application of the thermally stable, high energetic material 1,2,4-butanetriol trinitrate, is hindered by the lack of an economic route to synthesize its precursor, 1,2,4-butanetriol. Our goal is to design an efficient synthesis of 1,2,4-butanetriol using low-cost, renewable and abundantly available starting material. The application of cutting-edge molecular biology offers opportunities to create 1,2,4-butanetriol biocatalytic routes that can ultimately be scaled-up to commercially viable processes. 2-Deoxyribose-5-phosphate aldolase condenses acetaldehyde and glycoaldehyde into 3,4-dihydroxy-D-butanal. The success in coupling this enzymatic reaction with a Ru/C catalyzed hydrogenation yielded D-1,2,4-butanetriol in 30% yield. This chemoenzymatic synthesis represents the shortest route to synthesize a single 1,2,4-butanetriol enantiomer under relatively mild reaction conditions using inexpensive starting materials.

As an alternative to enzymatic processes, a large portion of this thesis was devoted to explore and improve fermentation processes using *E. coli* microbial catalysts that in a single step converts renewable carbohydrates into 1,2,4-butanetriol. To explore novel microbial alternatives, a biosynthesis was created to synthesize 1,2,4-butanetriol in a two-step conversion starting from erythritol. However, the identified erythritol dehydratase activity was not sufficient enough to elaborate this biosynthesis into a

microbial process. Earlier research efforts in the Frost group created a single *E. coli* microbe WN13/pWN7.126B that synthesized D-1,2,4-butanetriol from D-xylose at a concentration of 10.5 g/L when cultivating under an improved fermentation conditions. Attempts to identify a better 3-deoxy-*glycero*-pentulose decarboxylase did not lead to significantly improved alternatives. Approaches to improve *P. putida mdhC* gene encoding benzoylformate decarboxylase expression resulted in metabolic burden. *E. coli adhP* encoding alcohol dehydrogenase was identified to be involved in the conversion of 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol. *E. coli* microbe KIT10/pWN7.126B synthesized only 6.5 g/L D-1,2,4-butanetriol. A second generation D-1,2,4-butanetriol biocatalyst KIT18/pWN7.126B was created by overexpressing *adhP* chromosomally with two 2-keto acid dehydrogenase (YiaE and YcdW) knockouts. More recently, cultivation of *E. coli* KIT18/pWN7.126B yielded D-1,2,4-butanetriol at a concentration of 18 g/L in 50% (mol/mol) yield from D-xylose.

A downstream purification method for D-1,2,4-butanetriol was developed using Karr reciprocating column. Using 2-butanol as solvent, D-1,2,4-butanetriol was successfully recovered in 74% from the fermentation broth. Salt-free D-1,2,4-butanetriol with >99% purity was obtained by Karr column extraction and negative adsorption using Dowex 1×8 (OH⁻) followed by Dowex 50 (H⁺) ion-exchange resins.

Copyright by

Man Kit Lau

2007

To my parents
For their love and support

ACKNOWLEDGMENTS

This dissertation would not have been possible without the help of so many people. I would first like to express my warmest appreciation to Professor John W. Frost. His enthusiasm, integral view of science, and constructive criticism have all contributed a great deal to the success of my research. I would also like to thank the members of my graduate committee, Prof. Chris C. K. Chang, Prof. Babak Borhan, and Prof. Kevin D. Walker for their interest and intellectual input during the preparation of this thesis.

I am deeply indebted to Dr. Karen M. Draths for her guidance and invaluable inputs to my research. I wish to thank Carolyn Wemple for resolving most of the bureaucratic issues I have encountered throughout my studious years at MSU. There are no better colleagues and friends than Justas Jancauskas and Dr. Jihane Achkar, whom I must thank for their devoted friendship, encouragement and support. Brad Cox who struggled with my writing style certainly helped to put this thesis together. My gratitude is given to other Frost group members, past and present, including Dr. Vu Bui, Dr. Stephen Cheley, Dr. Sigal Lechno-Yossef, Dr. Wensheng Li, Dr. Jiantao Guo, Dr. Mapitso Molefe, Dr. Wei Niu, Dr. Ningqing Ran, Dr. Heather Stueben, Dr. Dongming Xie, Dr. Jinsong Yang, Dr. Jian Yi and Xiaofei Jia for providing a pleasant research environment. I am also grateful to my dear friends Cangel, Fei, Harrison, Sewa, Sing, Kostas, Thalia, Nineta, Chrysoula and Jennie for their friendship and help. The jokes and time we shared will never be forgotten. Finally, the love and support I have been given by my parents Joseph and Doris has been overwhelming. Even during the toughest time, I have always felt secure knowing they were there for me. Without all of you I would never have found my way.

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xix
 CHAPTER ONE	 1
Introduction.....	1
Metabolic pathway engineering for microbial catalysts	4
Shikimic acid and quinic acid	4
Indigo dye.....	7
1,3-Propanediol	8
The significance of 1,2,4-butanetriol and 1,2,4-butanetriol trinitrate	9
Chemical syntheses of 1,2,4-butanetriol	12
Creation of a 1,2,4-butanetriol biosynthetic pathway	14
A single microbe for D-1,2,4-butanetriol synthesis.....	16
Important chemicals derived from the microbial 1,2,4-butanetriol synthesis.	18
3,4-Dihydroxy-D-butanoic acid.....	18
3,4-Dihydroxy- <i>threo</i> -L-norvaline	19
REFERENCE.....	20
 CHAPTER TWO	 23
Chemo-Enzymatic Synthesis of D-1,2,4-Butanetriol Using 2-Deoxyribose-5-Phosphate Aldolase.....	23
Background	23
Synthesis of 3,4-dihydroxy-D-butanal using 2-deoxyribose-5-phosphate aldolase	30
Protein purification of 2-deoxyribose-5-phosphate aldolase	30
Chemical synthesis of 3,4-dihydroxy-D-butanal.....	33
Enzymatic synthesis of 3,4-dihydroxy-D-butanal	34
Catalytic hydrogenation of racemic 3,4-dihydroxybutanal.....	36
Chemo-enzymatic synthesis of D-1,2,4-butanetriol	38
Discussion and future work.....	40
REFERENCE.....	44
 CHAPTER THREE	 46
Improving Microbial Synthesis of 1,2,4-Butanetriol.....	46
Overview	46
Artificial biosynthesis of 1,2,4-butanetriol from erythritol.....	48
Background	48
Identification of erythritol dehydratase candidates	52
Plasmid Constructions.....	55

Enzyme assays for <i>in vitro</i> erythritol dehydratase activity	63
1,2,4-Butanetriol biosynthesis using erythritol as substrate.....	65
In search of 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase genes for the microbial synthesis of 1,2,4-butanetriol from pentoses	72
Background	72
Cloning and Screening for 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase candidates.....	76
Site-directed Mutagenesis of benzoylformate and pyruvate decarboxylase gene	85
Enzyme activities for 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase candidates.....	88
Codon usage	92
Amino acid and nucleotide sequence identity.....	92
Identification of novel benzoylformate decarboxylases	93
Directed evolution of 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase for microbial 1,2,4-butanetriol synthesis	95
Background	95
Subcloning 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase encoded genes	98
Chimera library construction for ssDNA family shuffling	102
Chimera library construction for dsDNA family shuffling.....	103
Chimera library screening for 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase.....	106
Bacteriophage T7 promoter and <i>E. coli</i> chaperones for <i>mdlC</i> expression....	112
Background	112
Expression of plasmid-borne <i>mdlC</i> gene under phage T7 promoter.....	113
<i>E. coli</i> chaperonins GroES-GroEL for <i>mdlC</i> gene expression.....	120
Discovery of alcohol dehydrogenases for D-1,2,4-butanetriol biosynthesis.	127
Background	127
Cloning and screening for D-1,2,4-butanetriol dehydrogenase candidates	127
Chromosomal <i>adhP</i> gene inactivation	134
Plasmid-based expression of <i>adhP</i> gene	137
Chromosomal expression of <i>E. coli adhP</i>	147
Construction of the second generation D-1,2,4-butanetriol synthesizing <i>E.</i> <i>coli</i> microbe.....	156
Future work	158
Directed Evolution of 3-deoxy-D- <i>glycero</i> -pentulosonate decarboxylase .	158
Chemo-microbial synthesis of (<i>S</i>)-3-hydroxy- γ -butyrolactone.....	161
Enzyme activity studies for <i>E. coli adhP</i> expressing microbes	164
REFERENCE.....	167
 CHAPTER FOUR.....	173
Downstream Purification of 1,2,4-Butanetriol	173
Background	173
Solvent candidates screening using continuous extraction method	179

Residual protein removal by membrane ultra-filtration.....	179
Liquid-liquid extraction using a single solvent system.....	181
Liquid-liquid extraction using a solvent-blend system	184
Development of a scalable extraction process using a bench-top Karr column	187
Bench-top Karr column unit setup	187
Distribution coefficient for solvent candidates	189
Start-up conditions for Karr column extractions.....	190
Preliminary extraction attempts using Karr column	192
Further optimization of Karr column extraction	194
Final step toward purified D-1,2,4-butanetriol	196
Alternative method for D-1,2,4-butanetriol purification	199
Future plan	201
REFERENCE.....	203
 CHAPTER FIVE	 204
Experimental	204
General chemistry	204
Reagents and solvents	204
Chromatography	205
Spectroscopic measurements	206
Microbial strains and plasmids	206
Storage of microbial strains and plasmids	209
Culture medium	210
General fed-batch fermentation conditions.....	212
Analysis of fermentation broth	213
Genetic manipulations	214
Large scale purification of plasmid DNA	216
Small scale purification of plasmid DNA	217
Restriction enzyme digestion of DNA	218
Determination of DNA concentration.....	219
Agarose gel electrophoresis	219
Isolation of DNA from agarose.....	219
Treatment of DNA with Klenow fragment.....	220
Treatment of vector DNA with calf intestinal alkaline phosphatase	221
Ligation of DNA	221
Preparation and transformation of competent <i>E. coli</i> cells	222
Purification of <i>E. coli</i> genomic DNA	224
P1 phage-mediated transduction	226
Enzyme assays	228
2-Deoxyribose-5-phosphate aldolase assay	228
Glycerol and diol dehydratase assay	228
2-Keto acid decarboxylase assay	229
Alcohol dehydrogenase assay	230
Protein SDS-PAGE analysis	231

Chapter 2.....	232
Purification of 2-deoxyribose-5-phosphate aldolase.....	232
Catalytic hydrogenation of 3,4-dihydroxy-D-butanal	234
Chapter 3	235
Artificial biosynthesis of 1,2,4-butanetriol.....	235
Plasmid pML5.176.....	235
Plasmid pML5.179.....	235
Plasmid pML5.200.....	235
Plasmid pML5.226.....	236
Plasmid pWN5.220A	236
Plasmid pWN5.260A	236
Plasmid pWN5.258A	237
Anaerobic cultivation of <i>Klebsiella</i> and <i>Lactobacillus</i> strains	237
In vitro dehydratase reaction of erythritol.....	238
In search of 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase genes for the microbial synthesis of 1,2,4-butanetriol from pentoses	238
Plasmid pML2.118.....	238
Plasmid pML2.123.....	238
Plasmid pML2.162.....	239
Plasmid pML2.208.....	239
Plasmid pML2.214.....	239
Plasmid pML2.256.....	240
Plasmid pML2.286.....	240
Plasmid pML3.040.....	240
Plasmid pML3.054.....	241
Plasmid pML3.062.....	242
Plasmid pML3.084.....	242
Plasmid pML3.086.....	243
Plasmid pML3.104.....	243
Directed evolution of 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase for microbial 1,2,4-butanetriol synthesis.....	244
Plasmid pML3.200.....	244
Plasmid pML3.224.....	244
Plasmid pML3.225.....	245
Plasmid pML3.226.....	245
Chimera library generation for ssDNA family shuffling.....	245
Chimera library generation for dsDNA family shuffling.....	247
Library screening for 3-deoxy- <i>glycero</i> -pentulosonate decarboxylase	248
Bacteriophage T7 promoter and <i>E. coli</i> chaperones for <i>mdlC</i> expression....	250
<i>E. coli</i> KIT2.....	250
<i>E. coli</i> KIT15.....	251
Plasmid pML7.128.....	251
Plasmid pML7.135.....	251
Plasmid pML7.166.....	252
Plasmid pML7.175.....	252
Plasmid pML7.180.....	252

Plasmid pML7.202.....	253
Discovery of alcohol dehydrogenases for D-1,2,4-butanetriol biosynthesis.....	253
<i>E. coli</i> KIT1.....	253
<i>E. coli</i> KIT3.....	254
<i>E. coli</i> KIT4.....	254
<i>E. coli</i> KIT5.....	255
<i>E. coli</i> KIT6.....	256
<i>E. coli</i> KIT7.....	256
<i>E. coli</i> KIT8.....	256
<i>E. coli</i> KIT9.....	257
<i>E. coli</i> KIT10.....	257
<i>E. coli</i> KIT16.....	258
<i>E. coli</i> KIT17.....	258
<i>E. coli</i> KIT18.....	259
Plasmid pML5.179.....	259
Plasmid pML6.090.....	259
Plasmid pML6.128.....	260
Plasmid pML6.133.....	260
Plasmid pML6.135.....	260
Plasmid pML6.166.....	260
Plasmid pML6.168.....	261
Plasmid pML6.185.....	261
Plasmid pML6.195.....	261
Plasmid pML6.255.....	262
Plasmid pML6.259.....	262
Plasmid pML6.261.....	262
Plasmid pML6.263.....	263
Plasmid pML6.272.....	263
Plasmid pML7.042.....	263
REFERENCE.....	265

LIST OF TABLES

Table 1. Purification table of 2-deoxyribose-5-phosphate aldolase.....	32
Table 2. Purification table of 2-deoxyribose-5-phosphate aldolase.....	33
Table 3. Influence of 2-deoxyribose-5-phosphate aldolase (DERA) purity on 3,4-dihydroxy-D-butanal yield.	35
Table 4. Catalytic hydrogenation of authentic 3,4-dihydroxybutanal.	38
Table 5. Glycerol and diol dehydratase candidates	54
Table 6. <i>In vitro</i> dehydratase activities.....	64
Table 7. 1,3-Propanediol oxidoreductase activities.	68
Table 8. <i>In vitro</i> enzymatic reactions using erythritol as substrate.....	69
Table 9. 3-Deoxy- <i>glycero</i> -pentulosonate decarboxylase candidates.....	85
Table 10. Forward primers used to generate mutations.....	88
Table 11. Activity table for wild-type pJF118EH clones.	90
Table 12. Activity table for benzoylformate and pyruvate decarboxylase mutants.	90
Table 13. Amino acid composition of the gene candidates.	91
Table 14. Rare codon analysis of the gene candidates.....	91
Table 15. Nucleotide and amino acid (in parentheses) sequence identity.	93
Table 16. Data reproducibility using <i>in vivo</i> GC screening assay.	108
Table 17. Test for false-negative results.....	108
Table 18. Test for mutants with growth issues.	108
Table 19. Shake flask experiments for improved decarboxylase activities.....	109
Table 20. <i>E. coli</i> alcohol dehydrogenase candidates.	134
Table 21. <i>E. coli</i> strains for chromosomal <i>adhP</i> inactivation study.	135
Table 22. Other alcohol dehydrogenase candidates.....	137

Table 23. <i>E. coli</i> strains for chromosomal <i>adhP</i> expression study.....	149
Table 24. Creation of the second generation D-1,2,4-butanetriol synthesizing <i>E. coli</i>	158
Table 25. Preliminary results on the cultivation of <i>E. coli</i> KIT18/pWN7.126B under fermentor-controlled conditions.....	158
Table 26. Oxidative decarboxylation of 3-deoxy-D- <i>glycero</i> -pentulosonic acid.....	163
Table 27. <i>E. coli</i> strains with <i>adhP</i> gene expression.	166
Table 28. Methods for residual protein removal from fermentation broth.	181
Table 29. Single solvent candidates.....	182
Table 30. Summary of single solvent screening experiments.....	183
Table 31. Liquid-liquid extraction using 2-butanone/ethyl acetate solvent-blend.	185
Table 32. Liquid-liquid extraction using ethanol/ethyl acetate solvent blend.....	186
Table 33. Distribution coefficients for various solvent candidates.....	189
Table 34. Preliminary Karr column extraction studies.	194
Table 35. Further optimization for Karr column extraction.	195
Table 36. Dowex 1×8 (OH ⁻) for extraction-free purification.	201
Table 37. D-1,2,4-Butanetriol recovery from Dowex 1×8 (OH ⁻).	201
Table 38. Microbial strains and plasmids.	206

LIST OF FIGURES

Figure 1. The shikimate pathway and biosynthesis of quinic acid.	5
Figure 2. Synthesis of hydroquinone from glucose.	7
Figure 3. Chemical and microbial syntheses of indigo.....	8
Figure 4. Schematic of the reaction network from D-glucose to 1,3-propanediol.....	10
Figure 5. Industrial synthesis of 1,2,4-butanetriol trinitrate.	11
Figure 6. Syntheses of 1,2,4-butanetriol.	12
Figure 7. Routes to 1,2,4-butanetriol previously developed in the Frost lab.....	14
Figure 8. Microbial syntheses of D- and L-1,2,4-butanetriol.	15
Figure 9. Native <i>E. coli</i> D-xylonic acid catabolic pathway.....	17
Figure 10. Microbial synthesis of D-1,2,4-butanetriol using WN13/pWN7.126B.	17
Figure 11. Drugs synthesized from 3,4-dihydroxy-D-butanoic acid and 4,5-dihydroxy- <i>threo</i> -L-norvaline.....	19
Figure 12. <i>In vivo</i> reaction catalyzed by 2-deoxyribose-5-phosphate aldolase.	23
Figure 13. 2-Deoxyribose-5-phosphate aldolase catalyzed tandem aldol reaction.....	24
Figure 14. Retro-synthesis of atorvastatin (Lipitor).	25
Figure 15. Catalytic hydrogenation of malic acid.....	27
Figure 16. Proposed chemo-enzymatic reaction using 2-deoxyribose-5-phosphate aldolase.	28
Figure 17. Industrial syntheses of acetaldehyde and glycolaldehyde.	29
Figure 18. Construction of plasmid pVH17.....	31
Figure 19. SDS-PAGE of 2-deoxyribose-5-phosphate aldolase purified from <i>E. coli</i> DH5 α	32
Figure 20. Chemical synthesis of racemic 3,4-dihydroxybutanal.....	34
Figure 21. NaBH ₄ reaction for 3,4-dihydroxy-D-butanal quantification..	34

Figure 22. Reaction parameters for the chemo-enzymatic synthesis of D-1,2,4-butanetriol.....	39
Figure 23. Proposed artificial biosynthetic pathway for 1,2,4-butanetriol from erythritol.	46
Figure 24. Anaerobic utilization of glycerol.....	49
Figure 25. Anaerobic utilization of 1,2-propanediol.	50
Figure 26. Plasmid maps of pCS120, pFL1 and pXY39.	56
Figure 27. Construction of plasmid pML5.176.	58
Figure 28. Constuction of plasmid pML5.200.....	59
Figure 29. Construction of plasmid pWN5.220A.	60
Figure 30. Construcion of plasmid pWN5.260A.	61
Figure 31. Construction of plasmid pWN5.258A.	62
Figure 32. <i>In vitro</i> dehydratase assays using different substrates.....	64
Figure 33. Construction of plasmid pML5.179.	66
Figure 34. Construction of plasmid pML5.226.	67
Figure 35. GC chromatogram for reaction 1 (control).....	69
Figure 36. GC chromatograms for reaction 2.	70
Figure 37. GC chromatograms for reaction 4.	71
Figure 38. Restriction enzyme map of plasmids.....	72
Figure 39. Mandelate pathway in <i>P. putida</i>	74
Figure 40. Reaction mechanism for the benzoylformate decarboxylase-catalyzed decarboxylation of benzoylformate.....	74
Figure 41. Construction of plasmid pML2.118.	77
Figure 42. Construction of plasmid pML2.123.	78
Figure 43. Construction of plasmid pML2.162.	79
Figure 44. Construction of plasmid pML3.040.	80

Figure 45. Construction of plasmid pML2.208.	81
Figure 46. Construction of plasmid pML2.214.	82
Figure 47. Construction of plasmid pML2.256.	83
Figure 48. Construction of plasmid pML2.286.	84
Figure 49. QuikChange Site directed mutagenesis.	87
Figure 50. Coupled enzyme assay for decarboxylase activities.	89
Figure 51. Multiple sequence alignment using ClustalW algorithm.	94
Figure 52. Schematic diagram for single-stranded DNA family shuffling.....	97
Figure 53. Construction of plasmid pML3.224.	99
Figure 54. Construction of plasmid pML3.225.	100
Figure 55. Construction of plasmid pML5.226.	101
Figure 56. Sequencing results of 20 randomly chosen clones from ssDNA family shuffling.....	103
Figure 57. DNA agarose gels for dsDNA family shuffling.	104
Figure 58. Sequencing results of 22 randomly chosen clones from dsDNA family shuffling.....	105
Figure 59. Sequence alignment of selected mutants.	111
Figure 60. Construction of plasmid pML7.128.	114
Figure 61. Construction of plasmid pML7.135.	115
Figure 62. Biosynthesis of 3-deoxy-D- <i>glycero</i> -pentulosonic acid under fermentor- controlled conditions.	119
Figure 63. Construction of plasmid pML7.166.	121
Figure 64. Construction of plasmid pML7.175.	122
Figure 65. Construction of plasmid pML7.180.	123
Figure 66. Construction of plasmid pML7.202.	124
Figure 67. Cultivation of chaperonin <i>groEL-groES</i> constructs under fermentor-controlled conditions.	126

Figure 68. Construction of plasmid pML6.166.	129
Figure 69. Construction of plasmid pML6.168.	130
Figure 70. Construction of plasmid pML6.259.	131
Figure 71. Construction of plasmid pML6.261.	132
Figure 72. Construction of plasmid pML2.263.	133
Figure 73. Cultivation of <i>E. coli</i> strains with <i>adhP</i> gene inactivation under fermentor-controlled conditions.	136
Figure 74. Construction of plasmid pML5.179.	138
Figure 75. Construction of plasmid pML6.185.	140
Figure 76. Construction of plasmid pML6.195.	141
Figure 77. Construction of plasmid pML6.133.	142
Figure 78. Construction of plasmid pML6.090.	143
Figure 79. Construction of plasmid pML6.128.	144
Figure 80. Construction of plasmid pML6.135.	145
Figure 81. Cultivation of <i>E. coli</i> constructs expressing plasmid-borne alcohol dehydrogenases under fermentor-controlled conditions.	146
Figure 82. Strategy to integrate <i>xdh-adhP-P_{tac}</i> gene cassette into <i>E. coli</i> W3110.	149
Figure 83. Construction of plasmid pML6.255.	150
Figure 84. Construction of plasmid pML6.272.	151
Figure 85. Construction of plasmid pML7.042.	153
Figure 86. Cultivation of <i>E. coli</i> constructs having a chromosomal <i>P_{tac}-adhP</i> gene insertion under fermentor-controlled conditions.	154
Figure 87. Cultivation of <i>E. coli</i> constructs having a chromosomal T5 promoter insertion under fermentor-controlled conditions.	155
Figure 88. Proposed chemo-microbial synthesis of (<i>S</i>)-3-hydroxy- γ -butyrolactone.	162
Figure 89. Enzyme activity time-course for microbial syntheses of D-1,2,4-butanetriol under fermentor-controlled conditions.	166

Figure 90. Single-stage liquid-liquid extractors.....	176
Figure 91. Agitated counter-current extraction columns.	178
Figure 92. Glass continuous liquid-liquid extractor (Sigma-Aldrich).	179
Figure 93. D-1,2,4-Butanetriol purified by 50% 2-butanone in ethyl acetate extraction.	185
Figure 94. D-1,2,4-Butanetriol purified by 10% ethanol in ethyl acetate extraction.	186
Figure 95. Benchtop Karr column in action.....	187
Figure 96. Kremser equation.....	197
Figure 97. Kremser type plot.	197
Figure 98. GC traces for Karr column extraction.	198
Figure 99. Large scale purification of <i>E. coli</i> WN13/pWN7.126B fermentation broth. .	199
Figure 100. Cyclic acetal formation of 1,2,4-butanetriol.....	202

LIST OF ABBREVIATIONS

ADH	alcohol dehydrogenase
Ap	ampicillin
ATP	adenosine triphosphate
bp	base pair
BSTFA	<i>N,O</i> -bis(trimethylsilyl)trifluoroacetamide
BT	1,2,4-butanetriol
CIAP	calf intestinal alkaline phosphatase
Cm	chloramphenicol
DEAE	diethylaminoethyl
DERA	2-deoxyribose-5-phosphate aldolase
DHBA	3,4-dihydroxybutanoic acid
DGP	3-deoxy- <i>glycero</i> -pentulosonic acid
DHN	4,5-dihydroxy- <i>threo</i> -norvaline
DNA	3-deoxyribonucleic acid
D.O.	dissolved oxygen
DPA	3-deoxy- <i>glycero</i> -pentanoic acid
Flp	flipase
FRT	flipase recognition target sequence
G3P	<i>sn</i> -glycerol 3-phosphate
GAP	D-glyceraldehyde 3-phosphate
GC	gas chromatography

GRAS	generally regard as safe
h	hour
His	L-histidine
HPLC	high pressure liquid chromatography
IPTG	isopropyl β -D-thiogalactopyranoside
Km	kanamycin
kg	kilogram
LB	Luria-Bertani broth
M	molar
mg	milligram
mL	milliliter
μ L	microliter
mM	millimolar
min	minute
NAD	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form
NADP	nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NMR	nuclear magnetic resonance
OD	optical density
ORF	open reading frame
PEG	polyethylene glycol
PEP	phosphoenolpyruvate

PCR	polymerase chain reaction
ppm	parts per million
psi	pounds per square inch
PTS	phosphoenolpyruvate:carbohydrate phosphotransferase systems
rpm	revolutions per minute
rt	room temperature
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
TCA cycle	tricarboxylic acid cycle
TSP	sodium 3-(trimethylsilyl)propionic-2,2,3,3- <i>d</i> ₄
UV	ultraviolet

CHAPTER ONE

Introduction

Industrial fermentation processes have existed for thousands of years to meet various human needs. Traditionally, most were accidentally discovered by mankind and incorporated into daily life to produce foods and beverages such as bread, wine, beer, vinegar and cheese.¹ The idea behind this technology was not clearly understood until the nineteenth century when a French chemist named Louis Pasteur first connected microorganisms to fermentation. His pioneering work that included the discovery of yeast, lactic acid bacillus, development of 'germ theory', and the invention of the Pasteurization process laid the foundation for modern biotechnology.² Modern industrial biocatalysis began when the first large scale fermentations were developed for the manufacturing of solvents, organic acids, vitamins etc. In the 1940's, antibiotic fermentations came into existence. Sir Alexander Fleming discovered microbe-synthesized penicillin, and its scale-up fermentation process was developed jointly by Merck & Co. and Princeton University.³ In 1972, the first recombinant DNA plasmid was made in the laboratories of Stanford University and University of California at San Francisco.³ In 1975, first DNA sequencing technique was established.⁴ In 1988, Kary Mullis published the first paper on the application of polymerase chain reaction (PCR).⁵ Since then, molecular biology has revolutionized traditional microbiology and established a new global biotechnology industry.

A prominent feature by which enzymes are favorably distinguished from chemical catalysts is their specificity. While enzymes are used to increase the regioselectivity of a reaction, their greatest advantage lies in the ability to differentiate between enantiomeric substrates, resulting in enantiomerically pure products. In addition, the high specificity of enzymes do not require highly purified substrate stream. Another advantage of biocatalysis over a chemical route is its mild, near ambient reaction conditions of temperature and pH. The preference of an aqueous reaction medium is often regarded as an important benefit to develop an environmentally benign manufacturing process. The possibility of using simple and inexpensive carbohydrates such as D-glucose and D-xylose as starting material makes this biocatalytic approach economically attractive and environmentally acceptable.

In terms of applications, enzymes have long been known to compete well with chemical methods for resolution. Ibuprofen, phenylethylamine and acrylamide are commonly cited in literature as examples that are synthesized using enzyme-based processes.⁶ With the advances in technology, such as enzyme immobilization in amino acids manufacturing,^{7, 8} new methods for accessing biodiversity to identify novel enzymes,^{7,9} and more recently, by directed evolution strategies to optimize enzymes for existing and new applications,¹⁰ new cutting-edge applications continue to be discovered. For instance, a fed-batch process of L-DOPA using *Erwinia herbicola* has been developed (Ajinomoto Co. Ltd.) to give a final product titer of 110 g/L.¹¹ This single step reaction utilizes catechol, pyruvic acid and ammonia as starting materials to afford L-DOPA, producing water as the only byproduct in a short production cycle of 3 days. This process provides a relatively cheaper route to L-DOPA, which is a metabolic precursor of

dopamine, an important drug for the treatment of Parkinson's disease. The other example is the microbial synthesis of shikimic acid developed in the Frost group.¹² Shikimic acid is the starting material for the manufacture of the anti-influenza drug Tamiflu. Traditionally, shikimic acid was extracted from star anise seeds. The Frost group has optimized an alternative process using a microbe to synthesize 60 g/L concentration of shikimic acid from glucose.¹² These examples show that biocatalysis has been evolving into a highly promising area of research, and the new generation of bioprocesses is becoming more competitive with conventional routes.

This dissertation presents research results for the syntheses of D-1,2,4-butanetriol using different methodologies. In chapter 2 of this thesis, a biocatalytic alternative to the industrially employed chemical synthesis of 1,2,4-butanetriol is described. A chemo-enzymatic approach utilized *E. coli deoC* encoding deoxyribose-5-phosphate aldolase (DERA) to condense acetaldehyde and glycoaldehyde into 3,4-dihydroxy-D-butanal¹³ followed by catalytic hydrogenation to form D-1,2,4-butanetriol. In chapter 3, development of a novel biosynthetic route to produce 1,2,4-butanetriol from erythritol will be discussed. Different approaches to improve the microbial synthesis of D-1,2,4-butanetriol from D-xylose using genetic engineered *E. coli* WN13/pWN7.126B¹⁴ will be presented as well. In contrast to the chemo-enzymatic approach, whole cell microbial synthesis exploits the metabolism of the microbe to supply necessary co-substrates and to regenerate cofactors. Although genetic manipulation of the microbe is necessary, at the end of the process, the organism produced can self-regenerate. In chapter 4, a downstream purification of D-1,2,4-butanetriol will be presented. Recovery of polyols from the complex and dilute fermentation broth is a challenging task due to relatively high boiling

points and inherently high hydrophilicity of related compounds. An efficient purification protocol is therefore critical for the development of a commercially viable bioconversion process for 1,2,4-butanetriol synthesis. To this end, a liquid-liquid extraction methodology was developed using a bench-top Karr reciprocating plate extraction column.

Metabolic pathway engineering for microbial catalysts

By manipulating metabolic pathways in microorganisms, a wide range of natural compounds, such as amino acids, organic acids, and nucleotides, can be produced. Essential for the successful production of chemicals, however, will be the integration of metabolic engineering with biochemical engineering. Successful industrially viable bioprocess development will achieve improvements in several aspects such as higher catalyst turnover, more efficient channeling of carbon and most importantly, higher product yield. Three independent cases will be discussed as examples.

Shikimic acid and quinic acid

Shikimic acid is the starting material for the synthesis of the neuraminidase inhibitor Tamiflu, which is an anti-influenza drug. Shikimic acid has traditionally been extracted from star anise seeds. More recently, fermentor-controlled microbial synthesis provides an alternative avenue to this molecule. Currently, these two routes supply the global demand for shikimic acid. The host strain used for microbial synthesis of shikimic acid is a genetically engineered *E. coli* that lacks shikimate kinase activity (Figure 1).¹² The first reported shikimate-producing *E. coli* synthesized 27 g/L shikimic acid in a 1 L

high density cell culture. After further genetic manipulation of the shikimate pathway and optimization of fermentor-controlled culture conditions, the current process can produce shikimic acid at a concentration of 62 g/L in yield from glucose of 33% (mol/mol).

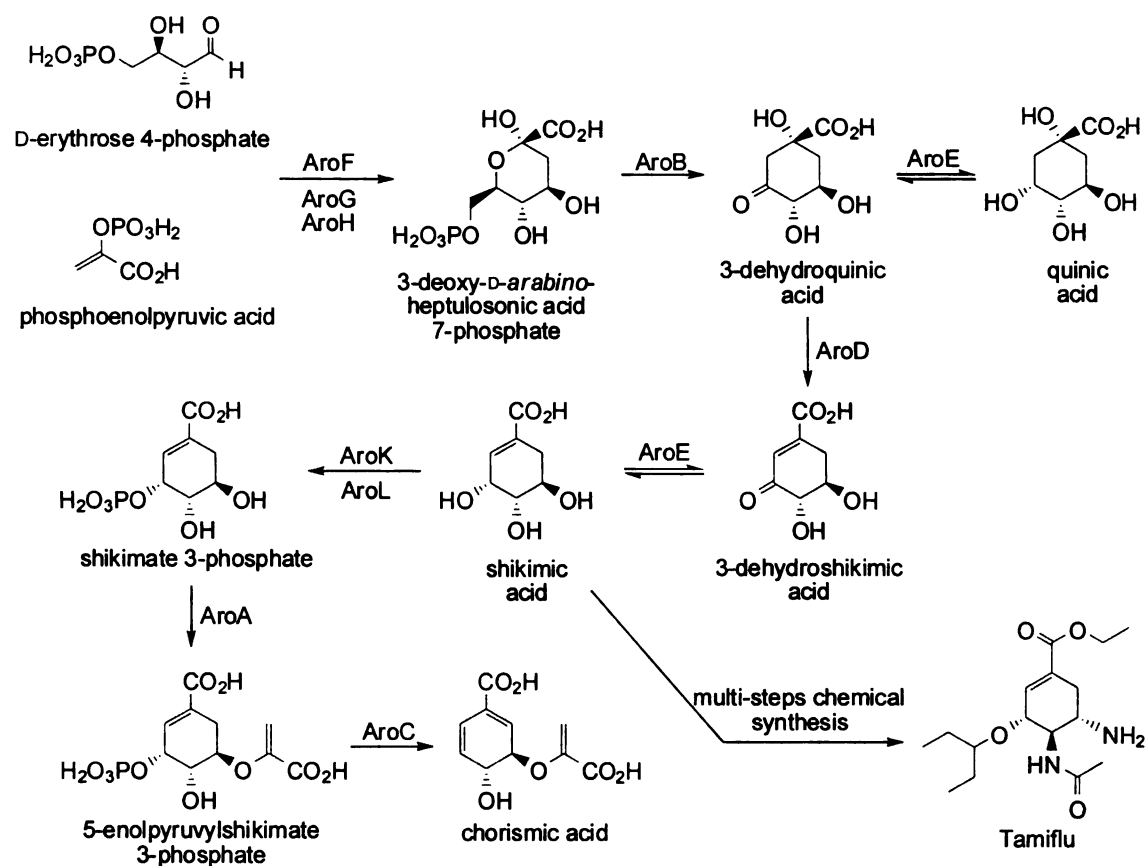


Figure 1. The shikimate pathway and biosynthesis of quinic acid. Enzymes: DAHP synthase (AroF, AroG, AroH); 3-dehydroquinase (AroB); 3-dehydroquinase dehydratase (AroD); shikimate dehydrogenase (AroE); shikimate kinase (AroK, AroL); 5-enolpyruvylshikimate 3-phosphate synthase (AroA); chorismate synthase (AroC).

Quinic acid is another shikimate pathway metabolite that is a useful natural product. Quinic acid has been identified to be biologically active and can be isolated from the bark of *Uncaria tomentosa*, which is better known as cat's claw.¹⁵ Cat's claw extract is widely used in South America as a historic medicinal treatment for gastric, arthritis, cancer and inflammatory conditions. Microbial synthesis of quinic acid using *E. coli* has been accomplished by exploiting the catalytic promiscuity of shikimate dehydrogenase, which in addition to catalyzing the reduction of 3-dehydroshikimate, can also catalyze the reduction of 3-dehydroquinic acid to quinic acid. High-titer, high-yielding microbial syntheses of quinic acid have subsequently been established based on overexpression of shikimate dehydrogenase in 3-dehydroquinic acid-synthesizing *E. coli* strains and have led to concentration of 54 g/L of quinic acid synthesized in 20% (mol/mol) yield from glucose.¹⁶ Quinic acid is also a key intermediate in the synthesis of hydroquinone from glucose.¹⁶ As a pseudocommodity chemical mainly used for photographic developing, the approximate 4.5×10^7 kg/yr production of hydroquinone is derived from aniline, phenol, and *p*-diisopropylbenzene, which are all manufactured from benzene.¹⁷ Quinic acid in partially purified fermentation broth was readily converted into hydroquinone employing HOCl or Ag₃PO₄/K₂S₂O₈ (Figure 2). More recently, quinic acid has been identified to have intriguing activities associated with stimulation of the immune system. In rats treated with the anticancer agent doxorubicin, cat's claw extract has been observed to accelerate the recovery of white blood cells.¹⁵

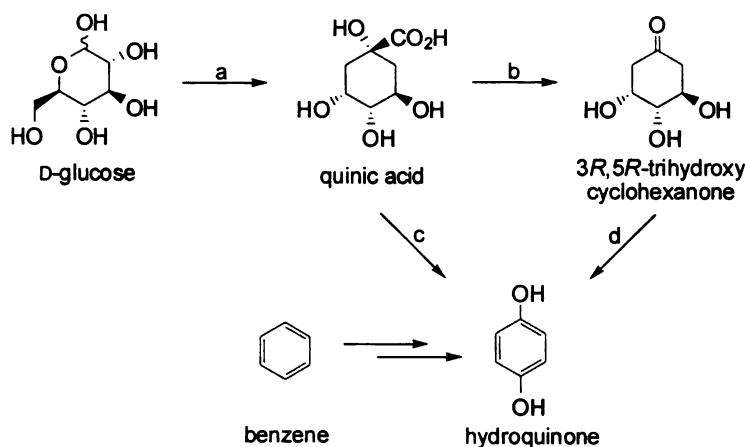


Figure 2. Synthesis of hydroquinone from glucose. (a) microbial synthesis; (b) $\text{Ag}_3\text{PO}_4/\text{K}_2\text{S}_2\text{O}_8$; (c) HOCl ; (d) heat.

Indigo dye

Another success of pathway engineering is found in the microbial synthesis of indigo dye. An estimated world market of 16,000 tons per year for indigo is dominated by its use to dye denim. In the chemical process, phenylglycine reacts with KOH-NaOH at $900\text{ }^\circ\text{C}$ with NaNH_2 to produce indoxyl, which subsequently oxidizes in air to form indigo (Figure 3).¹⁸ The microbial synthesis of indigo from D-glucose was developed by Genencor using metabolic engineering to incorporate a single *P. putida* gene encoding naphthalene dioxygenase (NDO) into *E. coli*. This heterologously expressed protein modified the tryptophan pathway in *E. coli* to cause high-level production of indole. Indole was converted by NDO into indoxyl, which spontaneously oxidized to form indigo (Figure 3).¹⁸ As another example of microbial synthesis of chemicals, genetically engineered *E. coli* has been successfully used to synthesized 1,3-propanediol that is not naturally produced in *E. coli*.¹⁹

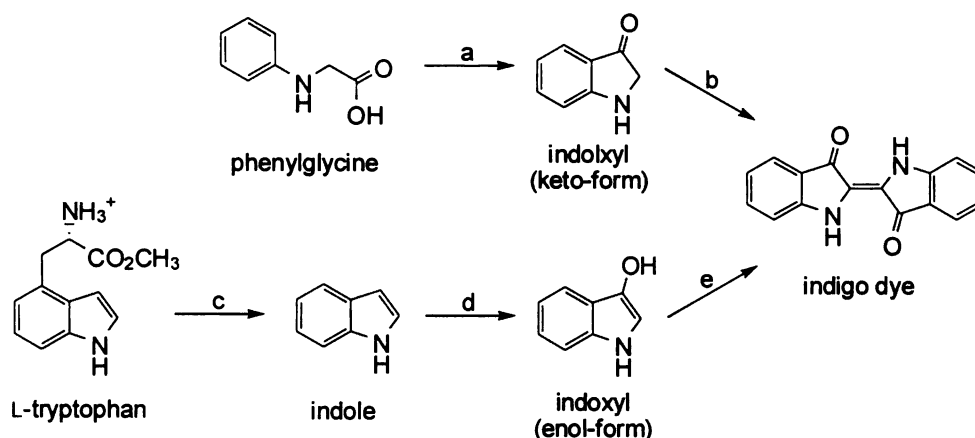


Figure 3. Chemical and microbial syntheses of indigo. (a) NaOH/KOH/NaNH₂, 900 °C; (b) spontaneous oxidation; (c) *E. coli* native tryptophanase; (d) *P. putida* naphthalene dioxygenase (NDO); (e) spontaneous oxidation.

1,3-Propanediol

1,3-propanediol is an important monomer in the production of poly(propylene terephthalate) (PPT) polymer and fibers,²⁰ which has long been known to have more preferable properties over poly(ethylene terephthalate) (PET), and poly(butylenes terephthalate) (PBT).²¹ Ethylene glycol and 1,4-butanediol, which are used to make PET and PBT, however, are much more expensive than 1,3-propanediol, which is the intermediate for PPT. Traditionally, 1,3-propanediol is produced either from the reductive hydration of acrolein (Degussa-DuPont process), or through reductive carbonylation of ethylene oxide (Shell process).

In the late 1990s, microbial synthesis of 1,3-propanediol jointly developed by DuPont and Genencor using renewable carbohydrates as the starting material came onto the scene.¹⁹ The process originally patented by DuPont in 1996 consisted of a two-stage fermentation, the first in *S. cerevisiae* leading to the synthesis of glycerol from glucose, and the second in *K. pneumoniae* to convert the glycerol into 1,3-propanediol.²² The yeast

produces glycerol from the glycolytic intermediate dihydroxyacetone-3-phosphate using glycerol-3-phosphate dehydrogenase (GPD1) and glycerol-3-phosphate phosphatase (GPP2). The natural pathway for the production of 1,3-propanediol from glycerol requires two enzymes, glycerol dehydratase (DhaB) and 1,3-propanediol dehydrogenase (DhaT). A single microbe synthesis of 1,3-propanediol was reported in 1998 based on the expression of GPD1, GPP2, DhaB and DhaT in a single *E. coli* microbe. *E. coli* cannot convert dihydroxyacetone-3-phosphate into glycerol or synthesize 1,3-propanediol from glycerol (Figure 4). *E. coli* provides several advantages over other organisms. It has been an intensively studied organism, an extensive range of methodologies is available for its genetic manipulation, and it has been used for some time for large scale production under fermentor-controlled conditions. In addition, since *E. coli* does not naturally produce glycerol or 1,3-propanediol, there is no natural regulation to overcome. Currently, microbial synthesis of 1,3-propanediol from glucose using a single microbe achieves a titer of 129 g/L in 34% yield (g of 1,3-propanediol/g of glucose consumed).²³

The significance of 1,2,4-butanetriol and 1,2,4-butanetriol trinitrate

Mixed acid nitration of racemic 1,2,4-butanetriol produces racemic 1,2,4-butanetriol trinitrate, which is used as an energetic plasticizer in propellant and explosive formulations (Figure 5). Although 1,2,4-butanetriol trinitrate and nitroglycerin are used in similar applications, 1,2,4-butanetriol trinitrate is less volatile, more thermally stable, and has a lower shock sensitivity than nitroglycerin. Substitution of 1,2,4-butanetriol trinitrate in plasticizer applications would therefore reduce hazards associated both with the propellant manufacturing process and use of the final product.

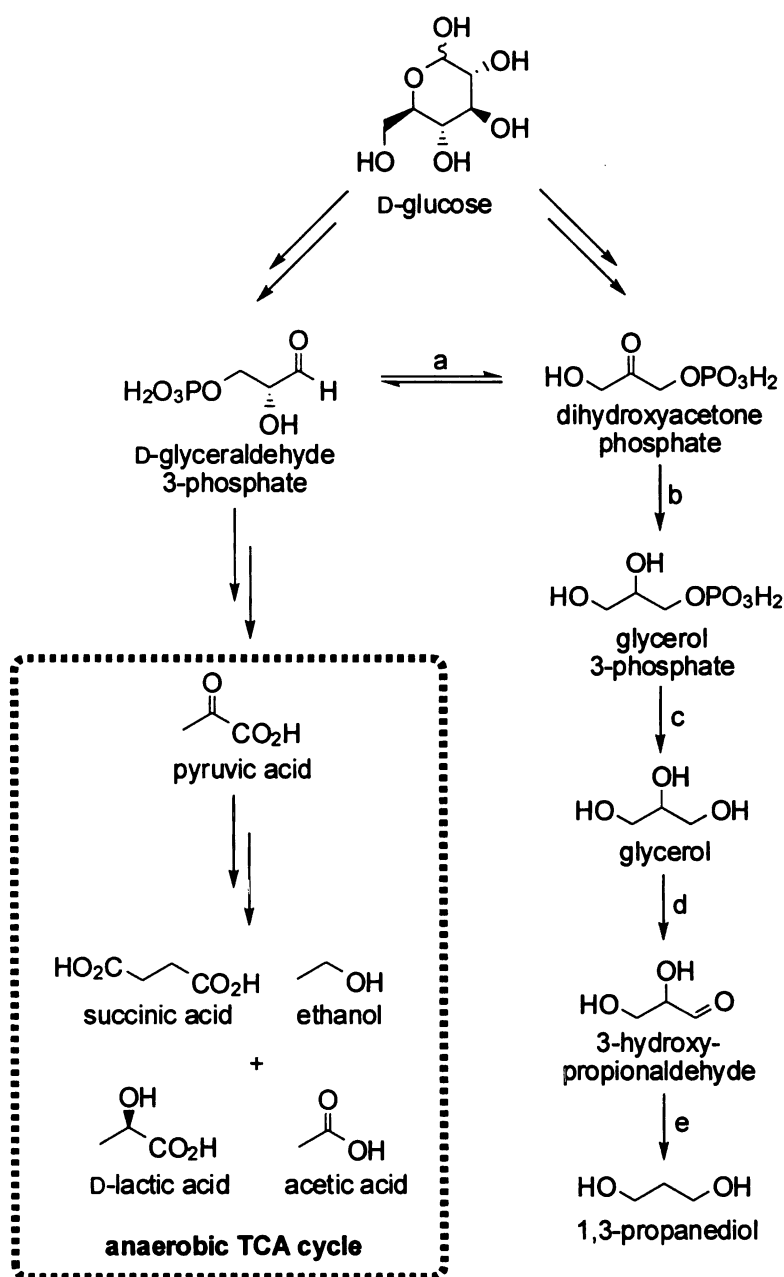


Figure 4. Schematic of the reaction network from D-glucose to 1,3-propanediol. (a) triose phosphate isomerase (TIM); (b) *S. cerevisiae* glycerol dehydrogenase (GPD1); (c) *S. cerevisiae* glycerol 3-phosphate phosphatase (GPP2); (d) *K. pneumoniae* glycerol dehydratase (DhaB); (e) *K. pneumoniae* 1,3-propanediol dehydrogenase (DhaT).



Beyond its use as an energetic material, 1,2,4-butanetriol could have significant impact in other aspects. Polyurethane foams made with 1,2,4-butanetriol have intriguing elastic properties. These foams have the same compression-bending characteristics as natural rubber.²⁵ D-1,2,4-Butanetriol trinitrate provides intriguing opportunities as an alternative to nitroglycerin for use as a vasodilator in the treatment of angina.²⁶ Beyond its advantages in manufacture and formulation over nitroglycerin, 1,2,4-butanetriol trinitrate is degraded more slowly by organic nitrate reductase than nitroglycerin.²⁶ This has the potential to translate into longer lasting plasma levels of D- and L-1,2,4-butanetriol trinitrate relative to nitroglycerin. Using only a single enantiomer of 1,2,4-butanetriol trinitrate could conceivably minimize the number of metabolites generated from the action of organic nitrate reductase and correspondingly minimize the chances of encountering adverse side effects.

Chemical syntheses of 1,2,4-butanetriol

1,2,4-Butanetriol has been commercially manufactured by the reduction of dimethyl malate using sodium borohydride in a mixture of a C₂₋₆-alcohols and tetrahydrofuran (Figure 6).²⁷ Such a stoichiometric reduction is expensive. In addition, each ton of 1,2,4-butanetriol produced results in the generation of 2-5 tons of borate salts, which constitute an environmental impediment to large scale manufacture.²⁷

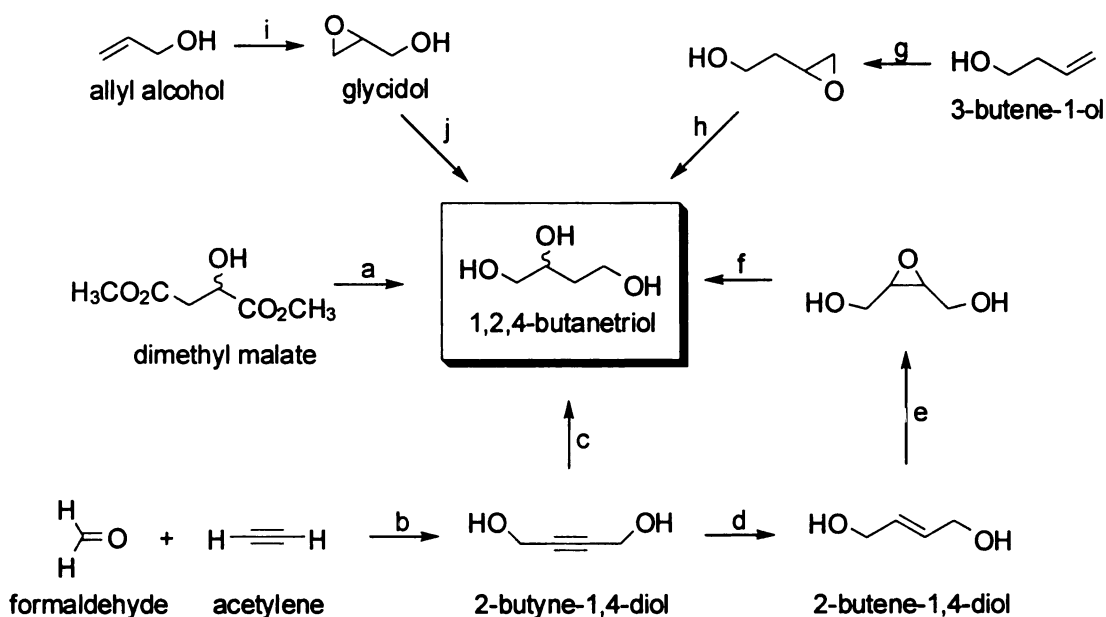


Figure 6. Syntheses of 1,2,4-butanetriol. (a) NaBH₄, THF, MeOH; (b) CuC₂, 50 atm; (c) i) HgSO₄, H₂SO₄, ii) 2CuO·Cr₂O₇, H₂; (d) Lindler's catalyst, H₂; (e) H₂O₂, H₂WO₄; (f) Pd/C, H₂; (g) H₂O₂, H₂WO₄; (h) H₂SO₄; (i) H₂O₂; (j) i) Co₂(CO)₈, (CO, H₂), ii) LiAlH₄.

A variety of other alternative routes to 1,2,4-butanetriol have also been explored (Figure 6). Reaction of acetylene and formaldehyde catalyzed by CuC₂ to form 2-butyne-1,4-diol provides two avenues to 1,2,4-butanetriol. Oxymercuration of 2-butyne-1,4-diol and subsequent reductive demercuration gives 1,4-dihydroxy-2-butanone, which is hydrogenated to afford 1,2,4-butanetriol. Alternatively, 2-butyne-1,4-diol can be converted

to 2-butene-1,4-diol. Oxidation to the corresponding epoxide followed by hydrogenation yields 1,3,4-butanetriol. A low-yielding route to 1,2,4-butanetriol consists of the reaction of glycidol with CO and H₂ followed by reduction of the 3-hydroxy- γ -butyrolactone intermediate. 3-Buten-1-ol, which is derived from acetaldehyde, can be oxidized to an epoxide followed by hydrolysis to obtain 1,2,4-butanetriol. The most direct chemical route to 1,2,4-butanetriol entails the low-yielding reaction of allyl alcohol with formaldehyde.

In the Frost lab, metal-mediated catalytic hydrogenation of malate was examined as an alternative to sodium borohydride and the salt streams attendant with this stoichiometric reduction. Ruthenium catalyzed hydrogenation of malic acid in water afforded 1,2,4-butanetriol in 74% yield. However, elevated H₂ pressure (5,000 psi) and elevated reaction temperature (135 °C) were required.²⁸ Competing deoxygenation and cracking reactions generated byproducts that were difficult to separate from 1,2,4-butanetriol.²⁸ A chemo-enzymatic synthesis of 1,2,4-butanetriol using L-ascorbic acid as starting material was also investigated (Figure 7). Oxidation of L-ascorbic acid using H₂O₂ gave L-threonate, which was dehydrated using dihydroxy-acid dehydratase. Cyclization of the resulting 4-hydroxy-2-ketobutyric acid during extraction from water afforded 2-hydroxy-2-buten-4-olide. High-pressure, high temperature catalytic hydrogenation of 2-hydroxy-2-buten-4-olide gave 1,2,4-butanetriol in 53% overall yield.²⁹ Given the shortcomings of the chemical and chemo-enzymatic routes, a microbial synthesis of 1,2,4-butanetriol as pursued.

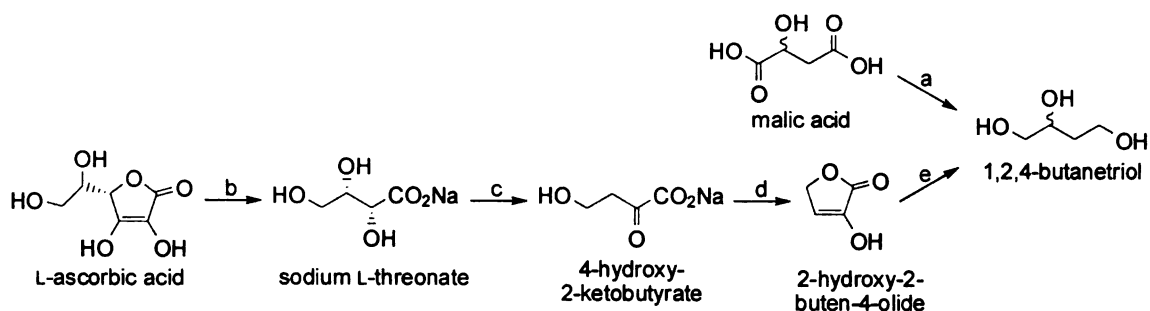


Figure 7. Routes to 1,2,4-butanetriol previously developed in the Frost lab. (a) 5% Ru on C at 1.3% Ru/malate, 340 atm H₂, 135 °C, 10 h, 74% yield; (b) Na₂CO₃, 30% H₂O₂; (c) *E. coli* JWF1/pON1.118B cell extract, pH 8.2, 37 °C, N₂; (d) HCl, pH 1.5; (e) 1 mol% Ru on C, 170 atm H₂, 125 °C, 53% overall yield from L-ascorbic acid.

Creation of a 1,2,4-butanetriol biosynthetic pathway

Microbial synthesis of 1,2,4-butanetriol poses a unique challenge, largely due to the fact that 1,2,4-butanetriol is not a natural product. Since no biosynthetic routes apparently exist in nature for D-1,2,4-butanetriol or L-1,2,4-butanetriol, a biosynthetic pathway had to be created *de novo*. The first route developed in the Frost group was a two-microbe synthesis of D-1,2,4-butanetriol from D-xylose (Figure 8). A wild-type microbe *Pseudomonas fragi* ATCC4973 was used to oxidize D-xylose, which afforded 77 g/L concentrations of D-xylonic acid in 70% yield.²⁸ An *E. coli* construct then catalyzed the conversion of D-xylonic acid into D-1,2,4-butanetriol. The Frost lab discovered that native *E. coli* expresses D-xylonate dehydratase (YjhG and YagF) and is able to use D-xylonic acid as a sole source of carbon to support growth. Therefore, transport of D-xylonic acid and its conversion to 3-deoxy-D-glycero-pentulosonic acid (Figure 8) relied solely on the native *E. coli* activity. Conversion of D-xylonic acid to 1.6 g/L of D-1,2,4-butanetriol in 25% yield required only the expression of plasmid-localized *Pseudomonas putida*

ATCC12633 *mdlC*-encoded benzoylformate decarboxylase in *E. coli* DH5 α /pWN6.168A.²⁸

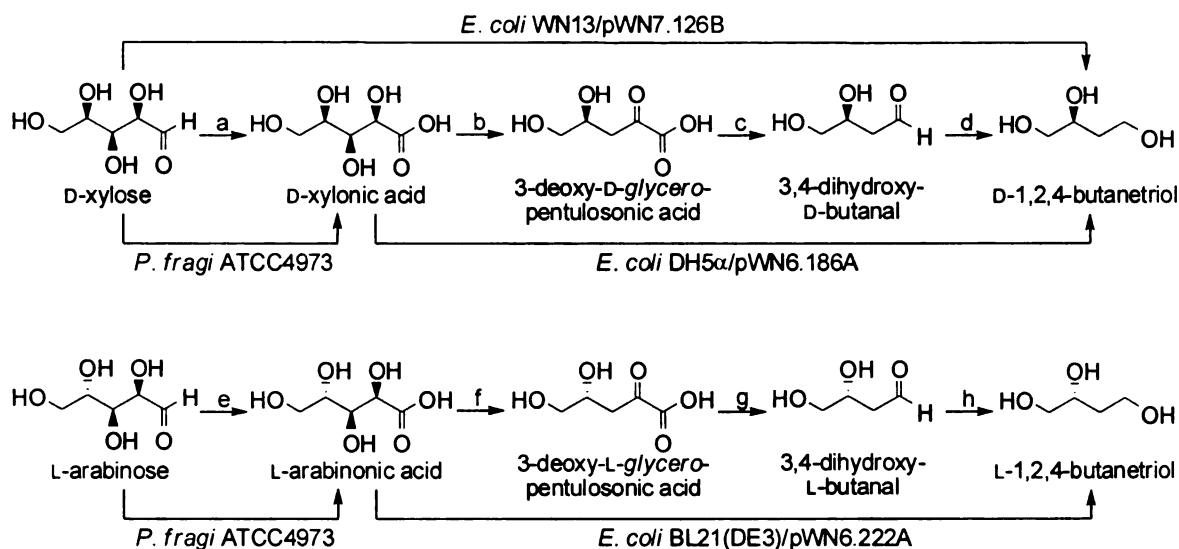


Figure 8. Microbial syntheses of D- and L-1,2,4-butanetriol. (a) D-xylonate dehydrogenase; (b) D-xylonate dehydratase; (c) benzoylformate decarboxylase; (d) alcohol dehydrogenase; (e) L-arabinonate dehydrogenase; (f) L-arabinonate dehydratase; (g) benzoylformate decarboxylase; (h) alcohol dehydrogenase.

A parallel biosynthetic pathway was also created for L-1,2,4-butanetriol. *P. fragi* ATCC4973 was again recruited for the oxidation of L-arabinose, which afforded 55 g/L L-arabinonic acid in 54% yield. The construction of the second *E. coli* catalyst, however, is more complicated than the former case since *E. coli* is unable to catabolize D-arabinonic acid. To circumvent this issue, a genomic library of *P. fragi* ATCC4973 was constructed in *E. coli* followed by selection for *E. coli* strains that acquired the ability to use L-arabinonic acid as a sole carbon source for growth and metabolism. This led to isolation of the gene designated *aadh* encoding L-arabinonate dehydratase. The *aatp* gene encoding arabinonate transport protein was also discovered to be contiguous to the *aadh* gene.

Conversion of L-arabinonic acid to 2.4 g/L of L-1,2,4-butanetriol in 19% yield was accomplished using *E. coli* BL21(DE3)/pWN6.222A, which expressed the plasmid localized *aatp*-encoded L-arabinonate transport protein, *aadh*-encoded L-arabinonate dehydratase, and *mdlC*-encoded benzoylformate decarboxylase (Figure 8).²⁸

A single microbe for D-1,2,4-butanetriol synthesis

Synthesis of D-1,2,4-butanetriol from D-xylose using a single microbe was then pursued. The major challenges to the direct, microbe-catalyzed conversion of D-xylose into D-1,2,4-butanetriol included the lack of a gene encoding D-xylonate dehydrogenase and competing catabolism of intermediate 3-deoxy-D-*glycero*-pentulosonic acid.

A partial amino acid sequence of D-xylonic acid dehydratase was isolated from *P. fragi* ATCC4973 by protein purification, by trypsin digestion, and N-terminal sequence analysis of the peptide fragments. A *Burkholderia fungorum* LB400 protein was identified in the ERGO bacterial genome database to encode D-xylonate dehydratase activity using this information.¹⁴ Because genes encoding catabolic enzymes for the same carbon source tend to form gene clusters in chromosomes, open reading frames adjacent to this *B. fungorum* D-xylonate dehydratase were examined leading to identification of the D-xylonate dehydrogenase in *B. fungorum*. A similar approach also led to the discovery of another D-xylonate dehydrogenase in *Caulobacter crescentus* CB15.¹⁴

E. coli has a different D-xylonic acid catabolism than *P. fragi* and elucidation of this pathway turned out to be critical for improving the yields and concentrations of biosynthesized D-1,2,4-butanetriol. Two sets of catabolic enzymes encoded by the *yjh* and *yag* gene clusters were discovered to be responsible for *E. coli* utilization of D-xylonic

acid as a sole source of carbon for growth (Figure 9).¹⁴ YagE and YjhH were further identified to catalyze the cleavage of 3-deoxy-*glycero*-D-pentulosonic acid to form pyruvic acid and glycolaldehyde.¹⁴

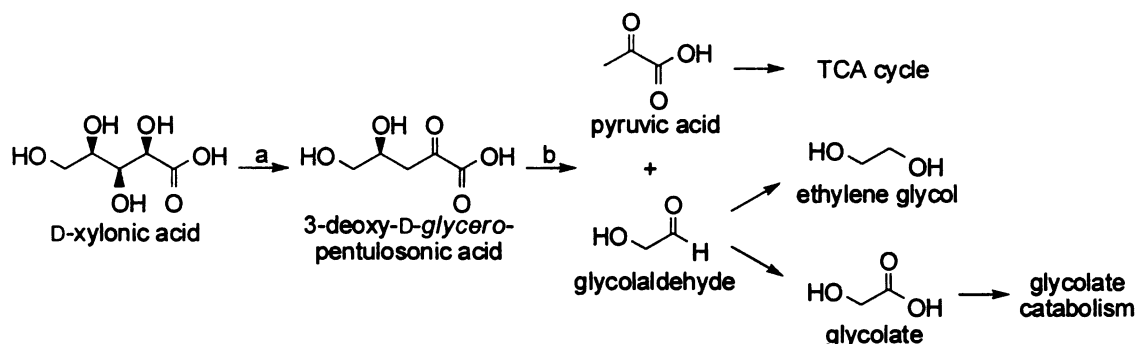


Figure 9. Native *E. coli* D-xylonic acid catabolic pathway. (a) *E. coli* D-xylonate dehydratase (YjhG, YagF); (b) 3-deoxy-D-*glycero*-pentulosonate aldolase (YagE, YjhH)

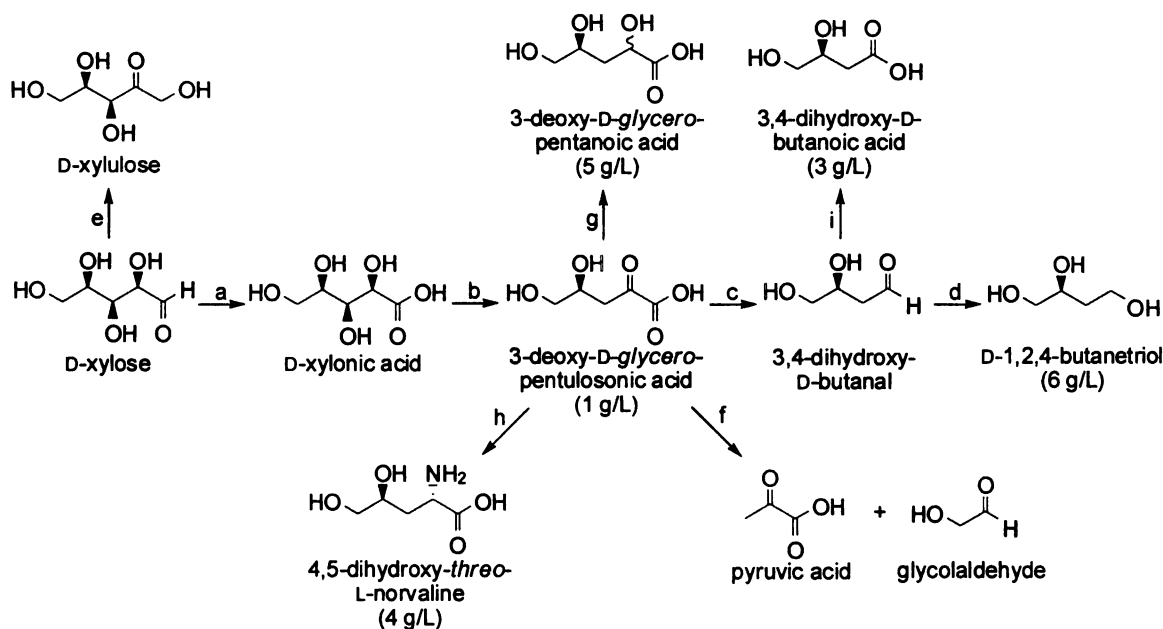


Figure 10. Microbial synthesis of D-1,2,4-butanetriol using WN13/pWN7.126B. (a) *C. crescentus* D-xylulose dehydrogenase (*xdh*); (b) *E. coli* D-xylulose dehydratase (*yjhG* and *yagF*); (c) *P. putida* benzoylformate decarboxylase (*mdlC*); (d) *E. coli* alcohol dehydrogenase(s); (e) inactivated *E. coli* D-xylulose isomerase (*xylA*); (f) inactivated *E. coli* 3-deoxy-D-*glycero*-pentulosonate aldolase (*yagE* and *yjhH*); (g) *E. coli* dehydrogenase(s); (h) *E. coli* transaminase; (i) *E. coli* dehydrogenase(s).

A single microbe, *E. coli* WN13/pWN7.126B, was subsequently constructed to express all four enzyme activities (Figure 10) required for the synthesis of D-1,2,4-butanetriol from D-xylose. The D-xylose dehydrogenase gene (*xdh*) from *C. crescentus* was integrated to the *xylAB* site in the *E. coli* WN13 chromosome to convert D-xylose into D-xylonic acid. Native *E. coli* D-xylonate dehydratases (YagF and YjhG) utilize D-xylonic acid to produce 3-deoxy-D-*glycero*-pentulosonic acid. Knockouts of the 2-keto acid aldolase genes (*yagE* and *yjhH*) eliminated catabolic competition for 3-deoxy-D-*glycero*-pentulosonate. Plasmid pWN7.126B contains the benzoylformate decarboxylase gene (*mdlC*) from *P. putida*. Expression of this gene from a *tac* promoter produced the enzyme activity for conversion of 3-deoxy-D-*glycero*-pentulosonic acid into 3,4-dihydroxy-D-butanal. As the final step in the biosynthesis, the conversion of 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol relied solely on unidentified native *E. coli* dehydrogenase activity. This *E. coli* microbe synthesized 6 g/L D-1,2,4-butanetriol in 28% yield (mol/mol) under fermentor-controlled conditions together with 3-deoxy-D-*glycero*-pentulosonic acid, 3-deoxy-D-*glycero*-pentanoic acid, 4,5-dihydroxy-*threo*-L-norvaline and 3,4-dihydroxy-D-butanoic acid as byproducts (Figure 10).

Important chemicals derived from the microbial 1,2,4-butanetriol synthesis

3,4-Dihydroxy-D-butanoic acid

3,4-dihydroxy-D-butanoic acid lactonizes to form (*S*)-3-hydroxy- γ -butyrolactone (Figure 11), which is an important synthetic precursor. This compound is used in the commercial manufacture of Astra Zeneca's cholesterol-lowering drug Crestor® (Figure

11). One of the two patented routes to the cholesterol-lowering drug Zetia® (Figure 11) also employs (*S*)-3-hydroxy- γ -butyrolactone as a chiral synthon. Merck and Schering-Plough jointly market Zetia®, which has a unique mode of action for lowering cholesterol.

3,4-Dihydroxy-*threo*-L-norvaline

3,4-dihydroxy-*threo*-L-norevaline has been used as a precursor for the synthesis of a clavam antibiotic, clavalanine.³⁰ This antibiotic was isolated in 1983 from *Streptomyces clavuligerus*,³¹ which is known to be a prolific β -lactam antibiotic producer. Clavalanine is unique for being an antimetabolite of *O*-succinylhomoserine which results in disruption of methionine biosynthesis.^{31c} This stands in marked contrast to disruption of peptidoglycan biosynthesis, which characterizes the mode of action of other β -lactam antibiotics.

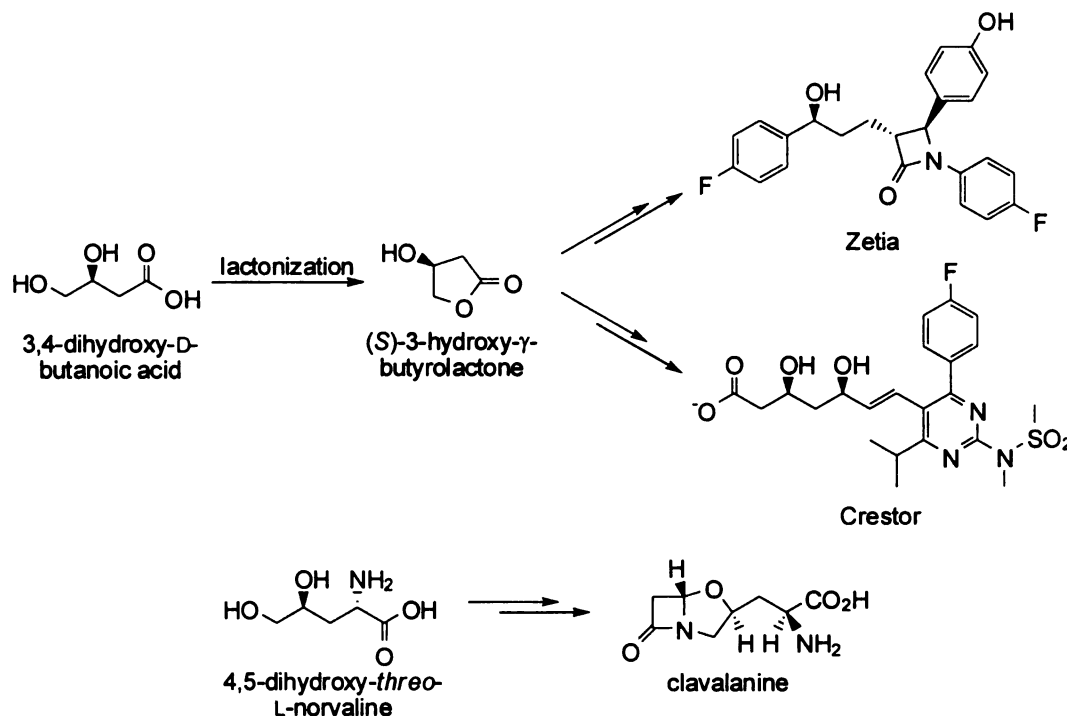


Figure 11. Drugs synthesized from 3,4-dihydroxy-D-butanoic acid and 4,5-dihydroxy-*threo*-L-norvaline.

REFERENCE

- ¹ Pannuri, S.; DiSanto, R.; Kamat, S. In *Kirk-Othmer Encyclopedia of Chemical Technology Online*; Biocatalysis, 2003, Wiley.
- ² Junker, B. In *Kirk-Othmer Encyclopedia of Chemical Technology Online*; Fermentation, 2004, Wiley.
- ³ Adrio, J. L.; Demain, A. L. In *Microbial Enzymes and Biotransformations*; Barrede, J. L. Ed.; Humana Press: Totowa, NJ, 2005; p1.
- ⁴ Cohen, S.; Chang, A. C. Y.; Boyer, H. W.; Helling, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 3240.
- ⁵ Saiki, R. K.; Scharf, S.; Faloona, F.; Mullis, K. B.; Horn, G. T.; Erlich, H. A.; Arnheim, N. *Science* **1985**, *4732*, 1350.
- ⁶ Vasic-Racki, D. In *Industrial Biotransformation*; Liese, A.; Seelbach, K.; Wandrey, C. Eds.; Wiley-VCH: Weinheim, 2006; p1.
- ⁷ Bornscheuer, U. T.; Buchholz, K. *Eng. Life Sci.* **2005**, *5*, 309.
- ⁸ Buchholz, K.; Poulsen, P. B. In *Applied Biocatalysis*; Straathof, A. J. J.; Adlercreutz, P. Eds.; Harwood Academic Publisher: Amsterdam, 2000.
- ⁹ Robertson, D. E.; Chaplin, J. A.; DeSantis, G.; Podar, M.; Madden, M.; Chi, E.; Richardson, T.; Milan, A.; Miller, M.; Weiner, D. P.; Wong, K.; McQuaid, J.; Farwell, B.; Preston, L. A.; Tan, X.; Snead, M. A.; Keller, M.; Mathur, E.; Kretz, P. L.; Burk, M. J.; Short, J. M. *Appl. Environ. Microbiol.* **2004**, *70*, 2429.
- ¹⁰ (a) Arnold, F. H.; Georgiou, G. In *Methods in Molecular Biology*, Vol. 230; Humana Press: Totowa, NJ, 2003. (b) Arnold, F. H.; Georgiou, G. In *Methods in Molecular Biology*, Vol. 231; Humana Press: Totowa, NJ, 2003. (c) Brakmann, S.; Johnsson, K. *Directed Molecular Evolution of Enzymes*; Wiley-VCH: Weinheim, 2002. (d) Brakmann, S.; Schwienhorst, A. *Evolutionary Methods in Biotechnology*; Wiley-VCH: Weinheim, 2004.
- ¹¹ Ghisalba, O. In *New Trends in Synthetic Medicinal Chemistry*; Gualtieri, F. Ed.; Wiley-VCH: Weinheim, 2000; p175.
- ¹² (a) Draths, K. M.; Knop, D. R.; Frost, J. W. *J. Am. Chem. Soc.* **1999**, *121*, 1603. (b) Frost, J. W.; Frost, K. M.; Knop, D. R. WO 2000044923 A1 20000803. (c) Knop, D. R.;

Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 10173. (d) Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. *Biotech. Prog.* **2003**, *19*, 808.

¹³ (a) Barbas, C. F.; Wang, Y.-F.; Wong, C.-H. *J. Am. Chem. Soc.* **1990**, *112*, 2013. (b) Chen, L.; Dumas, D. P.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 741.

¹⁴ Niu, W.; Frost, J. W. unpublished results.

¹⁵ Sheng, Y.; Akesson, C.; Holmgren, K.; Bryngelsson, C.; Giamapa, V.; Pero, R. W. *J. Ethnopharmacol.* **2005**, *96*, 577.

¹⁶ Ran, N.; Knop, D. R.; Draths, K. M.; Frost, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 10927.

¹⁷ Krumenacker, L.; Constantini, M.; Potal, P.; Sentenac, J. In *Kirk-Othmer Encyclopedia of Chemical Technology Online*; Hydroquinone, Resorcinol, and Catechol, 1995, Wiley.

¹⁸ (a) Ferreira, E. S. B.; Hulme, A. N.; McNab, H.; Quye, A. *Chem. Soc. Rev.* **2004**, *33*, 329. (b) Bommarius, A. S.; Riebel, B. R. *Biocatalysis*; Wiley-VCH: Weinheim, 2004.

¹⁹ Gatenby, A. A.; Haynie, S. L.; Nagarajan WO 9821339, 1998.

²⁰ Polk, M. B.; Vigo, T. L.; Turbak, A. F. In *Kirk-Othmer Encyclopedia of Chemical Technology Online*; High Performance Fibers, 2004, Wiley.

²¹ Hansen, S.; Atwood, K. B. In *Kirk-Othmer Encyclopedia of Chemical Technology Online*; Polyester Fibers, 2005, Wiley.

²² Haynie, S. L.; Wagner, L. W. WO 35799, 1996.

²³ Emptage, M.; Haylie, S. L.; Laffend, L. A.; Pucci, J. P.; Whited, G. M. US 7,067,300, 2006.

²⁴ (a) Charles Cappuccino, Operations Manager, Copperhead Chemical Company Inc. (b) Charles Painter, ManTec Center, Naval Surface Warfare Center, Indian Head Division.

²⁵ Gmitter, G. T. US 3,149,083, 1964. (b) Wilson, C. L.; Riedeman, W. L. DE 1047420, 1958.

²⁶ Needleman, P. *Annu. Rev. Pharmacol. Toxicol.* **1976**, *16*, 81.

- ²⁷ (a) Monteith, M. J.; Schofield, D.; Bailey, M. WO 98/08793, 1998. (b) ikai, K.; Amagasaki-shi, H. EP 1,061,060 A1, 1999.
- ²⁸ (a) Niu, W.; Molefe, M.; Frost, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 12998. (b) Frost, J. W. WO 2005/068642, 2005.
- ²⁹ Molefe, M. *PhD Thesis*, Michigan State University, 2005.
- ³⁰ Bernardo, S. D.; Tengi, J. P; Sasso, G. J.; Weigele, M. *J. Org. Chem.* **1985**, *50*, 3457.
- ³¹ (a) Higgins, C. E.; Kastner, R. E. *Int. J. Syst. Bacteriol.* **1971**, *21*, 326. (b) Evans, R. H.; Ax, H.; Jacoby, A.; Williams, T. H.; Jenkins, E.; Scannell, J. P. *J. Antibiotics* **1983**, *36*, 213. (c) Pruess, D. L.; Kellet, M. *J. Antibiotics* **1983**, *36*, 208. (d) Muller, J.-C.; Toome, V.; Pruess, D. L.; Blount, J. F.; Weigele, M. *J. Antibiotics* **1983**, *36*, 217.

CHAPTER TWO

Chemo-Enzymatic Synthesis of D-1,2,4-Butanetriol Using 2-Deoxyribose-5-Phosphate Aldolase

Background

Aldolases are emerging into effective and economically viable tools for the industrial syntheses of chiral molecules. They catalyze enantioselective carbon-carbon bond formations to generate up to two chiral centers under mild reaction conditions. Among the aldolases proven to be useful for large-scale synthesis is the *deoC* gene encoding 2-deoxyribose-5-phosphate aldolase (DERA, EC 4.1.2.4).¹ In *E. coli*, this aldolase is associated with thymidine phosphorylase (DeoA), purine nucleoside phosphorylase (DeoD), and phosphopentomutase (DeoB) in the *deo* operon.² 2-Deoxy-D-ribose-5-phosphate aldolase is a 28 kDa enzyme that reversibly catalyzes the retro-aldol reaction of 2-deoxy-D-ribose-5-phosphate, a degradation product of deoxyribonucleic acid to form acetaldehyde and D-glyceraldehyde-3-phosphate (Figure 12).³ Interestingly, the chemical equilibrium lies towards 2-deoxy-D-ribose-5-phosphate, with an equilibrium constant of $4.2 \times 10^3 \text{ M}^{-1}$.⁴

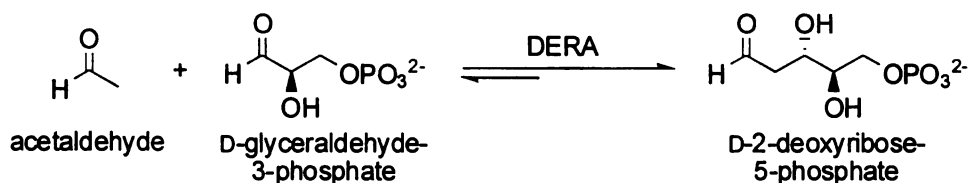


Figure 12. *In vivo* reaction catalyzed by 2-deoxyribose-5-phosphate aldolase.

Mechanistically, 2-deoxyribose-5-phosphate aldolase belongs to the type I class of aldolases with the enzymatic reaction proceeding through a Schiff base intermediate between the donor substrate and an active site lysine residue.⁵ It is the only aldolase known to date which accepts two aldehydes in the condensation reaction. Other aldolases require a ketone and an aldehyde. Several literature reports show that 2-deoxyribose-5-phosphate aldolase has high substrate specificity towards using acetaldehyde as the donor substrate.⁶ Its application in organic synthesis has been emerging since the first report of an unprecedented one-pot tandem aldol reaction to synthesize six-membered lactol derivatives, in which two equivalents of acetaldehyde were added in sequence to a variety of two-carbon aldehyde acceptors (Figure 13).⁷ Since 2-deoxyribose-5-phosphate aldolase catalyzes reactions in both directions, the intermediate 4-carbon adduct is reversibly formed under the reaction conditions. The second condensation took place between this adduct and a second equivalent of acetaldehyde to form a more stable lactol.

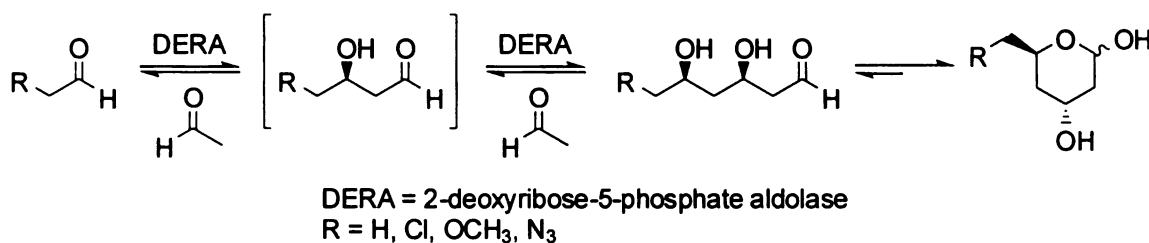


Figure 13. 2-Deoxyribose-5-phosphate aldolase catalyzed tandem aldol reaction.

The industrial interest in 2-deoxyribose-5-phosphate aldolase lies in the structural similarity of its enzymatic products to the lactone moiety of the cholesterol lowering 3-hydroxy-3-methylglutaryl (HMG)-CoA CoA reductase inhibitors, which are collectively known as statins. Unlike the earlier generation of statins that were mostly fungal

metabolites or semi-synthetic molecules, Pfizer's Lipitor (atorvastatin) (Figure 14) and AstraZeneca's Crestor (rosuvastatin) (Figure 11) are completely synthetically made. A (3*R*,5*S*)-dihydroxyhexanoate moiety that is commonly shared by these two drugs turns out to be essential for its activity. Chemical construction of the two chiral centers involved in this side chain proved to be a challenging task. For instance, the (3*R*,5*S*)-dihydroxyhexanoate in atorvastatin accounts for about 25% of the compound's total molecular weight. In addition, a synthesis of this intermediate with greater than 99.5% enantiomeric and 99% distereomeric excess is required to meet the standard for pharmaceutical use. To circumvent these issues associated with the traditional chemical synthesis, DSM⁸ and the enzyme developer Diversa⁹ have independently developed two improved 2-deoxyribose-5-phosphate aldolase variants for statin intermediate synthesis. These improved bioprocesses allows the intermediate (3*R*,5*S*)-6-chloro-2,4,6-trideoxyhexapyranoside to be synthesized in a large quantity and compete with other traditional routes to supply the 220 tons per year market demand worldwide.

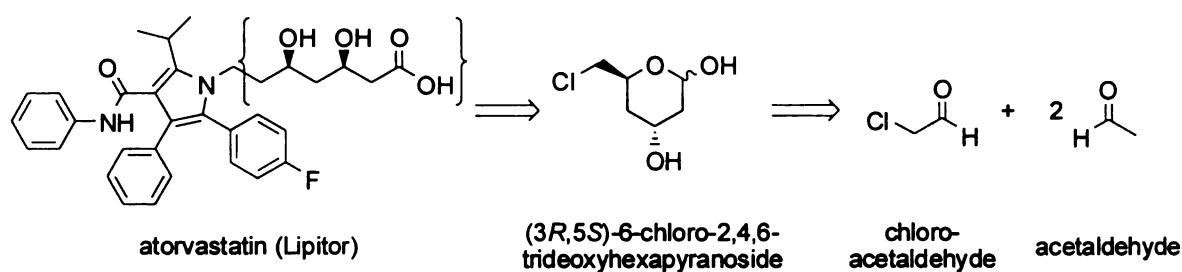


Figure 14. Retro-synthesis of atorvastatin (Lipitor).

The Frost group developed several approaches to the environmental benign synthesis of 1,2,4-butanetriol as discussed previously in Chapter 1. Among them is the direct catalytic hydrogenation of D- and L-malic acid. Hydrogenation of malic acid in aqueous solution using 5% ruthenium on carbon as catalyst was investigated in detail.^{10,11} Carboxylates with electron withdrawing groups attached to the adjacent carbon atom are known to be activated towards metal-mediated reduction.¹² However, the presence of the second unactivated carboxylate in malic acid is far less reactive and its hydrogenation requires elevated H₂ pressure and temperature. Under an H₂ pressure of 1,000 psi and temperature of 135 °C, malic acid was converted into 1,2,4-butanetriol in 25% yield using 5% ruthenium on carbon at 1.3% ruthenium to malic acid ratio. Exclusive reduction of the activated carboxylate on malic acid led to the formation of 3,4-dihydroxybutanoic acid and 3-hydroxybutyrolactone as the side product in a combined yield of 60%.¹¹ Further optimization of the hydrogenation process resulted in complete reduction of malic acid but also formed other side products. Under an elevated H₂ pressure of 5,000 psi and 135 °C, 1,2,4-butanetriol was synthesized in a maximum yield of 74% with other reaction parameters unchanged.¹¹ The elevated H₂ pressure and temperature led to the formation of various polyol byproducts due to carbon-carbon and carbon-oxygen bond cleavage reactions (Figure 15). Due to the structural similarity of 1,2,4-butanetriol and these byproducts, downstream purification becomes challenging. In addition, the expense of an industrial scale hydrogenation at 5,000 psi H₂ pressure was projected to be quite high. Although malic acid can be obtained using microbial synthesis with a renewable feedstock, its catalytic hydrogenation to synthesize 1,2,4-butanetriol is problematic.

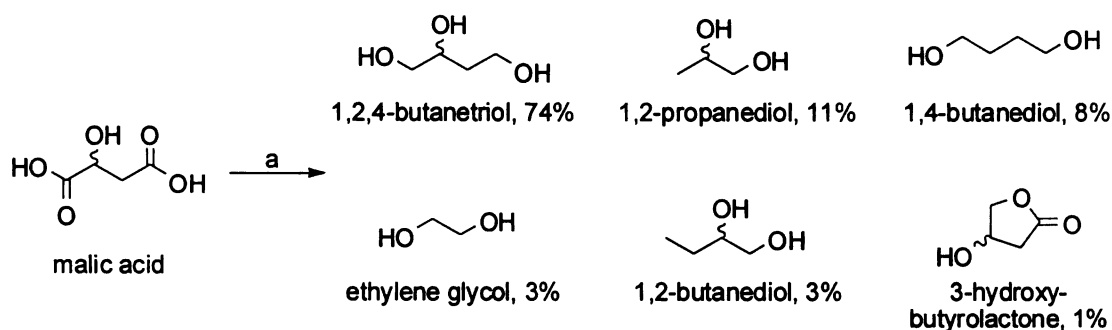


Figure 15. Catalytic hydrogenation of malic acid. (a) 5% Ru on C at 1.3% Ru/malate, 5,000 psi H₂, 135 °C, 10 h, 74% yield.

An alternate route to 1,2,4-butanetriol has also been examined in the Frost group. 2-Hydroxy-buten-4-olide (Figure 7) was targeted as the substrate for catalytic hydrogenation to afford 1,2,4-butanetriol based on its single carboxylate possessing an activating hydroxyl group attached to the adjacent carbon atom. Milder reaction conditions for the hydrogenation of butenolide relative to those required for the hydrogenation of malic acid were of interest to avoid the formation of polyol byproducts. 1,2,4-butanetriol was synthesized from 2-hydroxy-2-buten-4-olide under H₂ pressure of 2,500 psi at 125 °C using 1.0 mol% ruthenium catalyst (5% ruthenium on carbon) in a yield of 96% without observable byproduct formation.¹¹ Synthesis of 2-hydroxy-2-buten-4-olide started from L-ascorbic acid and proceeded through intermediacy of L-threonate and 4-hydroxy-2-ketobutyrate using a chemo-enzymatic route in a yield of 55% (Figure 7). Compared with the catalytic hydrogenation of malic acid, the hydrogenation of 2-hydroxy-2-buten-4-olide afforded a significant improvement in both catalytic hydrogenation reaction yield and product purity. However, the overall yield for the chemo-enzymatic process to 1,2,4-butanetriol from L-ascorbic acid was only 53% (Figure 7). Furthermore, replacing an industrially well-established one-step sodium borohydride reduction by this four-step process is not economically feasible.

Accordingly, a two-step chemo-enzymatic synthesis of D-1,2,4-butanetriol has been elaborated. It involves an enzymatic asymmetric aldol reaction followed by catalytic hydrogenation (Figure 16). The first step in our synthetic scheme is an enantioselective aldol condensation reaction using acetaldehyde and glycolaldehyde. Beyond the traditional demand of enantioselective control, the proposed reaction also requires high regioselectivity preference. The acetaldehyde molecule participates strictly as a donor substrate in the aldol reaction, while glycolaldehyde acts as receptor to form the corresponding 3,4-dihydroxy-D-butanal as the sole product and thus avoided the non-enzymatic mixed-aldol products formation. On the other hand, the reaction environment has to be controlled to avoid the retro-aldol reaction to take place. The success of developing this chemo-enzymatic process relied on a strategy using 2-deoxyribose-5-phosphate aldolase, the only enzyme that accepts two aldehydes (acetaldehyde and glycolaldehyde) as substrates to produce 3,4-dihydroxy-D-butanal. The resulting intermediate 3,4-dihydroxy-D-butanal was anticipated to undergo facile catalytic hydrogenation relative to the carboxylates employed in previous studies and yield D-1,2,4-butanetriol as the target product.

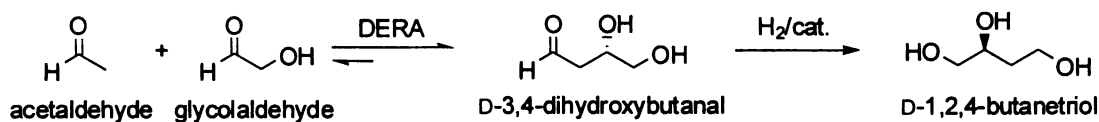


Figure 16. Proposed chemo-enzymatic reaction using 2-deoxyribose-5-phosphate aldolase.

Another important advantage of this proposed synthesis lies on its simple and inexpensive starting materials. Acetaldehyde and glycolaldehyde are readily available in large quantities. Acetaldehyde is manufactured by liquid phase Wacker oxidation of ethylene catalyzed by palladium(II) chloride and copper(II) chloride. This is a process first developed in 1957 – 1959 by Wacker-Chemie and Farbwerke Hoechst (Figure 17).¹³ Glycolaldehyde is produced by rhodium metal catalyzed reaction between syn gas with formaldehyde at 30 MPa. Formaldehyde is industrially made by a silver catalyzed oxidation of methanol at 600 – 650 °C (Figure 17).¹⁴

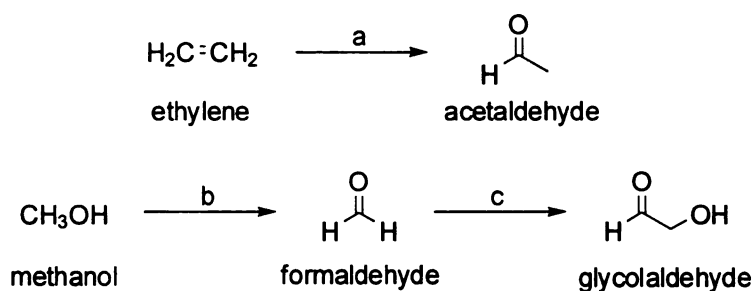


Figure 17. Industrial syntheses of acetaldehyde and glycolaldehyde. (a) PdCl₂, CuCl₂, O₂. (b) Ag, O₂, 1 atm, 600 – 650 °C. (c) CO, H₂, Rh, 30 MPa, 150 °C.

In this chapter, the 2-deoxyribose-5-phosphate aldolase catalyzed aldol reaction and the catalytic hydrogenation will first be presented independently. This allows us to determine all necessary reaction parameters. In the last part of this chapter, the two-step synthetic process was scaled up. Downstream purification using short-path distillation and Kugelrohr distillation will be discussed. Both yields determined by gas chromatography and isolated yields will be reported.

Synthesis of 3,4-dihydroxy-D-butanal using 2-deoxyribose-5-phosphate aldolase

Protein purification of 2-deoxyribose-5-phosphate aldolase

Over-expressed *E. coli* 2-deoxyribose-5-phosphate aldolase was obtained by culturing the strain *E. coli* DH5 α /pVH17 (ATCC86963) (Figure 18).¹⁵ A 4 L culture generally provides 80,000 units of crude 2-deoxyribose-5-phosphate aldolase lysate activity. The protein quality is of paramount importance to the isolated yield of D-1,2,4-butanetriol, as was unambiguously proven in the course of our study. To obtain large quantities of pure 2-deoxyribose-5-phosphate aldolase for the reaction trials, classical enzyme purification was performed. This technique includes protein precipitation, ion-exchange chromatography using an open-column, and FPLC. During the first attempt, a literature purification protocol was followed with minor modifications.⁶ *E. coli* DH5 α /pVH17 crude lysate was subjected to protein precipitation using 1% streptomycin sulfate. The resulting supernatant was fractionated using ammonium sulfate (40 – 60%). Subsequently the protein pellet was resuspended into buffer and loaded on a Mono-Q 10/10 anion exchange column controlled by an Amersham Bioscience AKTA Purifier. Near-homogeneous 2-deoxyribose-5-phosphate aldolase was obtained and analyzed by SDS-PAGE protein electrophoresis (Figure 19, left, lane 2 and 3). Purified protein migrated as a major band on the gel with a molecular weight of approximately 28 kDa. During the purification, 2-deoxyribose-5-phosphate aldolase activity was monitored by a coupled enzyme assay using glycerol 3-phosphate dehydrogenase and triose phosphate isomerase.⁶ The activity was determined by the depletion of NADH as indicated by measuring optical density at 340 nm (Table 1).

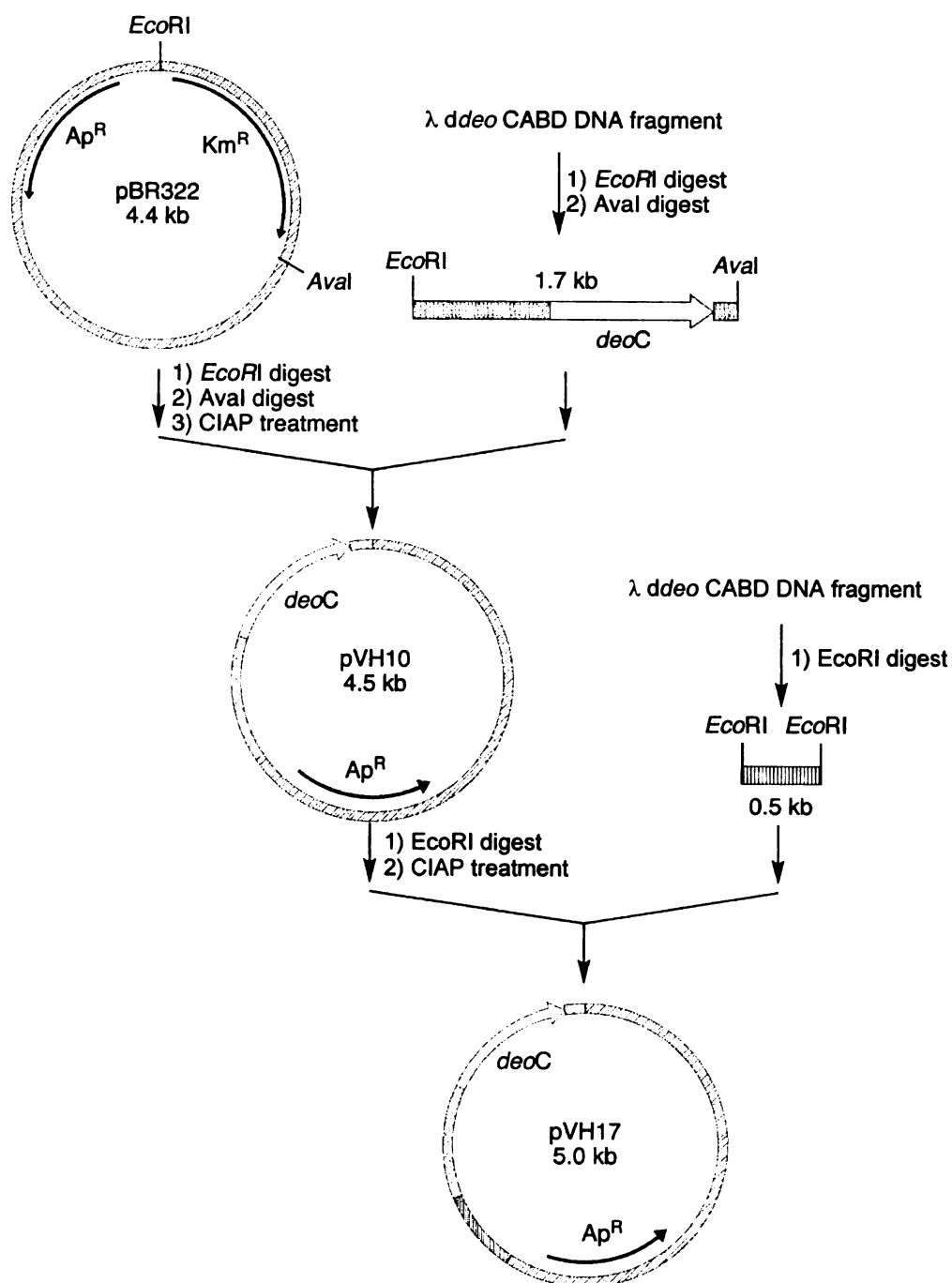


Figure 18. Construction of plasmid pVH17.¹⁵

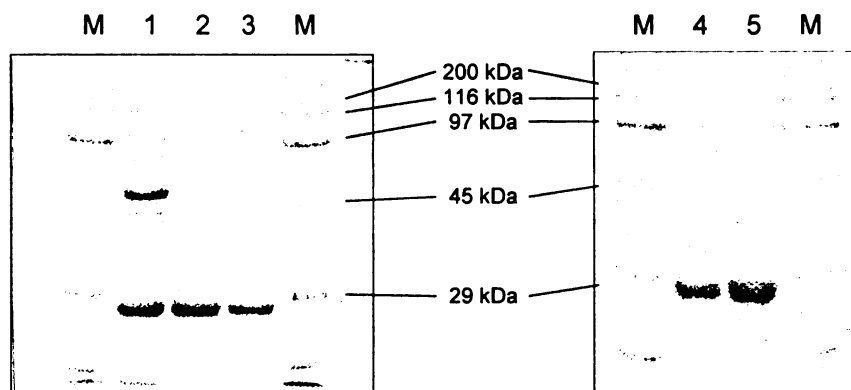


Figure 19. SDS-PAGE of 2-deoxyribose-5-phosphate aldolase purified from *E. coli* DH5α. (left) using literature protocol; (right) using simplified protocol. M, Molecular size marker; lane 1, after $(\text{NH}_4)_2\text{SO}_4$ fractionation; lane 2, fractions collected after Mono-Q purification; lane 3, Mono-Q eluent after dialysis; lane 4, after Q-Sepharose Fast-Flow; lane 5, after Mono-Q.

Table 1. Purification table of 2-deoxyribose-5-phosphate aldolase.

step	volume (mL)	[protein] (mg/mL)	total protein (mg)	specific activity ^a (U/mg)	total activity (U)	recovery (%)	purification (fold)
crude	193	29.7	5732	10.1	57,900	100	1
strep. sul. ^b	206	36.9	7601	8	60,800	100	0.8
$(\text{NH}_4)_2\text{SO}_4$ ^c	70	48.1	3367	14	47,100	81	1.4
dialysis	40	33.1	1324	11.2	14,800	26	1.1
Mono-Q	40	6.7	268	38	10,200	17	3.8

^aone unit (U) of 2-deoxyribose-5-phosphate aldolase corresponds to the formation of 1 μmol of glyceraldehyde-3-phosphate per min at 25 °C.

^bstreptomycin sulfate precipitation (strep. sul.)

^cammonium sulfate precipitation ($(\text{NH}_4)_2\text{SO}_4$)

As indicated in Table 1, the two protein precipitation steps resulted in 1.4 fold purification. Surprisingly, about 32,300 units total enzyme activity was lost after dialyzing the protein solution against buffer, presumably due to the instability of 2-deoxyribose-5-phosphate aldolase in high salt solution. FPLC protein purification using

Mono-Q 10/10 anion exchange column turned out to be successful and 3.8 fold purification was obtained in this single step. Further investigation resulted in a simplified purification protocol involving two column chromatography steps using anion exchange resin. An open column containing 300 mL Q-Sepharose Fast-Flow anion exchange resin was used to purify 3,000 mg crude protein in a single passage. The protein solution was eluted with a gradient up to 1 M NaCl in buffer resulting in a 2.5-fold purification and a 77% recovery of enzyme activity. Near-homogeneous 2-deoxyribose-5-phosphate aldolase was then obtained by chromatography using a Mono-Q column. These protein samples were analyzed by SDS-PAGE protein electrophoresis (Figure 19, right, lane 4-5).

Table 2. Purification table of 2-deoxyribose-5-phosphate aldolase.

step	volume (mL)	[protein] (mg/mL)	total protein (mg)	specific activity ^a (U/mg)	total activity (U)	recovery (%)	purification (fold)
crude	68	31.8	2160	17.6	38,000	100	1
Q-Sep. ^b	29	23.1	670	43.9	29,400	77	2.5
Mono-Q	9	32.7	294	52.4	15,400	40	3

^aone unit (U) of 2-deoxyribose-5-phosphate aldolase corresponds to the formation of 1 μ mol of glyceraldehyde-3-phosphate per min at 25 °C.

^bQ-Sepharose Fast-Flow anion exchange column (Q-Sep.)

Chemical synthesis of 3,4-dihydroxy-D-butanal

3,4-dihydroxy-D-butanal is not a commercially available compound. To provide enough authentic material for the titled chemo-enzymatic process, racemic 3,4-dihydroxy-D-butanal was synthesized from racemic 1,2,4-butanetriol.¹⁶ The 1,2-diol moiety was first selectively protected with acetone as the corresponding acetal in the presence of *p*-toluenesulfonic acid. The remaining terminal hydroxyl group was then oxidized to the aldehyde using pyridinium chlorochromate. These two intermediates were easily purified

by using short-path distillation under reduced pressure. Target racemic 3,4-dihydroxybutanal was obtained in 50% overall yield upon the deprotection of the acetal. The final product was characterized by ^1H and ^{13}C NMR spectroscopy and was identical to literature values.

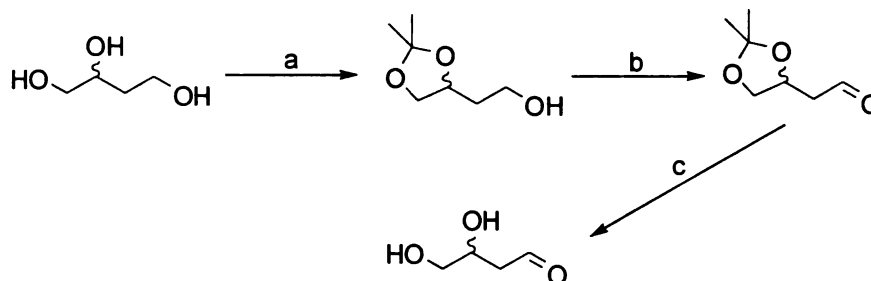


Figure 20. Chemical synthesis of racemic 3,4-dihydroxybutanal. (a) acetone, *p*-toluenesulfonic acid, 85%; (b) pyridium chlorochromate, CH_2Cl_2 , 60%; (c) Dowex 5(H^+), H_2O , quantitative.

Enzymatic synthesis of 3,4-dihydroxy-D-butanal

Small scale enzymatic reactions were carried out using literature conditions for tandem aldol reactions to synthesize six-membered lactol derivatives.⁷ In order to establish the relationship between the enzyme purity and the yield of 3,4-dihydroxy-D-butanal, this unstable aldehyde intermediate was quantitatively converted into the corresponding D-1,2,4-butanetriol by stoichiometric sodium borohydride reduction.

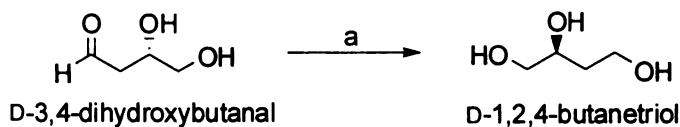


Figure 21. NaBH_4 reaction for 3,4-dihydroxy-D-butanal. (a) NaBH_4 , H_2O .

Upon sodium borohydride reduction, the reaction mixture was passed through a small pad of Dowex 50(H⁺) resin. Boric acid was removed by repeated azeotropic distillation with methanol. This protocol was used to determine the influence of enzyme purity on the yield of 3,4-dihydroxy-D-butanal as summarized in Table 3. Interestingly, using the same amount but higher purity of 2-deoxyribose-5-phosphate aldolase resulted in a significant improvement in the final yield of 3,4-dihydroxy-D-butanal. Using 1,000 activity units of aldolase purified by protein precipitation increased the yield of 3,4-dihydroxy-D-butanal by 9% as supposed to reaction catalyzed by crude cell lysate (Table 3, entries 1 and 2). The same amount of aldolase purified by anion exchange columns generally doubled the product obtained, up to a maximum yield of 50% (Table 3, entries 3 and 4). Furthermore, the amount of 3,4-dihydroxy-D-butanal obtained using Q-Sepharose purified 2-deoxyribose-5-phosphate aldolase (43.9 U/mg) was comparable to the same reaction using enzyme purified by Mono-Q equipped FPLC (52.4 U/mg) (Table 3, entries 3 and 4). This suggested that passing the crude lysate through Q-Sepharose resin offers a one-step method to obtain 2-deoxyribose-5-phosphate aldolase for synthetic purposes.

Table 3. Influence of 2-deoxyribose-5-phosphate aldolase (DERA) purity on 3,4-dihydroxy-D-butanal yield.

entry	DERA batch	specific activity ^a (U/mg)	total units (U)	reaction time (days)	product yield (%)
1	crude	12.0	1,000	3	14
2	protein ppt. ^b	14.0	1,000	3	22
3	Q-Sep. ^c	43.9	1,000	3	45
4	Mono-Q	52.4	1,000	3	50

^aOne unit (U) of 2-deoxyribose-5-phosphate aldolase corresponds to the formation of 1 μmol of glyceraldehyde-3-phosphate per min at 25 °C.

^bProtein precipitation steps using streptomycin sulfate followed by ammonium sulfate (protein ppt.)

^cQ-Sepharose Fast-Flow anion exchange column (Q-Sep.)

Notably, the lactol 2,4-dideoxy-D-hexapyranoside, which resulted from the condensation of one molecule of glycolaldehyde and two molecules of acetaldehyde catalyzed by 2-deoxyribose-5-phosphate aldolase was also formed.⁷ However, this double aldol product was only observed as a minor byproduct in contrast to other α -substituted aldehyde receptors.

Catalytic hydrogenation of racemic 3,4-dihydroxybutanal

Having succeeded in the enzymatic-catalyzed condensation, focus was changed to the catalytic hydrogenation step in the proposed chemo-enzymatic synthesis of D-1,2,4-butanetriol. Based on the previous successful use of commercially available 5 wt% ruthenium on carbon in the hydrogenation of D- and L-malic acid in aqueous medium, the same catalyst was employed to hydrogenate chemically synthesized 3,4-dihydroxybutanal (Figure 16).^{10,11} Variation of pressure, temperature, reaction time, concentration of substrate, rate of stirring and catalyst loading were explored. Representative reaction trials are presented in Table 4. A typical hydrogenation was conducted by dissolving authentic racemic 3,4-dihydroxybutanal in distilled, deionized water (100 mL) in a glass liner. The catalyst was suspended in the 3,4-dihydroxybutanal solution. The liner was inserted into a 500 mL Parr 4575 stainless steel high temperature-high pressure reactor, and the vessel was sealed. The temperature and stir rate were controlled by a Parr 4842 temperature controller. Hydrogen was bubbled through the reaction mixture for 10 – 15 min to remove air while stirring at 100 rpm. The vessel was then charged with hydrogen gas to a pressure below the desired value. After heating the reaction to desired temperature, the hydrogen pressure was adjusted to the desired pressure. The reaction was

stirred at 200 rpm for 1 – 10 h at a constant temperature. When the reaction was completed, the vessel was cooled to room temperature and the pressure was released. After removal of the catalyst by filtration, the reaction mixture was concentrated to dryness under vacuum to afford a colorless oil, which was derivatized with *N,O*-bis-(trimethylsilyl)-trifluoroacetamide (BSTFA) and analyzed by gas chromatography.

In the case of catalytic hydrogenation of D- and L-malic acid, elevated temperature and pressure at 135 °C and 5,000 psi is required.^{10,11} Competing deoxygenation and cracking reactions generate byproducts that are very difficult to separate from 1,2,4-butanetriol. In the chemo-enzymatic synthesis, a carboxylate functionality is substituted by a more reactive aldehyde and thus milder reaction conditions could be employed. In Table 4, catalytic hydrogenation of a 250 mM aqueous solution of 3,4-dihydroxybutanal at 400 psi H₂, a 700 rpm stir rate, and a reaction temperature of 50 °C for 5 h resulted in a 23% yield of D-1,2,4-butanetriol (entry 1). Increasing the catalyst loading had a pronounced impact on the yield of 1,2,4-butanetriol even under reduced temperature and pressure. The yield of 1,2,4-butanetriol increased to 68% using 0.5 mol% of the same ruthenium on carbon catalyst at room temperature (entry 2). Lowering the pressure from 400 psi to 100 psi did not change the yield of product (entry 3). Increasing the catalyst loading to 1 mol% led to another increase in product yield (entry 4). Attempts to increase the catalyst loading beyond 1 mol% did not translate to further improvement in 3,4-dihydroxybutanal reduction yields (result not shown). In addition, reducing the stir rate to 400 rpm resulted in no difference in 3,4-dihydroxybutanal reduction yields (entry 5). However, a pronounced decrease in product yield was noticed with doubled substrate concentration (entry 6). At last, both reaction time and temperature were found to be the

keys to complete conversion. A reaction with product yield exceeding 90% was achieved simply by doubling the reaction time (entry 7 vs. entry 6). Increasing reaction temperature led to a 10% increase in the final yield of 1,2,4-butanetriol (entry 10 vs. 8). Catalytic hydrogenation of a 250 mM aqueous solution of 3,4-dihydroxybutanal at 200 psi H₂, a 700 rpm stir rate, and a reaction temperature of 30 °C for 5 h resulted in a quantitative yield of 1,2,4-butanetriol.

Table 4. Catalytic hydrogenation of authentic 3,4-dihydroxybutanal.

entry	Ru on C (mol%)	substrate (mM)	temp (°C)	pressure (psi)	stir rate (rpm)	time (h)	GC yield (%)
1	0.1	250	50	400	700	5	23
2	0.5	250	25	400	700	5	68
3	0.5	250	25	100	700	5	71
4	1	250	25	100	700	5	82
5	1	250	25	100	400	5	80
6	1	500	25	100	400	5	62
7	1	500	25	100	400	10	91
8	1	250	25	200	700	5	93
9	1	250	20	200	700	5	81
10	1	250	30	200	700	5	quant.

Chemo-enzymatic synthesis of D-1,2,4-butanetriol

Having developed the steps separately, the chemo-enzymatic synthesis was applied to the preparation of D-1,2,4-butanetriol using acetaldehyde and glycolaldehyde as starting material. As shown in Figure 22, the conversion began with the condensation of acetaldehyde and glycolaldehyde using 2-deoxyribose-5-phosphate aldolase. The intermediate, 3,4-dihydroxy-D-butanal was allowed to pass through a pad of Dowex 50

(H⁺). The reaction mixture was then concentrated to about 150 mL in volume and subjected to the high pressure-high temperature parr reactor for catalytic hydrogenation. The optimized hydrogenation conditions using authentic racemic 3,4-dihydroxybutanal was applied in this step. The yield of product D-1,2,4-butanetriol in the final reaction mixture was found to be 30% as quantified by using ¹H NMR and gas chromatography.

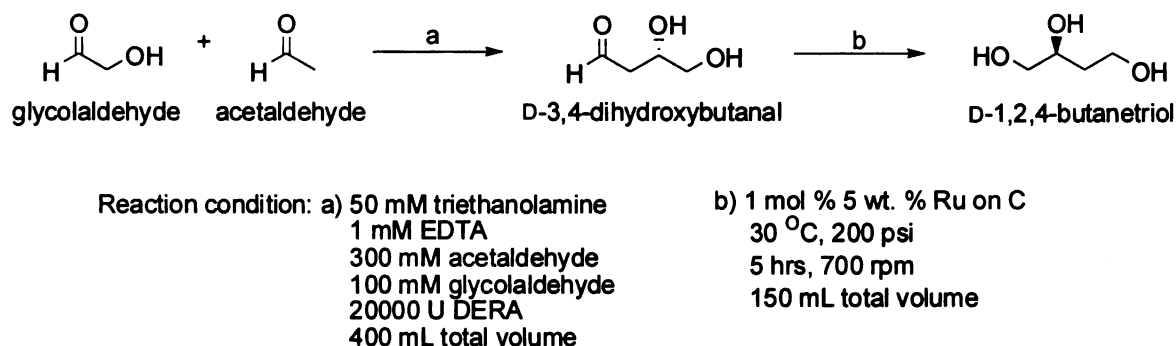


Figure 22. Reaction parameters for the chemo-enzymatic synthesis of D-1,2,4-butanetriol.

To determine the enantiomeric purity of chemo-enzymatically synthesized D-1,2,4-butanetriol, product sample in the reaction mixture was partially purified and derivatized using Mosher's reagent, (*S*)-(+)- α -methoxy- α -(trifluoromethyl)phenylacetyl chloride.¹⁷ Analysis of esterified 1,2,4-butanetriol using an HPLC Chiralpak AD column (Daicel Chemical) revealed that the ee% of chemo-enzymatically synthesized D-1,2,4-butanetriol was >99%.

Purification of product D-1,2,4-butanetriol, however, was not straightforward. Attempts to purify crude D-1,2,4-butanetriol in the reaction mixture using short path distillation in vacuo was unsuccessful. Distillation using a Kugelrohr apparatus resulted in

a 17% yield of pure D-1,2,4-butanetriol as characterized by both ^1H NMR and gas chromatography.

Discussion and future work

The commercially employed route for the manufacture of 1,2,4-butanetriol relies on a stoichiometric sodium borohydride reduction of dimethyl malate.¹⁸ The major issues associated with this current synthesis include the disposal of byproduct borate salts and the substantial cost of using a stoichiometric reducing agent sodium borohydride. Although its derivative 1,2,4-butanetriol trinitrate offers superior physical properties over nitroglycerin as an energetic plasticizer, these factors significantly limit the application of 1,2,4-butanetriol trinitrate for military and civilian purposes. To this end, novel alternatives involve catalytic hydrogenation of malic acid using heterogeneous ruthenium on carbon catalyst,^{10,11} chemo-enzymatic synthesis using L-ascorbic acid¹¹ and the whole-cell microbial synthesis using a created biosynthetic pathway^{10,16} had been developed in the Frost group to synthesize precursor 1,2,4-butanetriol. Catalytic malic acid hydrogenation does not generate a salt stream as a reaction byproduct. However, the cost of the ruthenium metal catalyst and the capital investment for constructing a high pressure-high temperature reactor for large-scale 1,2,4-butanetriol production are problematic. In addition, the formation of a number of polyol byproducts complicates the downstream processing of the target molecule.^{10,11} The four-step chemo-enzymatic approach using L-ascorbic acid afforded 1,2,4-butanetriol in 50% overall yield.¹¹ Although competing deoxygenation and cracking reactions are eliminated, a long synthesis and loss of half of the starting material makes this process less attractive. The

microbial synthesis methodology offers an alternative that ultimately turned out to be the best approach. However, at the time when the research was conducted, this complicated two microbe, four enzyme synthesis suffered from low yields, low product concentrations and byproduct formation.¹⁰ This prompted elaboration of a chemo-enzymatic synthesis using 2-deoxyribose-5-phosphate aldolase.

Chemo-enzymatic synthesis of D-1,2,4-butanetriol presents several interesting features that have not been explored by other approaches. This two-step route has the advantage of relying on an enzyme that is produced in great abundance by a commercially available strain and requires the catalytic hydrogenation of an aldehyde as opposed to a carboxylate. Intermediate 3,4-dihydroxybutanal underwent catalytic hydrogenation over ruthenium on carbon at a temperature of 30 °C and H₂ pressure of 200 psi, which sharply contrasts with the conditions (135 °C, and 5,000 psi H₂) that were required for catalytic malic acid hydrogenation. In addition, starting materials for this synthesis are achiral molecules acetaldehyde and glycolaldehyde. These small molecular building blocks are commercially available.

1,2,4-Butanetriol trinitrate that has been used as an energetic material has been a racemic mixture, which reflects the use of methyl D-, L-malate as the starting material for commercial synthesis of precursor racemic 1,2,4-butanetriol. The difference between the melting points of the pure enantiomers relative to the melting point for the racemic mixture of enantiomers is an important consideration in the utilization of energetic materials. Moreover, 1,2,4-butanetriol trinitrate was proven to be an alternative to nitroglycerin for use as a vasodilator in the treatment of angina.¹⁹ Using a pure chiral chemical could conceivably minimize the chances of encountering adverse side effects. In

the 1,2,4-butanetriol syntheses employing D-, L-malic acid or L-ascorbic acid as starting materials, racemization was observed during the problematic reduction of carboxylates under elevated temperature and pressure.¹¹ Fortunately, synthesizing a single D-1,2,4-butanetriol using this chemo-enzymatic synthesis was found to be possible due to the milder catalytic hydrogenation reaction condition involved. Unlike the microbial approach where the stereochemistry is defined by the starting D-xylonic acid or L-arabinonic acid, 2-deoxyribose-5-phosphate aldolase generates the product chirality using an enantioselective aldol condensation.

Despite all these advantages of utilizing this chemo-enzymatic synthesis using 2-deoxyribose-5-phosphate aldolase, the enzymatic condensation of glycolaldehyde and acetaldehyde and the catalytic hydrogenation only led to a 30% overall yield of D-1,2,4-butanetriol. The major concern lies in the enzymatic aldol condensation reaction using 2-deoxyribose-5-phosphate aldolase. First, the catalyst loading was 20,000 units per gram of D-1,2,4-butanetriol synthesized, which is considered to be very high for an enzyme-catalyzed reaction. Use of such a high concentration of enzyme would make the process prohibitively expensive, in addition to making isolation of product from the reaction mixture difficult. Second, the reaction time was on the order of several days, therefore leading to a considerable increase in anticipated production costs. These issues can be addressed by developing an improved 2-deoxyribose-5-phosphate aldolase. As a case study, there are generally two different approaches pursued in the industrial community to improve this enzyme toward the synthesis of atorvastatin. DSM engineered 2-deoxyribose-5-phosphate aldolase for higher affinity and resistance toward the substrate chloroacetaldehyde, which is the key building block for atorvastatin side chain synthesis.⁸

They employed error-prone PCR and DNA shuffling techniques to generate the aldolase mutant library. Subsequent screening led to the isolation of a ten-fold improved variant for the reaction of interest. In 2002, Diversa also discovered a 2-deoxyribose-5-phosphate aldolase alternative from an environmental genomic DNA library collected from a variety of habitats.⁹ By using a combination of enzyme selection and screening, the newly discovered enzyme resulted in a production rate of 30.6 g/L/h. Catalyst loading was also improved by a 10-fold decrease while maintaining high enantio- and diastereoselectivity. These examples suggest that it may be possible to increase the affinity and resistance of 2-deoxyribose-5-phosphate aldolase towards glycolaldehyde. At the time when our chemo-enzymatic synthesis of D-1,2,4-butanetriol was completed, our research focus turned to the development of microbial syntheses using pentoses as starting materials. We believed that a synthesis employing a single microbe would ultimately represent a general, practical, and economical route to 1,2,4-butanetriol enantiomers.

REFERENCE

- ¹ Valentin-Hansen, P.; Boetius, F.; Hammer-Jespersen, K.; Svendsen, I. *Eur. J. Biochem.* **1982**, *125*, 561.
- ² Blank, J.; Hoffe, P. *Mol. Gen. Genet.* **1972**, *116*, 291.
- ³ Racker, E. *J. Biol Chem.* **1952**, *196*, 347.
- ⁴ Pricer, W. E.; Horecker, B. L. *J. Biol. Chem.* **1960**, *5*, 1292.
- ⁵ (a) Hoffee, P. *Arch. Biochem. Biophys.* **1968**, *126*, 795. (b) Horecker, B. L.; Pontremoli, S.; Ricci, C.; Cheng, T. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1942.
- ⁶ (a) Barbas, C. F. III; Wang, Y.-F.; Wong, C.-H. *J. Am. Chem. Soc.* **1990**, *112*, 2013. (b) Chen, L.; Dumas, P. D.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 741.
- ⁷ (a) Gijsen, H. J. M.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 8422. (b) Wong, C.-H.; Garcia-Junceda, E.; Chen, L.; Blanco, O.; Gijsen, H. J. M.; Steensma, D. H. *J. Am. Chem. Soc.* **1995**, *117*, 3333.
- ⁸ Jennewein, S.; Schurmann, M.; Wolberg, M.; Hilker, I.; Luiten, R.; Wubbolts, M.; Mink, D. *Biotechnol. J.* **2006**, *1*, 537.
- ⁹ (a) Liu, J.; Hsu, C.-C.; Wong, C.-H. *Tetrahedron Lett.* **2004**, *45*, 2439. (b) Greenberg, W. A.; Varvak, A.; Hanson, S. R.; Wong, K.; Huang, H.; Chen, P.; Burk, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5788.
- ¹⁰ Niu, W.; Molefe, M. N.; Frost, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 12998.
- ¹¹ Molefe, M. N. *PhD Thesis*, Michigan State University, 2005.
- ¹² (a) Folkers, K.; Adkins, H. *J. Am. Chem. Soc.* **1932**, *54*, 1145. (b) Trenner, N. R.; Bacher, F. A. *J. Am. Chem. Soc.* **1949**, *71*, 2352. (c) Zhang, Z.; Jackson, J. E.; Miller, D. *J. Appl. Cat. A* **2001**, *219*, 89.
- ¹³ Hagemeyer, H. J. In *Kirk-Othmer Encyclopedia of Chemical Technology*; Wiley: 2002, DOI: 10.1002/0471238961.0103052008010705.a01.pub2.
- ¹⁴ Gerberich, H. R.; Seaman, G. C.; Hoechst-Celanese Corporation In *Kirk-Othmer Encyclopedia of Chemical Technology*; Wiley: 1994, DOI: 10.1002/0471238961.0615181307051802.a01

- ¹⁵ (a) Valentin-Hansen, P.; Aiba, H.; Schumperli, D. *EMBO J.* **1982**, *1*, 317. (b) Jorgensen, P.; Collins, J.; Valentin-Hansen, P. *Mol. Gen. Genet.* **1977**, *155*, 93.
- ¹⁶ Niu, W. *PhD Thesis*, Michigan State University, 2004.
- ¹⁷ Dale, J. A.; Mosher, H. S. *J. Am. Chem. Soc.* **1973**, *95*, 512.
- ¹⁸ (a) Monteith, M. J.; Schoefield, D.; Bailey, M. PCT Int. Appl. WO 98/08793, 1998. (b) Ikai, K.; Mikami, M.; Furukawa, Y.; Ho, S. PCT Int. Appl. WO 99/44976, 1999.
- ¹⁹ Needleman, P. *Annu. Rev. Pharmacol. Toxicol.* **1976**, *16*, 81.

CHAPTER THREE

Improving Microbial Synthesis of 1,2,4-Butanetriol

Overview

The objective of this chapter is to develop a novel and improve an existing microbial synthesis for 1,2,4-butanetriol as a precursor for the production of 1,2,4-butanetriol trinitrate. A newly proposed biosynthesis of 1,2,4-butanetriol using erythritol as the starting material under an anaerobic environment was explored (Figure 23). Efforts were also diverted to improve the existing aerobic, pentose-based microbial process to D-1,2,4-butanetriol using D-xylose as starting material (Figure 10).

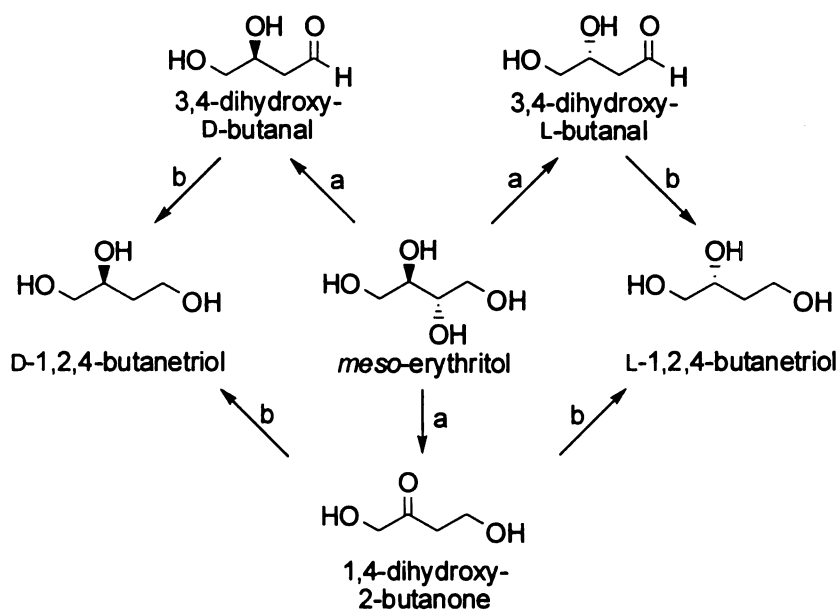


Figure 23. Proposed artificial biosynthetic pathway for 1,2,4-butanetriol from erythritol. (a) diol or glycerol dehydratase. (b) alcohol dehydrogenase

Biosyntheses of D-1,2,4-butanetriol and L-butanetriol has been accomplished by creating biosynthetic pathways not found in nature. A two-step enzymatic conversion of erythritol into 1,2,4-butanetriol was also thought to be possible. For example, an active erythritol dehydratase could convert erythritol into 3,4-dihydroxybutanal or 1,4-dihydroxy-2-butanone (Figure 23). Alcohol dehydrogenase activity could then reduce D- or L-3,4-dihydroxybutanal (Figure 23) or 1,4-dihydroxy-2-butanone (Figure 23) into 1,2,4-butanetriol. Although the desired erythritol dehydratase activity was identified, the activity was too low to be practically exploited. Instead of elaborating this newly explored route into a microbial process, the research focus was changed to improving the pentose-based pathway previously developed in the Frost group using D-xylose. In the second part of this chapter, efforts toward isolating a more active 3-deoxy-*glycero*-pentulosonate decarboxylase will be discussed. Potential wild-type decarboxylases were identified *in silico* using bioinformatics tools and decarboxylase mutants were generated by site directed mutagenesis and directed evolution methodologies. UV/vis spectrophotometry, gas chromatography and high-throughput 96-well plate assays were developed to evaluate these enzyme candidates. Finally, *E. coli* WN13/pWN7.126B, which synthesizes D-1,2,4-butanetriol from D-xylose was subjected to further metabolic engineering to improve the concentrations and yields of D-1,2,4-butanetriol. Differential expression of *P. putida* 3-deoxy-*glycero*-pentulosonate decarboxylase (MdlC), *E. coli* molecular chaperonins (GroES and GroEL), D-1,2,4-butanetriol dehydrogenase (AdhP) and keto-acid dehydrogenases (YiaE and YcdW) were studied. The resulting second generation microbial catalyst of D-1,2,4-butanetriol was evaluated under fermentor-controlled conditions.

Artificial biosynthesis of 1,2,4-butanetriol from erythritol

Background

The first step in the proposed bioconversion of erythritol to D-, and L-1,2,4-butanetriol involves the dehydration of erythritol to D- and L-3,4-dihydroxybutanal (Figure 23). It was believed that either a glycerol dehydratase or diol dehydratase would be able to catalyze this dehydration reaction. Glycerol dehydratase plays a critical role in the ability of microbes to ferment glycerol to 1,3-propanediol while diol dehydratase catalyzes a similar dehydration essential to the microbial fermentation of 1,2-propanediol.¹ The final step of the erythritol-based route requires the reduction of D- and L-3,4-dihydroxybutanal to the corresponding D-, and L-1,2,4-butanetriol, respectively (Figure 23).

Erythritol is commercially used as a sweetening agent in foods and is derived from glucose via a microbial fermentation route.² Cerester, which has been the major producer of erythritol, was acquired by Cargill in 2002.³ The starch fraction obtained during corn wet-milling is an abundant source of inexpensive glucose. Titters as high as 130 g/L and yields of 46% have been reported for the microbial synthesis of erythritol from glucose.⁴ Considerable progress had been made in detailing the steps involved in microbial synthesis of erythritol from glucose.⁵ This progress suggests, over the long term, that it may be possible to construct a microbe capable of synthesizing 1,2,4-butanetriol directly from glucose with erythritol as an *in vivo* biosynthetic intermediate.

Glycerol dehydratase (EC 4.2.1.30) and diol dehydratase (EC 4.2.1.28) can catalyze the conversion of glycerol (Figure 24), 1,2-propanediol (Figure 25) and 1,2-ethanediol to the corresponding aldehydes.^{6,7} These enzymatic reactions proceeds through

a mechanism involving coenzyme B₁₂ as an essential cofactor. Coenzyme B₁₂ contains a cobalt-carbon bond that is stable in water, although easily undergoes a homolytic cleavage reaction when a suitable substrate is bound to the enzyme. This homolytic bond cleavage initiates the catalytic cycle of all coenzyme B₁₂ dependent enzymatic processes.^{8,9}

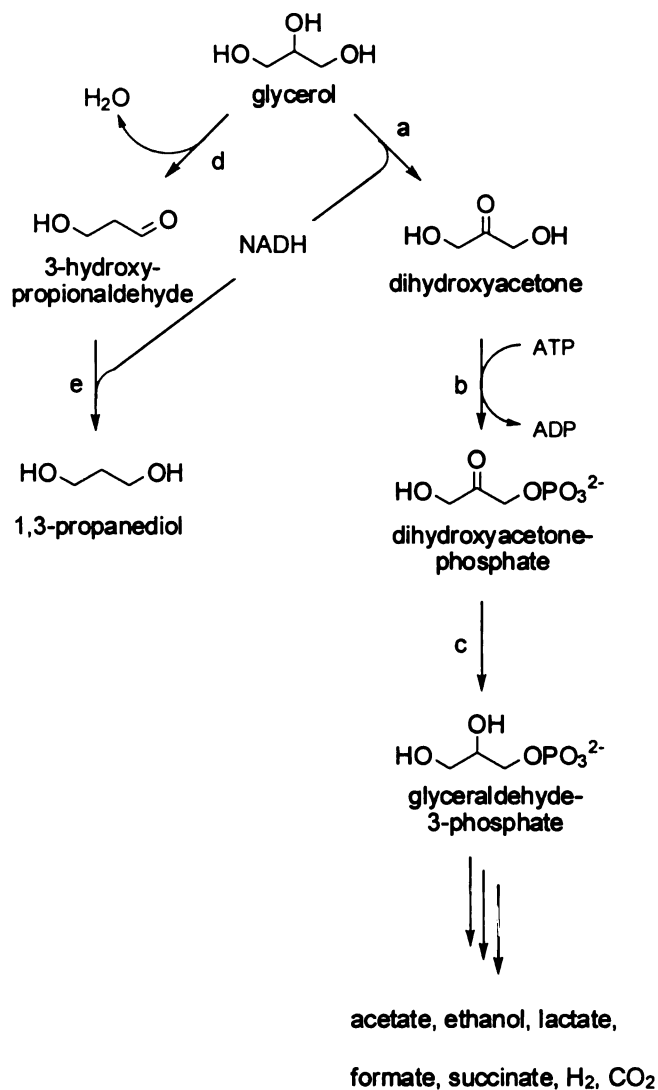


Figure 24. Anaerobic utilization of glycerol. (a) glycerol dehydrogenase; (b) dihydroxyacetone kinase; (c) glyceraldehyde-3-phosphate dehydrogenase; (d) glycerol dehydratase; (e) 1,3-propanediol dehydrogenase.

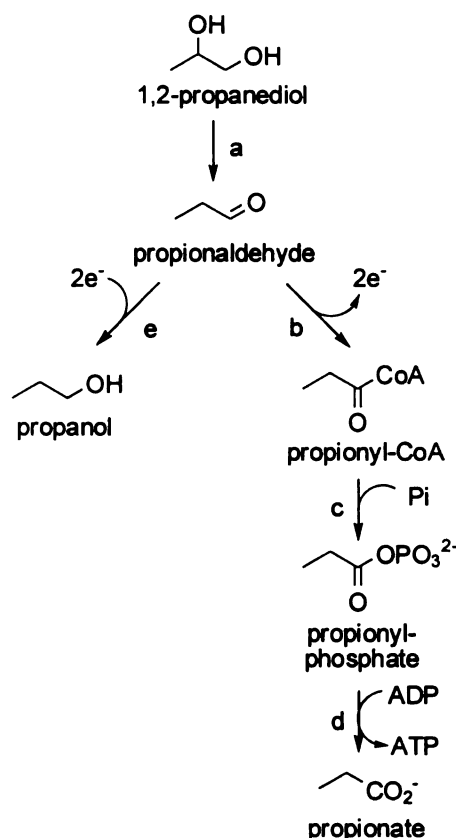


Figure 25. Anaerobic utilization of 1,2-propanediol. (a) diol dehydratase; (b) propionaldehyde dehydrogenase; (c) phosphotransacylase; (d) propionate kinase; (e) 1-propanol dehydrogenase.

Glycerol and diol dehydratases are involved in the utilization of small molecule under anaerobic conditions (Figure 24 and 3). These enzymes catalyze molecular rearrangements that generate an aldehyde, which can be reduced during glycerol fermentation. On the other hand, the aldehyde generated can also dismutate to oxidized and reduced compounds in 1,2-propanediol fermentation. The anaerobic degradation of glycerol is initiated by two enzymes (Figure 24). Glycerol dehydrogenase forms dihydroxyacetone and glycerol dehydratase produces 3-hydroxypropionaldehyde which is further reduced to 1,3-propanediol.^{6,10,11} 1,2-Propanediol fermentation shares a very

similar mechanism as ethanolamine ammonia-lyase, which utilizes coenzyme B₁₂ (Figure 25). Diol dehydratase first converts 1,2-propanediol into propionaldehyde.^{12,13,14} This aldehyde dismutates into propanol and propionyl-CoA. These reactions allow the generation of ATP during the propionate kinase catalyzed reaction (Figure 25).^{12,15} Glycerol and its closely related diol dehydratases were extensively studied. It had been shown that these enzymes have similar molecular masses and substrate spectra.^{6,16,17,18}

Glycerol and diol dehydratases are known to undergo irreversible mechanism-based inactivation by glycerol during catalysis.^{7,19} Inactivation by glycerol involves irreversible cleavage of the cobalt-carbon bond of coenzyme B₁₂, forming 5'-deoxyadenosine and an alkylcobalamin-like species.^{19a,20} Irreversible inactivation is brought about by tight binding of the modified coenzyme. Such suicide inactivation seems enigmatic, since glycerol is essential for glycerol metabolism. However, it had been shown that glycerol-inactivated dehydratases can be reactivated by an exchange of the modified coenzyme for an intact coenzyme B₁₂ in the presence of ATP, Mg²⁺ and a rescue protein.²¹ The rescue protein is made up of two subunits and is located in the same gene cluster as its corresponding dehydratase.

The proposed alternate biosynthesis of 1,2,4-butanetriol utilizes only two enzyme-catalyzed steps, which appears to make this erythritol-based synthesis a simpler route relative to the pentose-based syntheses previously developed in the Frost group. This presumption is reinforced with the successful identification of dehydrogenases capable of reducing D- and L-3,4-dihydroxybutanal. However, the stereoselectivity as well as the regioselectivity in the enzyme-catalyzed erythritol dehydration reactions cannot be

predicted. This contrasts with the predictable, stereospecific synthesis of D-1,2,4-butanetriol from D-xylose and L-1,2,4-butanetriol from L-arabinose.

Dehydratase-catalyzed formation of an aldehyde from erythritol could lead to either D- or L-3,4-dihydroxybutanal. The enantioselectivity of this desymmetrization cannot be predicted (Figure 23). Furthermore, the regioselectivity of the dehydratase-catalyzed reaction cannot be predicted. It is possible that 1,4-dihydroxy-2-butanone may be produced from erythritol by the dehydratase. Formation of 1,4-dihydroxy-2-butanone would wipe out the prochirality of erythritol. Whether pure enantiomers or enantiomeric mixture of 1,2,4-butanetriol were synthesized would be dictated by the enantioselectivity of the alcohol dehydrogenase catalyzing the reduction of 1,4-dihydroxy-2-butanone. Irrespective of 3,4-dihydroxybutanal or 1,4-dihydroxy-2-butanone intermediacy, synthesis of mixtures of 1,2,4-butanetriol enantiomers is an acceptable outcome. This follows from the current exclusive use of racemic 1,2,4-butanetriol trinitrate in high energy material applications.

Identification of erythritol dehydratase candidates

A wide range of microorganisms are able to ferment glycerol and 1,2-propanediol. To investigate whether these naturally occurring diol and glycerol dehydratase converts a structurally similar substrate erythritol into its corresponding aldehyde or ketone, enzyme candidates from seven bacteria were identified and are shown in Table 5. Fermentation of glycerol by *Citrobacter freundii* and *Klebsiella pneumoniae* had been studied exclusively.²² In the absence of an external oxidant, glycerol is fermented by a dismutation process involving two pathways, leading to glycerol oxidation, the other being

responsible for oxidation of NADH (Figure 24). All four key enzymes and the corresponding genes have been characterized in *C. freundii* and *K. pneumoniae*.^{6,11,23,24,25,26} In *C. freundii*, the structural genes of the glycerol dehydratase (*dhaBCE*) are part of the *dha* regulon along with the genes encoding three other key enzymes (*dhaD*, *dhaK*, *dhaT*) of the pathway. The glycerol dehydratase in *K. pneumoniae* is comprised of three different subunits (*gldABC*) and catalyzes the same anaerobic conversion of glycerol. The *pddABC* encoding diol dehydratase in *K. pneumoniae* has also been identified and characterized.²⁷ *Klebsiella oxytoca*^{17,28} and *Salmonella typhimurium*²⁹ were reported to produce a similar diol dehydratase when grown anaerobically in 1,2-propanediol or glycerol. *Klebsiella oxytoca pddABC* genes were further characterized by cloning and expressing in *E. coli*.¹⁶ Interestingly, although the *pduCDE* genes of *S. typhimurium* and the *pddABC* genes of *K. oxytoca* are over 90% identical in nucleotide sequence, *S. typhimurium* cannot use 1,2-propanediol nor glycerol as a sole carbon source to grow under anaerobic conditions.²⁹ The glycerol fermentation pathway of *Clostridium pasteurianum* is slightly different. In addition to the products mentioned above, it synthesizes butyric acid, butanol and ethanol.^{30,31} The *dhaBCE* gene cluster encoding glycerol dehydratase in *Cl. pasteurianum* were identified and had a high amino acid sequence identity to those characterized in *K. pneumoniae* and *C. freundii*. The last candidates were *Lactobacillus brevis* strains LB18 and 19. Although genetic information were not available for these strains, they were the only known strains that transform *meso*-2,3-butanediol into 2-butanol. They are of particular interest due to the structural similarity between *meso*-2,3-butanediol and erythritol.³²

Table 5. Glycerol and diol dehydratase candidates

entry	strain	gene locus	plasmid
1	<i>Citrobacter freundii</i> DSM30040	<i>dhaBCE</i> ^a	pCS120 ³³
2	<i>Klebsiella pneumoniae</i> ATCC25955	<i>gldBC</i>	pML5.176
		<i>gldABC</i> ^b	pML5.200
		<i>dhaT</i> ^c (<i>P</i> _{tac})	pML5.179
		<i>dhaT</i> ^c (<i>P</i> _{T7})	pWN5.220A
		<i>gldABC-dhaT</i>	pML5.226
		<i>pddABC</i> ^d - <i>dhaT</i>	pWN5.260A
3	<i>Klebsiella oxytoca</i> ATCC8724	<i>pddABC</i> ^d - <i>ddrAB</i> ^f - <i>dhaT</i>	pWN5.258A
4	<i>Klebsiella oxytoca</i> ATCC43165		
5	<i>Salmonella typhimurium</i> TT10324	<i>pduCDE</i> ^e	pXY39 ²⁹
6	<i>Clostridium pasteurianum</i> DSM525	<i>dhaBCE</i> ^a	pFL1 ³⁴
7	<i>Lactobacillus brevis</i> LB18		
8	<i>Lactobacillus brevis</i> LB19		

^{a,b} glycerol dehydratase

^{d,e} diol dehydratase

^c 1,3-propanediol oxidoreductase

^f rescue protein

Bacterial strains *K. pneumoniae* ATCC25955, *K. oxytoca* ATCC8724 and ATCC43165 were obtained from the American Type Culture Collection. *L. brevis* LB18 (CNRZ734) and LB19 (CNRZ735) were obtained from the Centre National de Recherches Zootechniques in France. As a preliminary study, these five wild-type strains were screened *in vivo* by culturing anaerobically in medium containing erythritol as substrate. The resulting culture medium was subject to sodium borohydride reduction followed by BSTFA derivatization and analyzed by gas chromatography. No 1,2,4-butanetriol production was observed.

Plasmid Constructions

A total of six plasmids were employed in the *in vitro* enzyme screening assays using glycerol, 1,2-propanediol and erythritol as substrates. The genes encoding coenzyme B₁₂ dependent glycerol or diol dehydratase in *C. freundii* DSM30040, *Cl. Pasteurianum* DSM525, *K. pneumoniae* ATCC25955, *K. oxytoca* ATCC8724 and *S. typhimurium* TT10324 were expressed in *E. coli*. These plasmids were either obtained from individual laboratories or prepared in-house (Table 5). A 2.7 kb PCR product containing the *dhaBCE* gene fragment was amplified using *C. freundii* DSM30040 genomic DNA as a template. The plasmid pCS120 was constructed by cloning this PCR product into pBluescript SK⁺ (Stratagene).³³ The *dhaBCE* gene cluster was transcribed in the same direction as the *lac* promoter on the pBluescript SK⁺ cloning vector. Recombinant cosmid pFL1 contains a 13.5 kb insert of *Cl. Pasteurianum* genomic DNA and harbors the genes encoding the reductive branch of glycerol breakdown.³⁴ The presence of *dhaBCE* genes encoding glycerol dehydratase was confirmed by DNA sequencing. Plasmid pXY39 was constructed by cloning *S. typhimurium pduCDE* into vector *placI^QPO-BgIII*.²⁹ This vector expresses the LacI^Q protein and transcription of the cloned dehydratase genes will therefore be induced by the addition of IPTG to the medium. These three plasmids were obtained as gifts from other research laboratories.

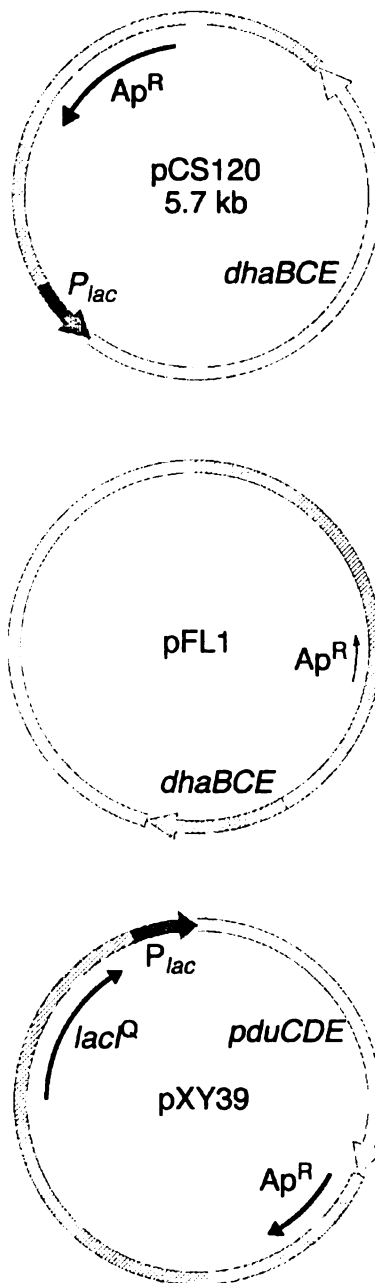


Figure 26. Plasmid maps of pCS120, pFL1 and pXY39.

To construct a plasmid with a *K. pneumoniae* *gldABC*, a 1 kb DNA fragment encoding *gldBC* was amplified from *K. pneumoniae* ATCC25955 genomic DNA and subsequently digested with *Sa*I and *Hind*III. Ligation of this fragment into the *Sa*I/*Hind*III digested pJF118EH resulted in pML5.176 (Figure 27). Another 1 kb fragment encoding glycerol dehydratase subunit GldA was amplified using the same genomic DNA followed by restriction enzyme digestion using *Eco*RI and *Sa*I. Plasmid pML5.200 was obtained by ligating this fragment into *Eco*RI/*Sa*I digested pML5.176 (Figure 28). The resulting *gldABC* gene cluster was transcribed in the same direction as the *tac* promoter located on pJF118EH. Transcription of the glycerol dehydratase was initiated by the pJF118EH-localized *tac* promoter, which was under the control of LacI^Q repressor protein encoded by the *lacI*^Q gene.

The plasmid that expresses *K. pneumoniae* *pddABC* encoding diol dehydratase and *dhaT* encoding 1,3-propanediol oxidoreductase was derived from cloning vector pT7-7. This particular 1,3-propanediol oxidoreductase is known to be active toward the conversion of 3,4-dihydroxybutanal into 1,2,4-butanetriol.³⁵ A 1.2 kb fragment containing the *K. pneumoniae* ATCC25955 *dhaT* gene was excised from plasmid pWN5.022³⁶ using *Bam*HI. This fragment was ligated into the *Bam*HI site of pT7-7 to afford plasmid pWN5.220A (Figure 29). A 2.9 kb DNA fragment containing the *pddABC* gene was amplified from the genomic DNA of the same *K. pneumoniae* strain and subsequently digested with *Eco*RI. Ligation of this fragment into the *Eco*RI site of pWN5.220 yielded plasmid pWN5.260A (Figure 30). On plasmid pWN5.260A, the *pddABC* and *dhaT* genes are transcribed in the same direction as the phage T7 promoter on pT7-7.

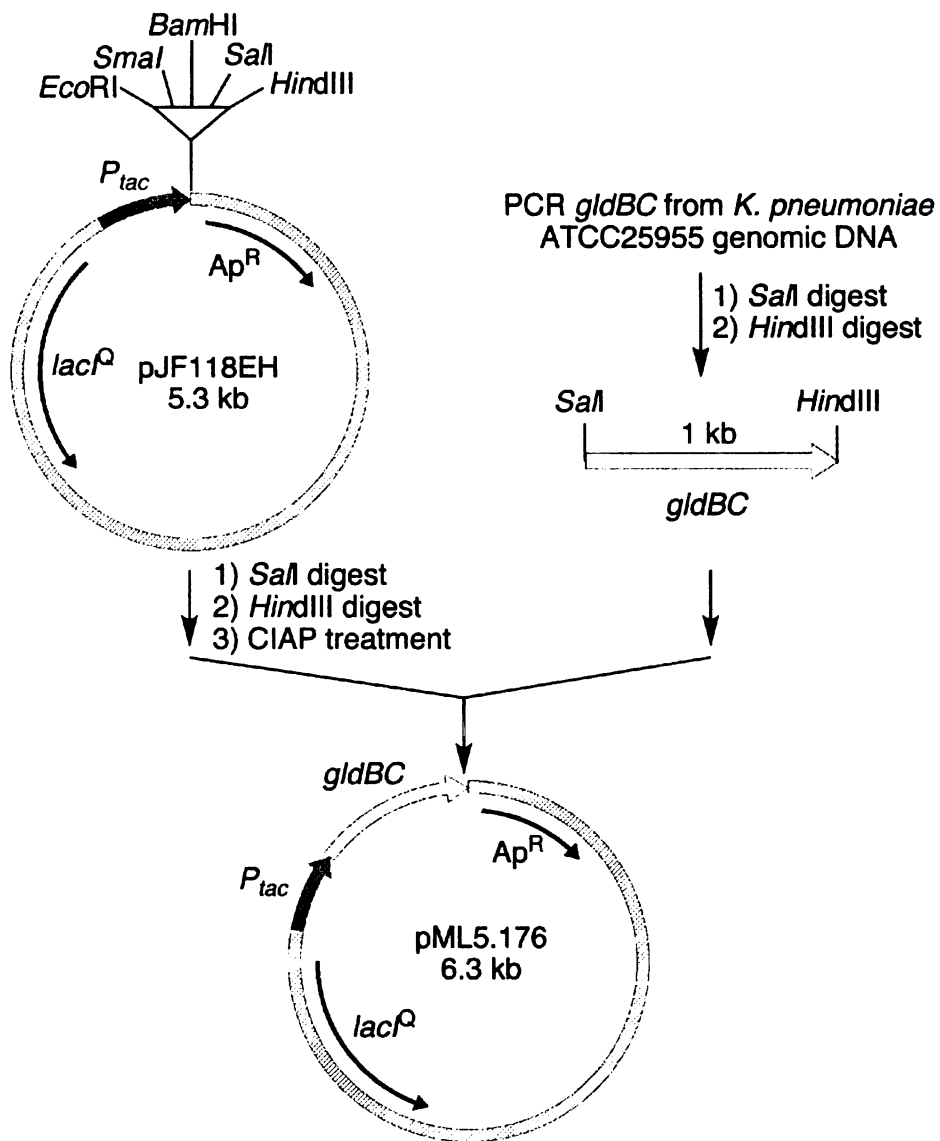


Figure 27. Construction of plasmid pML5.176.

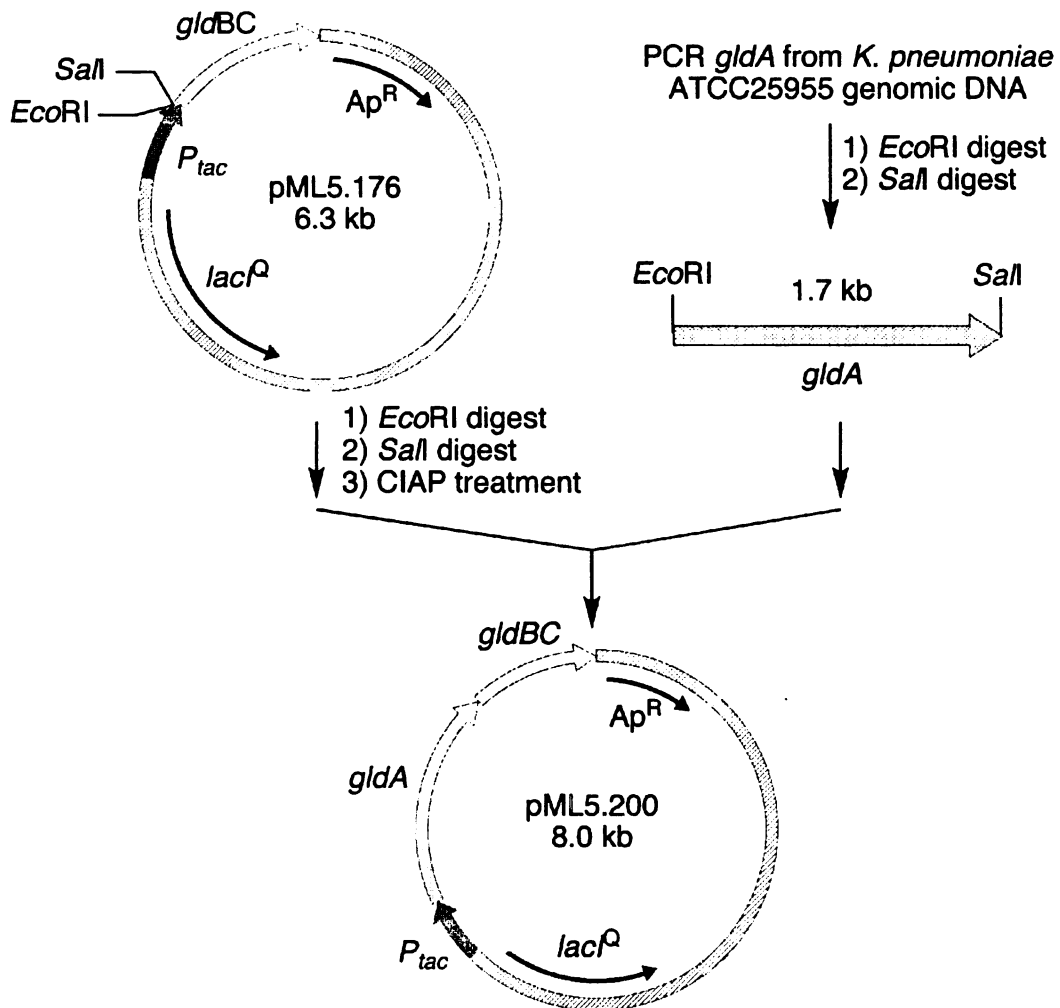


Figure 28. Constuction of plasmid pML5.200.

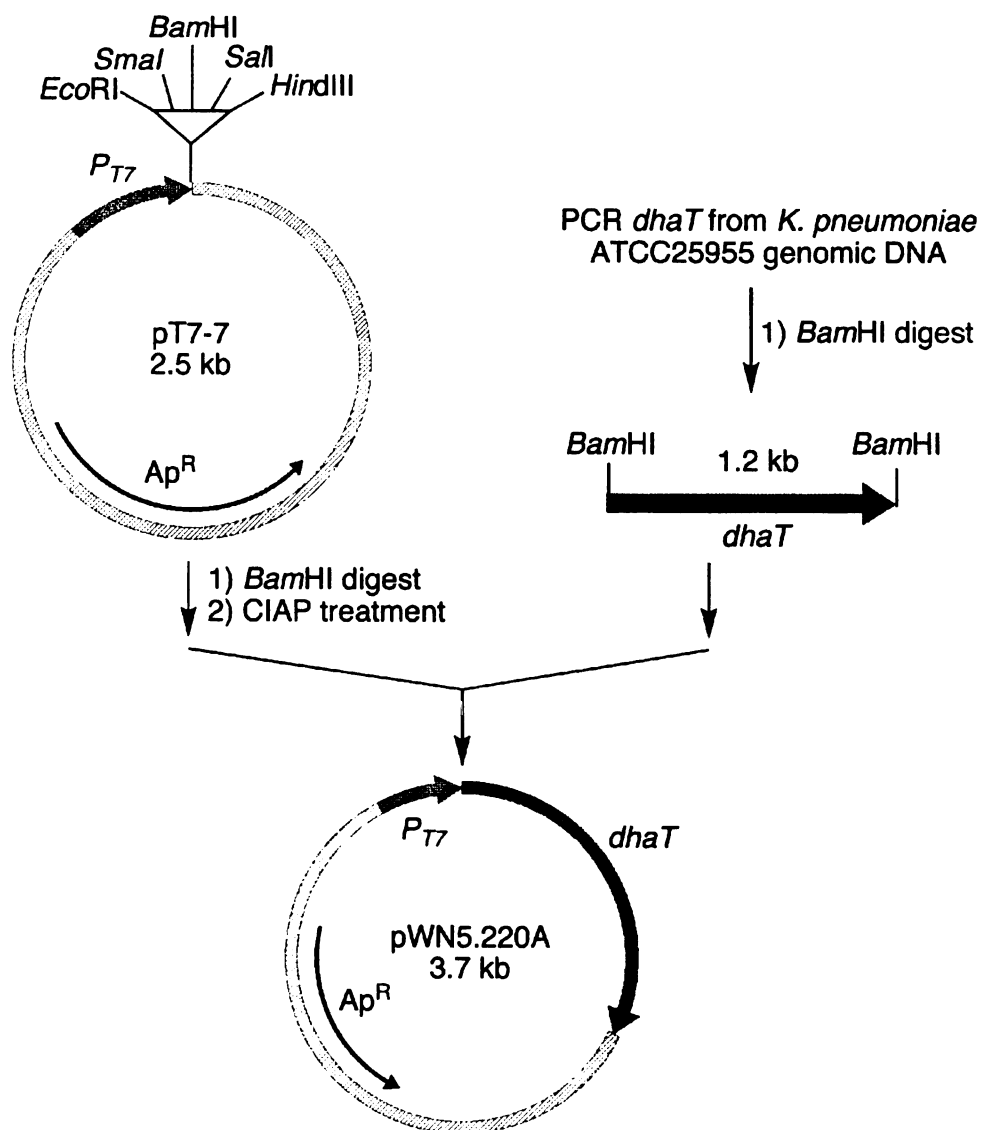


Figure 29. Construction of plasmid pWN5.220A.

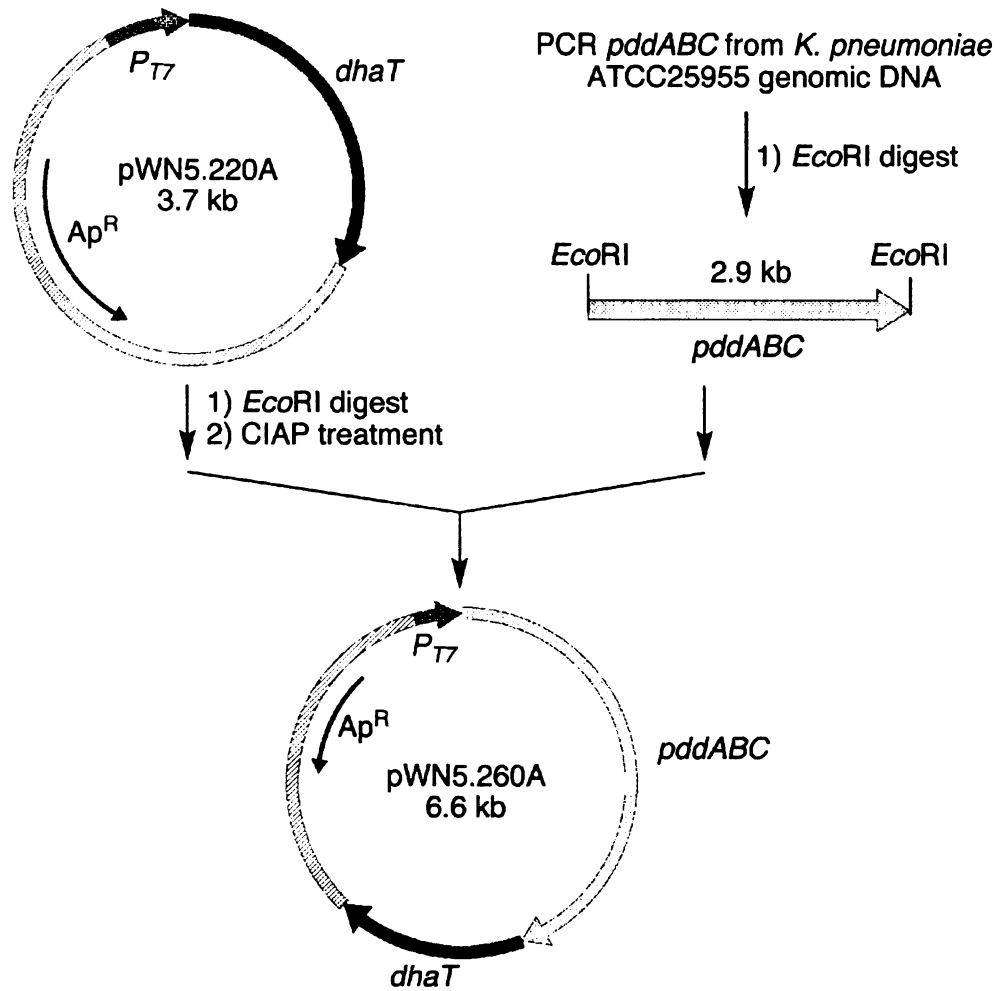


Figure 30. Construcion of plasmid pWN5.260A.

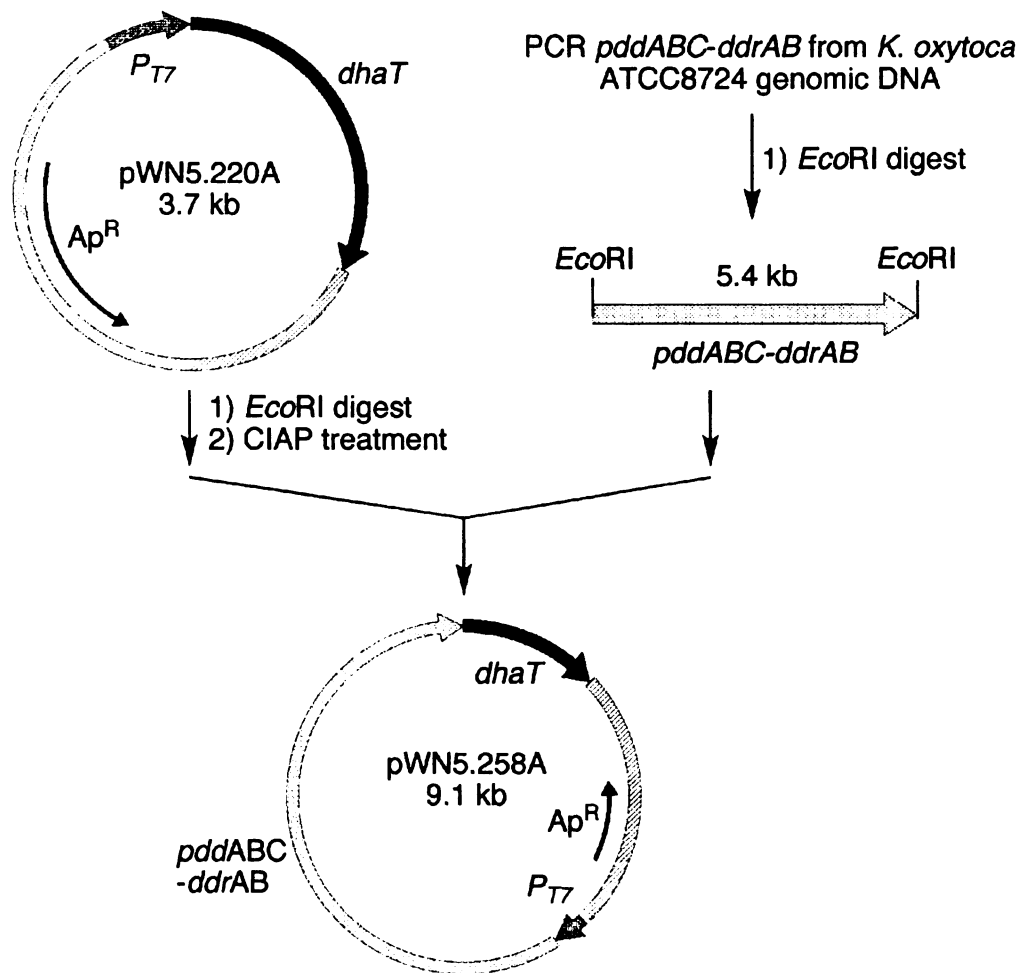


Figure 31. Construction of plasmid pWN5.258A.

The final plasmid pWN5.258A harboring the *K. oxytoca pddABC* encoding diol dehydratase, *ddrAB* encoding reactivating protein, and *K. pneumoniae dhaT* encoding 1,3-propanediol oxidoreductase was derived from the same cloning vector pT7-7. Due to substrate similarity, erythritol might inactivate dehydratases under the same glycerol inactivation mechanism. This glycerol inactivated dehydratase is known to be reactivated in the presence of ATP, Mg^{2+} and the reactivating protein. It was interesting whether diol dehydratase could be revived in the case when erythritol inactivation exists. A 5.35 kb fragment containing *pddABC* and *ddrAB* genes was amplified from the *K. oxytoca* ATCC8724 genomic DNA and subsequently digested with *EcoRI*. Ligation of this fragment into the *EcoRI* site of pWN5.220 led to the formation of plasmid pWN5.258A. Genes *pddABC*, *ddrAB* and *dhaT* in plasmid pWN5.258A were transcribed in the same orientation as the phage T7 promoter.

Enzyme assays for in vitro erythritol dehydratase activity

In vitro dehydratase activity was assayed according to the widely applied literature procedure by Abeles (Figure 32).³⁷ A typical assay mixture was composed of the substrate of interest, yeast alcohol dehydrogenase, NADH and coenzyme B₁₂ in a phosphate buffer at pH 8 in a cuvette. A unit of activity equals 1 μ mol per min of NADH oxidized at 37 °C. The heterologously expressed dehydratases were first examined by assaying their activity using erythritol as substrate. The possible oxidation reaction in the coupled enzyme assay was followed spectrophotometrically by monitored formation of NAD at 340 nm. Enzyme assays of these dehydratases were carried out in crude cell lysate without further purification. The catalytic activity of each enzyme towards its

native substrates glycerol and 1,2-propanediol was also determined. Catalytic activity towards erythritol was detected for *K. pneumoniae* GldABC in crude lysate (Table 6).

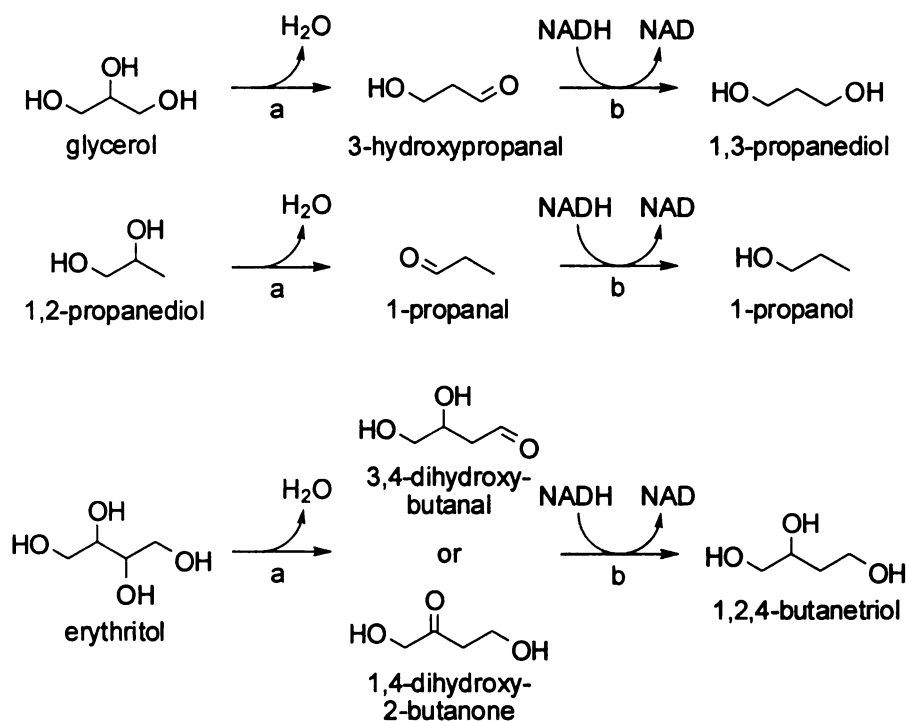


Figure 32. *In vitro* dehydratase assays using different substrates. (a) glycerol or diol dehydratase; (b) yeast alcohol dehydrogenase

Table 6. *In vitro* dehydratase activities.

construct	gene(s)	specific activity (U/mg)	
		native substrate	erythritol
DH5 α /pCS120	<i>C. freundii dhaBCE</i>	no activity ^a	no activity
JM109/pFL1	<i>Cl. Pasteurianum</i> 13.5 kb genomic fragment containing <i>dhaBCE</i>	0.25 ^a	no activity
DH5 α /pXY39	<i>S. typhimurium pduCDE</i>	0.34 ^b	no activity
W3110/pML5.200	<i>K. pneumoniae gldABC</i>	0.61 ^a	0.001
BL21(DE3)/pWN5.260A	<i>K. pneumoniae pddABC-dhaT</i>	1.32 ^b	no activity
BL21(DE3)/pWN5.258A	<i>K. oxytoca pddABC-ddrAB-dhaT</i>	1.45 ^b	no activity

^aglycerol as substrate

^b1,2-propanediol as substrate

1,2,4-Butanetriol biosynthesis using erythritol as substrate

The successful identification of erythritol dehydratase activity using *K. pneumoniae* GldABC makes erythritol a possible starting material for the biosynthesis of 1,2,4-butanetriol. To further establish this biosynthetic route, it was necessary to identify all the enzymes responsible to the biosynthetic conversion. Our goal was to construct a single plasmid capable of synthesizing all the enzymes necessary in the pathway, ultimately leading to a microbe that synthesizes 1,2,4-butanetriol from erythritol. To this end, two plasmids were constructed harboring the *K. pneumoniae dhaT* gene encoding 1,3-propanediol oxidoreductase. A 1.2 kb fragment containing the *dhaT* gene was amplified from the *K. pneumoniae* genomic DNA and subsequently digested with *Hind*III. This fragment was then ligated into the *Hind*III site of pJF118EH to yield plasmid pML5.179 (Figure 33). To obtain a construct that harbors both *gldABC* and *dhaT* genes, the *Hind*III digested *dhaT* fragment was ligated into pML5.200, which had been linearized by *Hind*III to afford plasmid pML5.226 (Figure 34). Transcription of *gldABC* and *dhaT* was initiated by the pJF118EH-localized *tac* promoter. The 1,3-propanediol oxidoreductase activity was determined by a reverse assay proposed by Johnson and Lin.³⁸ The assay mixture contained the substrate of interest and NAD in a bicarbonate buffer at pH 9. A unit of activity is given as μmol per min at 25 °C. *In vitro* oxidoreductase activities of these two plasmids expressed in wild-type W3110 are shown in Table 7. Activities were observed using both 1,3-propanediol and 1,2,4-butanetriol as substrates, indicating that the oxidoreductase activities were expressed heterologously in both constructs. Our control experiment with construct *E. coli* W3110/pML5.200 showed no detectable activity.

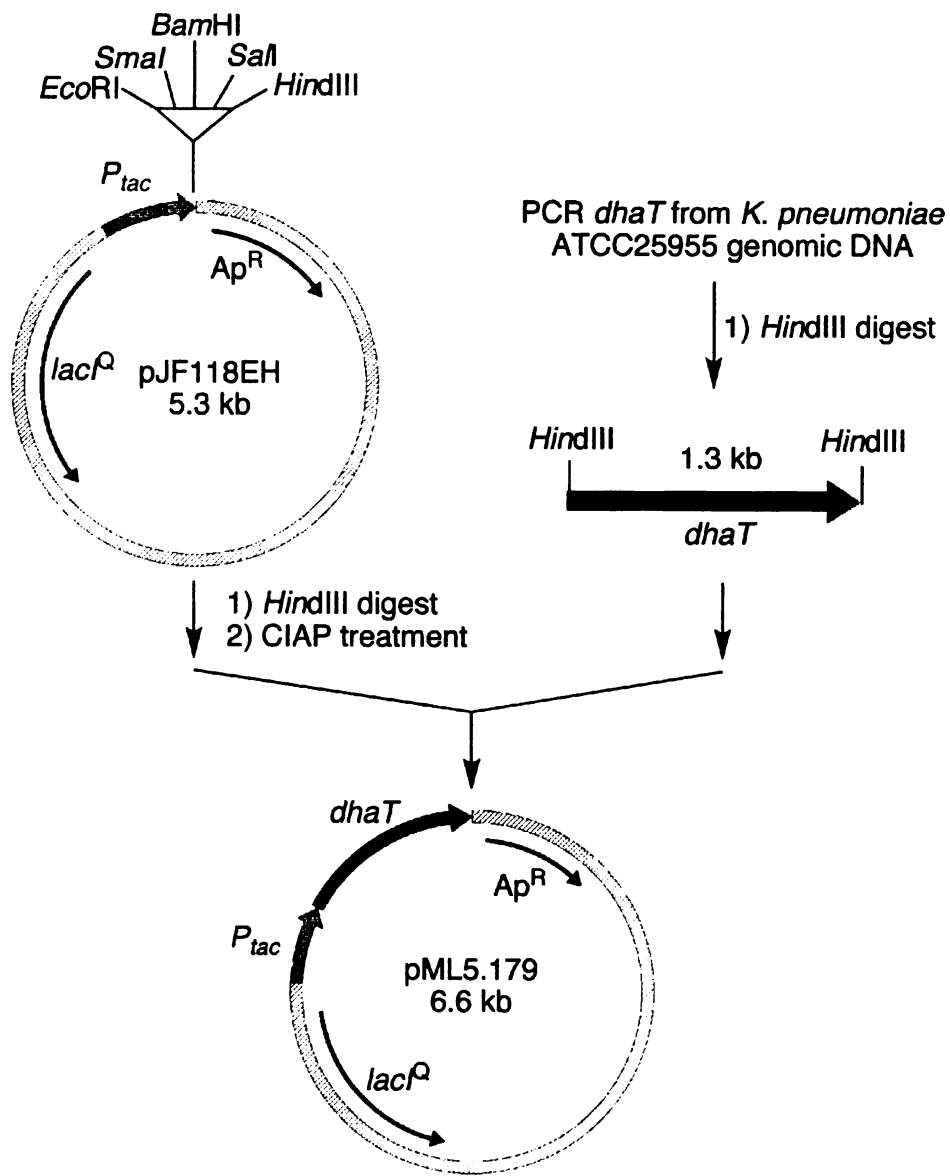


Figure 33. Construction of plasmid pML5.179.

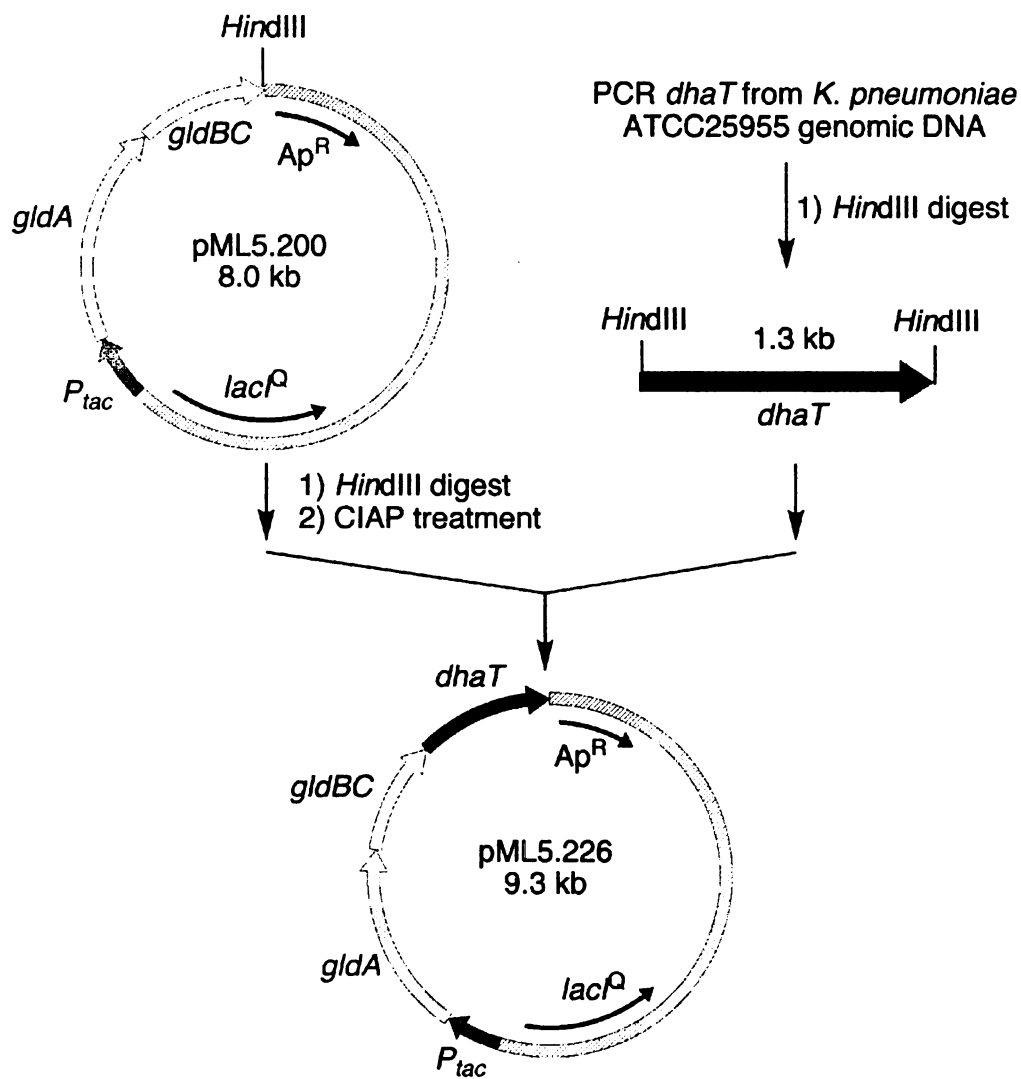


Figure 34. Construction of plasmid pML5.226.

Table 7. 1,3-Propanediol oxidoreductase activities.

construct	gene(s)	specific activity (U/mg)	
		1,3-propanediol	1,2,4-butanetriol
W3110/pML5.179	<i>K. pneumonia dhaT</i>	4.3	0.01
W3110/pML5.200	<i>K. pneumoniae gldABC</i>	no activity	no activity
W3110/pML5.226	<i>K. pneumoniae gldABC-dhaT</i>	1.1	0.003

Previous experiments suggested that *K. pneumoniae gldABC* encoding glycerol dehydratase, while coupled with *dhaT* encoding 1,3-propanediol oxidoreductase, converts erythritol to 1,2,4-butanetriol. As a second line of evidence, *in vitro* enzymatic reactions using erythritol as substrate was setup to verify this finding. Clarified *E. coli* W3110/pML5.179, W3110/pML5.200 and W3110/pML5.226 cell-free lysate were prepared. A total of four reactions were set up under different conditions (Table 8). A typical enzymatic reaction mixture contained 100 mM erythritol and 0.13 mM coenzyme B₁₂ in 40 mM phosphate buffer at pH 8 and was incubated overnight with the cell-free lysate of interest. The first two reactions employed lysate solution obtained from *E. coli*. W3110/pML5.200 as a catalyst. The resulting enzyme reaction mixture was directly subject to sodium borohydride reduction, BSTFA derivatization and analyzed by gas chromatography for the presence of 1,2,4-butanetriol. 1,2,4-butanetriol was detected in the gas chromatogram of reaction 2 with a total yield of 0.004% based on the amount of erythritol in the reaction mixture. No 1,2,4-butanetriol was detected in the control reaction 1 without supplemented coenzyme B₁₂. Reaction 3 employed W3110/pML5.226 lysate as a catalyst while reaction 4 was incubated in W3110/pML5.179 and W3110/pML5.200 lysate mixture in a 1:1 ratio. Both reactions were derivatized using BSTFA and analyzed by gas chromatography directly without prior treatment. No 1,2,4-butanetriol was detected in reaction 3. However, to our surprise, 1,2,4-butanetriol

formation was observed in reaction 4, which might indicate the lack of either dehydratase or oxidoreductase expression in reaction 3. These results provide a proof of concept in using erythritol as substrate for 1,2,4-butanetriol biosynthesis. The discovery that the *K. pneumoniae* *gldABC* gene cluster encodes erythritol dehydratase activity will serve as parent genes for directed enzyme evolution experiments designed to improve this dehydratase activity.

Table 8. *In vitro* enzymatic reactions using erythritol as substrate.

reaction	lysate	reaction conditions ^a	1,2,4-butanetriol
1	W3110/pML5.200	enzyme reaction (no B ₁₂ added) was subject to NaBH ₄ reduction	not detected
2	W3110/pML5.200	enzyme reaction (with B ₁₂) was subject to NaBH ₄ reduction	detected
3	W3110/pML5.226	enzyme reaction with NADH (0.58 mM)	not detected
4	W3110/pML5.200 and W3110/pML5.179 (1:1)	enzyme reaction with NADH (0.58 mM)	detected

^aAll enzymatic reactions were performed in potassium phosphate buffer (40 mM pH 8), coenzyme B₁₂ (0.13 mM) and erythritol (100 mM) at rt.

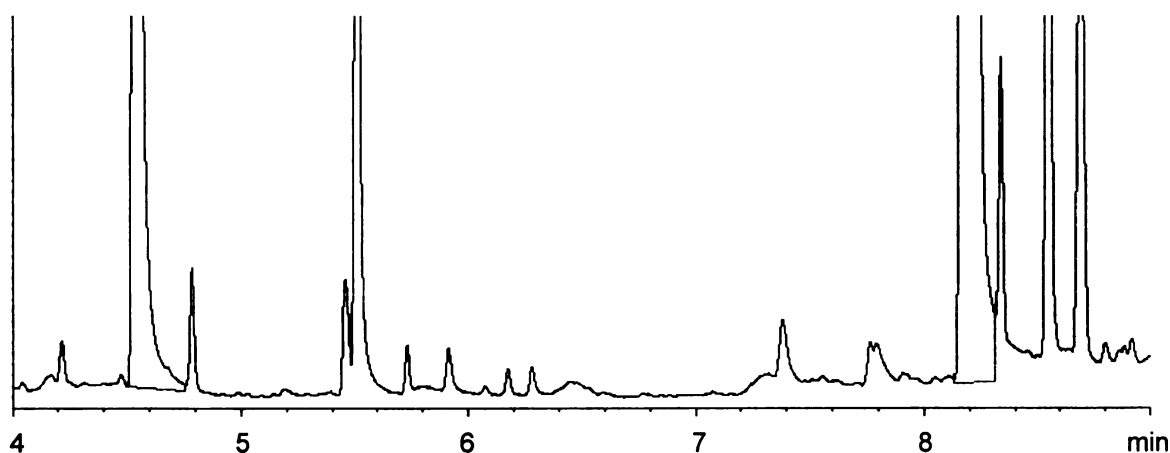


Figure 35. GC chromatogram for reaction 1 (control).

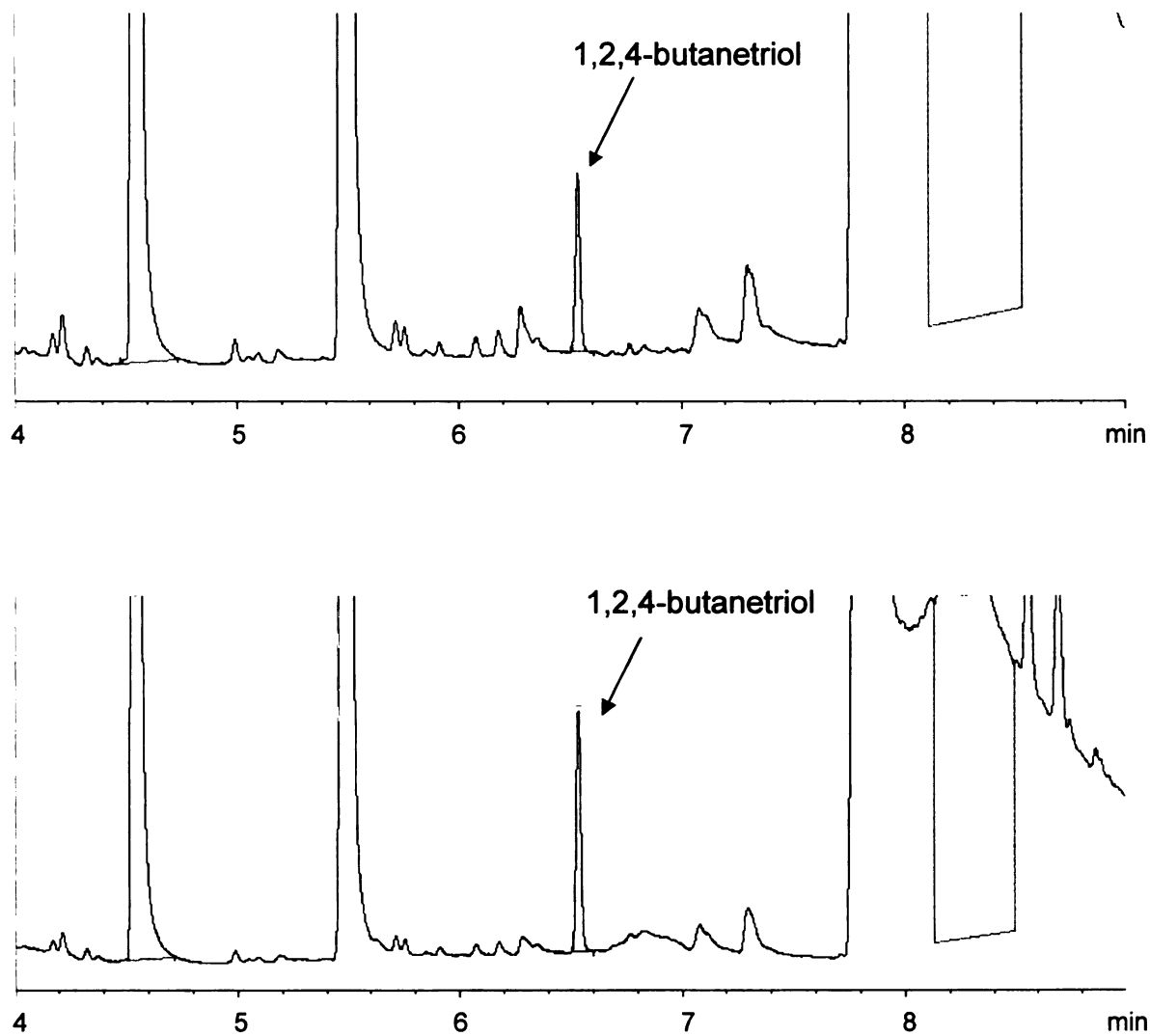


Figure 36. GC chromatograms for reaction 2. (upper) reaction 2; (lower) reaction 2 and authentic 1,2,4-butanetriol co-injection.

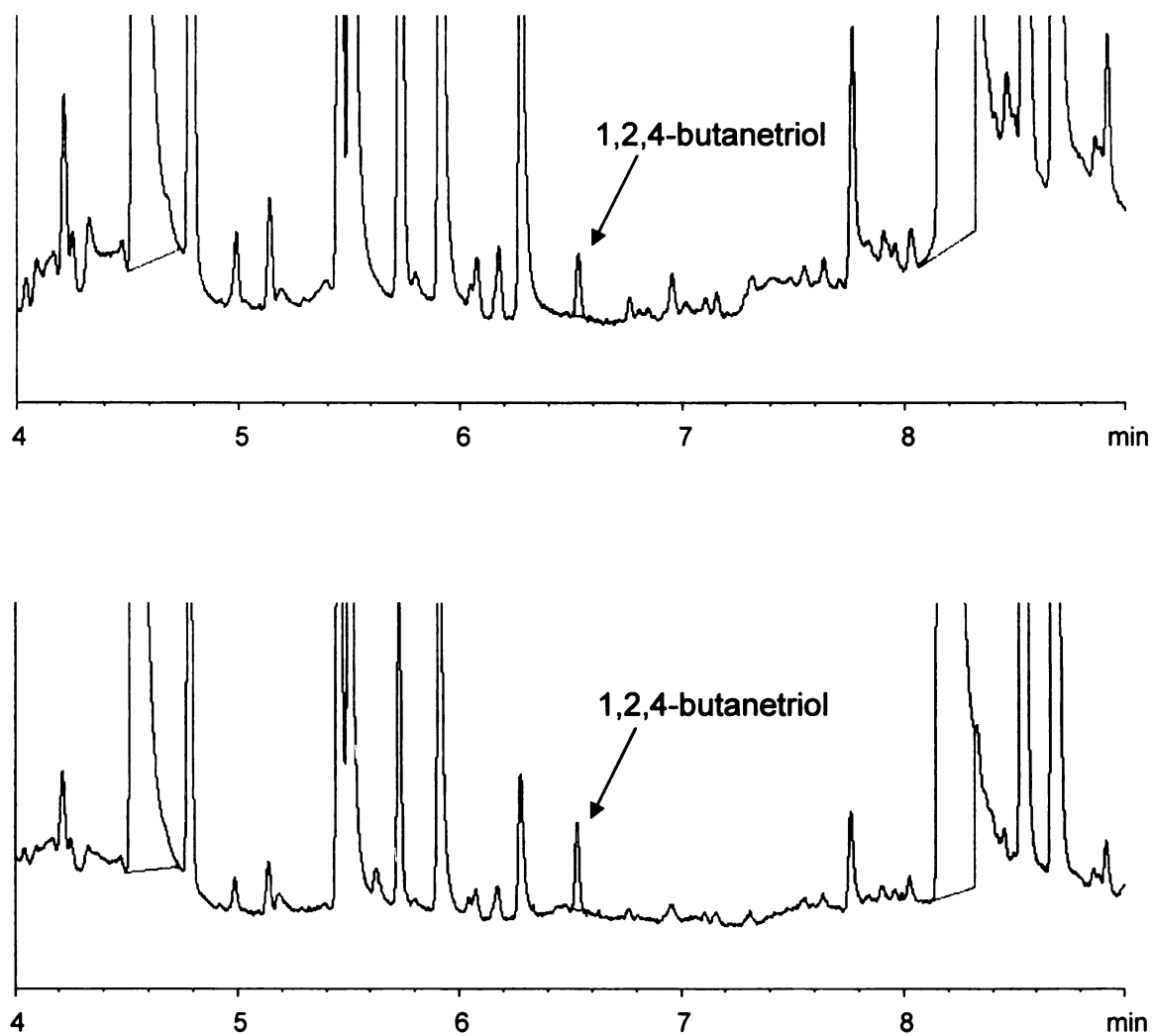


Figure 37. GC chromatograms for reaction 4. (upper) reaction 4; (lower) reaction 4 and authentic 1,2,4-butanetriol co-injection.

In search of 3-deoxy-*glycero*-pentulosonate decarboxylase genes for the microbial synthesis of 1,2,4-butanetriol from pentoses

Background

Earlier research efforts in the Frost group elucidated two parallel biocatalytic pathways that start from D-xylose and L-arabinose to give D- and L-1,2,4-butanetriol, respectively.³⁹ Each biocatalytic system involves two microbes that synthesize all four enzymes required for the conversion. *Pseudomonas fragi* ATCC4973 oxidizes the pentoses into the corresponding carboxylic acids. Recombinant *E. coli* DH5 α /pWN6.186A and BL21(DE3)/pWN6.222A catalyze the conversion of D-xylonic acid to D-1,2,4-butanetriol and L-arabinonic acid to L-1,2,4-butanetriol, respectively.

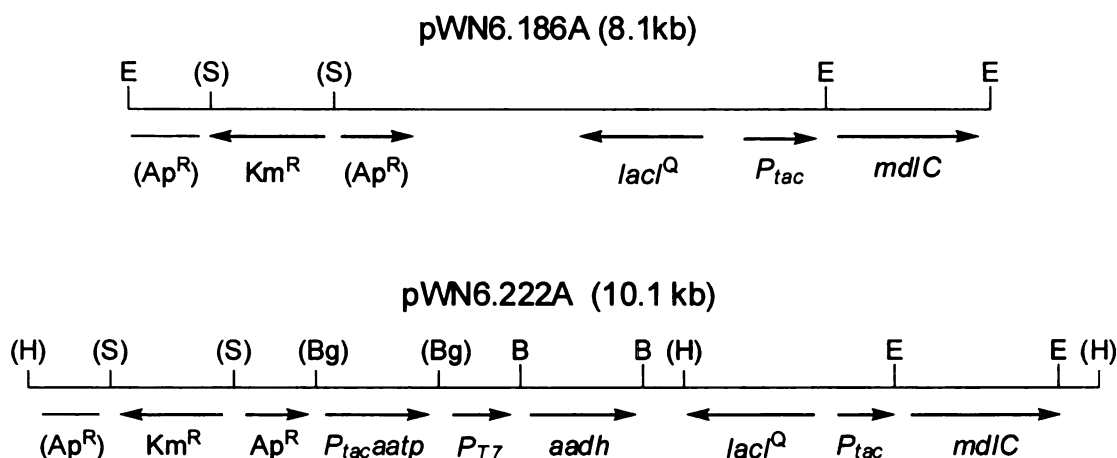


Figure 38. Restriction enzyme map of plasmids. Sites are abbreviated as follow: E = *EcoRI*; S = *ScaI*; H = *HindIII*; Bg = *BglII*; B = *BamHI*.

Although using whole-cell recombinant microbes to produce 1,2,4-butanetriol has proven to be the most efficient approach so far, *E. coli* constructs DH5 α /pWN6.186A and BL21(DE3)/pWN6.222A produced 1,2,4-butanetriol enantiomers in unsatisfactory yields and concentrations. The problem is believed to lie in the decarboxylation step.^{35,36} To

tackle this issue, gene candidates encoding decarboxylase activity towards 3-deoxy-*glycero*-pentulosonic acid were identified using bioinformatics tools. Cloning and enzyme activity assays of the resulting recombinant proteins followed.

Pseudomonas putida ATCC12633 benzoylformate decarboxylase (EC 4.1.1.7) is a thiamin-diphosphate dependent enzyme that catalyzes the non-oxidative decarboxylation of benzoylformate to form benzaldehyde and carbon dioxide. The genes for benzoylformate decarboxylase, (*S*)-mandelate dehydrogenase and mandelate racemase are arranged in an operon (*mdlCBA*), while the genes for the benzaldehyde dehydrogenases are independently transcribed (*mdlDE*). The enzymes in the mandelate pathway allow both enantiomers of mandelic acid to be utilized as a sole carbon source in pseudomonads.⁴⁰ (*R*)-Mandelic acid is first converted into (*S*)-mandelic acid, which is then oxidized to its keto-acid. Benzoylformic acid is decarboxylated to form benzaldehyde and finally degraded into benzoic acid (Figure 39). Benzoic acid is subsequently metabolized by the β -ketoadipate pathway and citric acid cycle. The mechanism of the *mdlC* gene encoding benzoylformate decarboxylase catalyzed reaction has been studied in detail (Figure 40) and the crystal structure of this enzyme is available.⁴¹ In the initial step, the imino group of thiamin-diphosphate abstracts a proton from the thiazolium ring, resulting in the formation of an ylide. The ylide attacks the carbonyl of the substrate to form the tetrahedral intermediate 2- α -mandelyl complex of thiamin-diphosphate. Decarboxylation of this complex results in the enamine intermediate. Protonation of the enamine provides another tetrahedral 2-hydroxybenzyl-thiamin complex. Finally, benzaldehyde is eliminated from the enzyme active site (Figure 40).

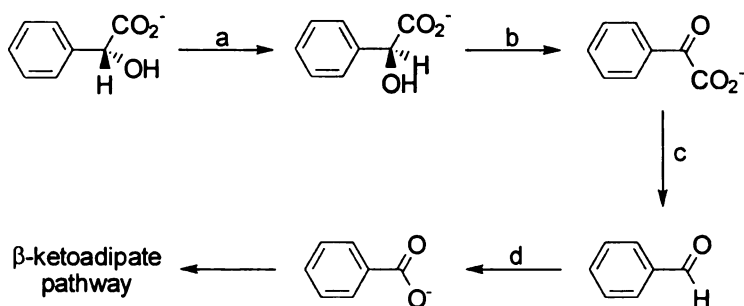


Figure 39. Mandelate pathway in *P. putida*. (a) mandelate racemase (*mdlA*); (b) (*S*)-mandelate dehydrogenase (*mdlB*); (c) benzoylformate decarboxylase (*mdlC*); (d) NAD and NADP benzaldehyde dehydrogenase (*mdlD*, *mdlE*)

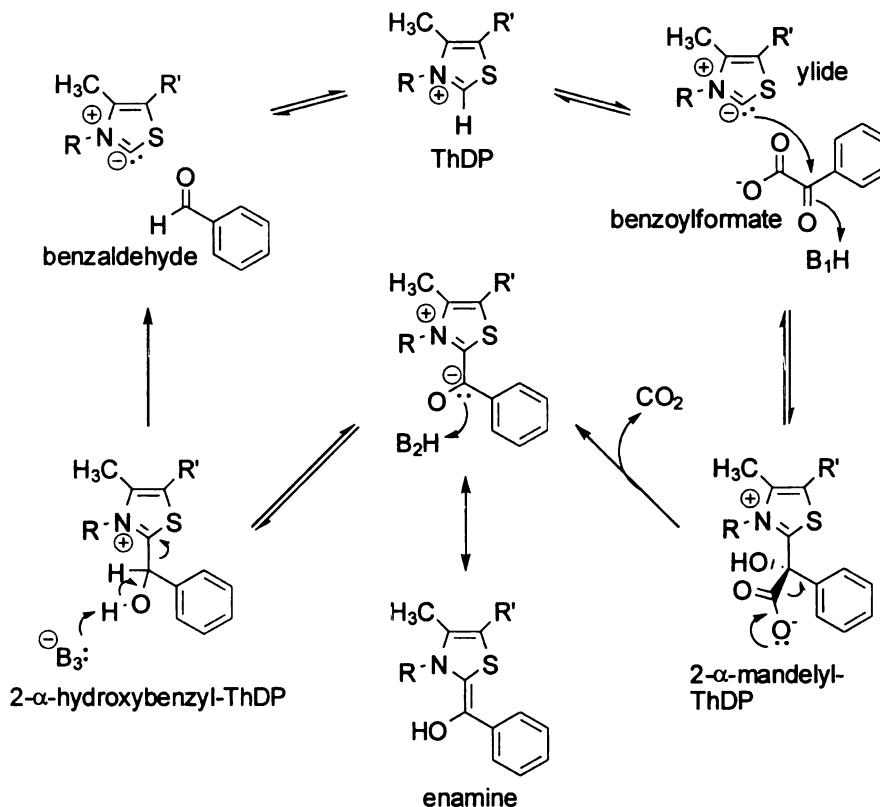


Figure 40. Reaction mechanism for the benzoylformate decarboxylase-catalyzed decarboxylation of benzoylformate.⁴² R = 4'-amino-2-methyl-5-pyrimidyl; R' = β -hydroxyethylidiphosphate.

Historically, the first attribute established about an enzyme is the reaction that it catalyzed. A particular enzyme is either cloned in another microorganism or overexpressed in its native host. Crude extract is then purified and subjected to activity assay using the substrate of interest. The purified enzyme can be used to deduce the amino acid sequence, and subsequently the gene sequence is obtained. The discovery of 3-deoxy-*glycero*-pentulosonate decarboxylase activity for 1,2,4-butanetriol biosynthesis using *P. putida mdhC* encoding benzoylformate decarboxylase represents an example. This enzyme was obtained after screening for decarboxylase activities through more than 10 independent decarboxylase candidates.³⁹ Since *P. putida* benzoylformate decarboxylase is well studied, the gene sequence was therefore obtained simply by accessing a gene database.^{40,41} In recent years, a different path has been established which follows almost the reverse direction of the conventional route. By now moving from sequence to activity, there is a higher chance of success with fewer resources, including less time investment for more hits. Gene sequences, which are often publicly available, can be accessed by computers. Bioinformatics, which is a discipline combining biology and information technology, is the essential tool to process these enormous amounts of information. Gene-mining among the available genome sequences and comparing gene sequence similarity can be used to identify suitable or interesting open reading frames for different functions.⁴³

In general, web-based bioinformatics tools were used using links provided through web-sites such as National Center for Biotechnology Information (NCBI),⁴⁴ European Bioinformatics Institute (EBI)⁴⁵ and the ERGO Bioinformatics Suite provided by Integrated Genomics, Inc.⁴⁶ Using BLAST, one can search either for similar proteins

within a database or for new genes similar to the protein sequence at hand.⁴⁷ The results are given in a graphic display with the scores in different colors, which is then deciphered into a hit list. Multiple sequence alignment between sequences of interest can be done using ClustalW.⁴⁸ It is particularly useful to compare whole protein/DNA sequences or domains. Even in the area of structural information, long regarded as one with access for experts only, there are handy molecular visualization softwares like PyMOL available to help.⁴⁹ PyMOL reads in molecular coordinates of atoms in the protein structure from a .pdb file and these files can be found at the Protein Data Bank.⁵⁰ With the ability to visualize an enzyme's three-dimensional structure, rational protein design provides an alternative avenue toward an enzyme mutant with improved properties.

Cloning and Screening for 3-deoxy-glycero-pentulosonate decarboxylase candidates

Eight genes bearing high sequence identity with *P. putida mdhC* gene were selected using sequence alignment algorithms such as BLAST (Table 9). The strains or genomic DNA were obtained mainly from ATCC and revived according to the enclosed instructions. Genomic DNA was extracted from each of the candidates using the standard cetyltrimethylammonium bromide (CTAB) protocol. PCR reactions using *PfuTurbo* DNA polymerase (Stratagene) were employed to amplify the target decarboxylase genes and the plasmid pJF118EH was used as the cloning vector. Both the vector and the insert were digested with the suitable restriction enzymes and ligated. The resulting recombinant plasmids were transformed into *E. coli* BL21 expression host strain and grown to evaluate enzyme activities (Figure 41-26).

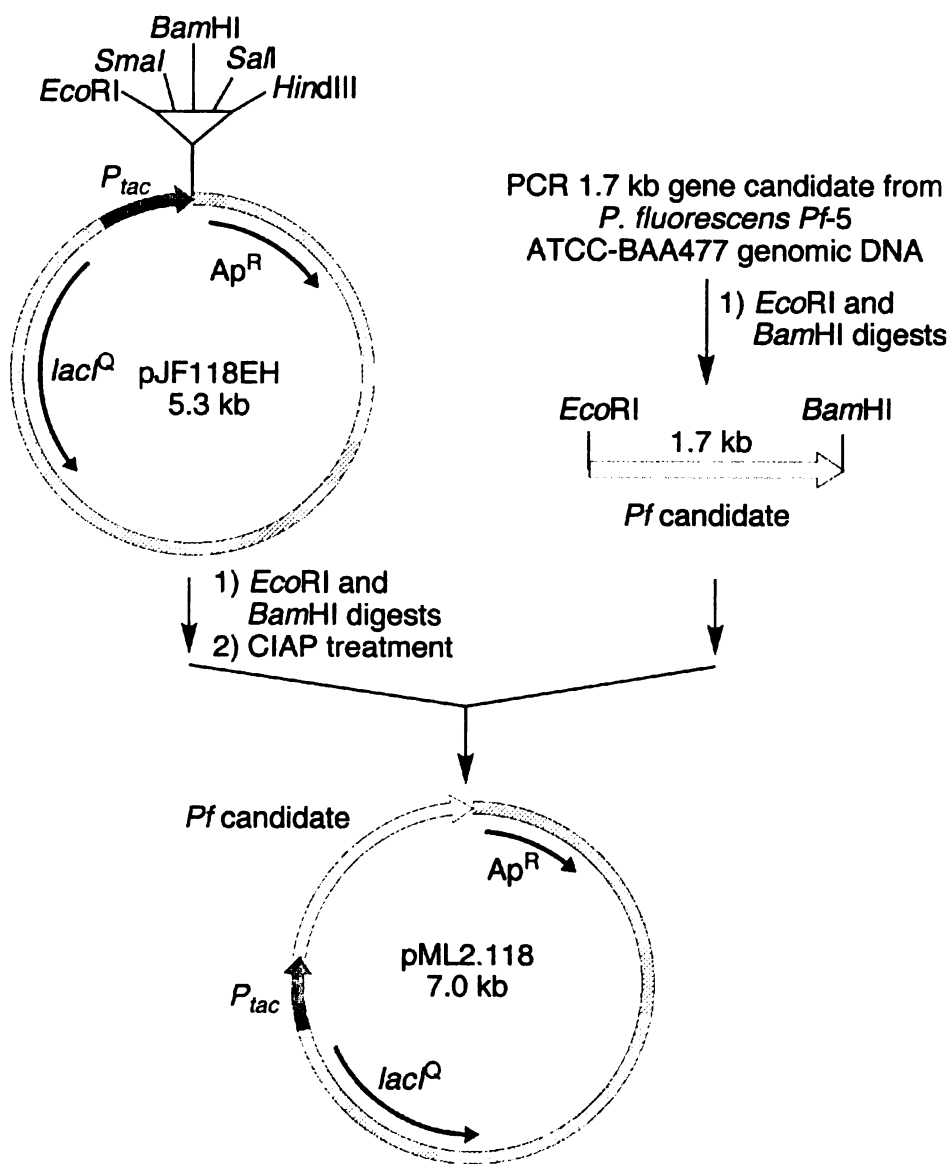


Figure 41. Construction of plasmid pML2.118.

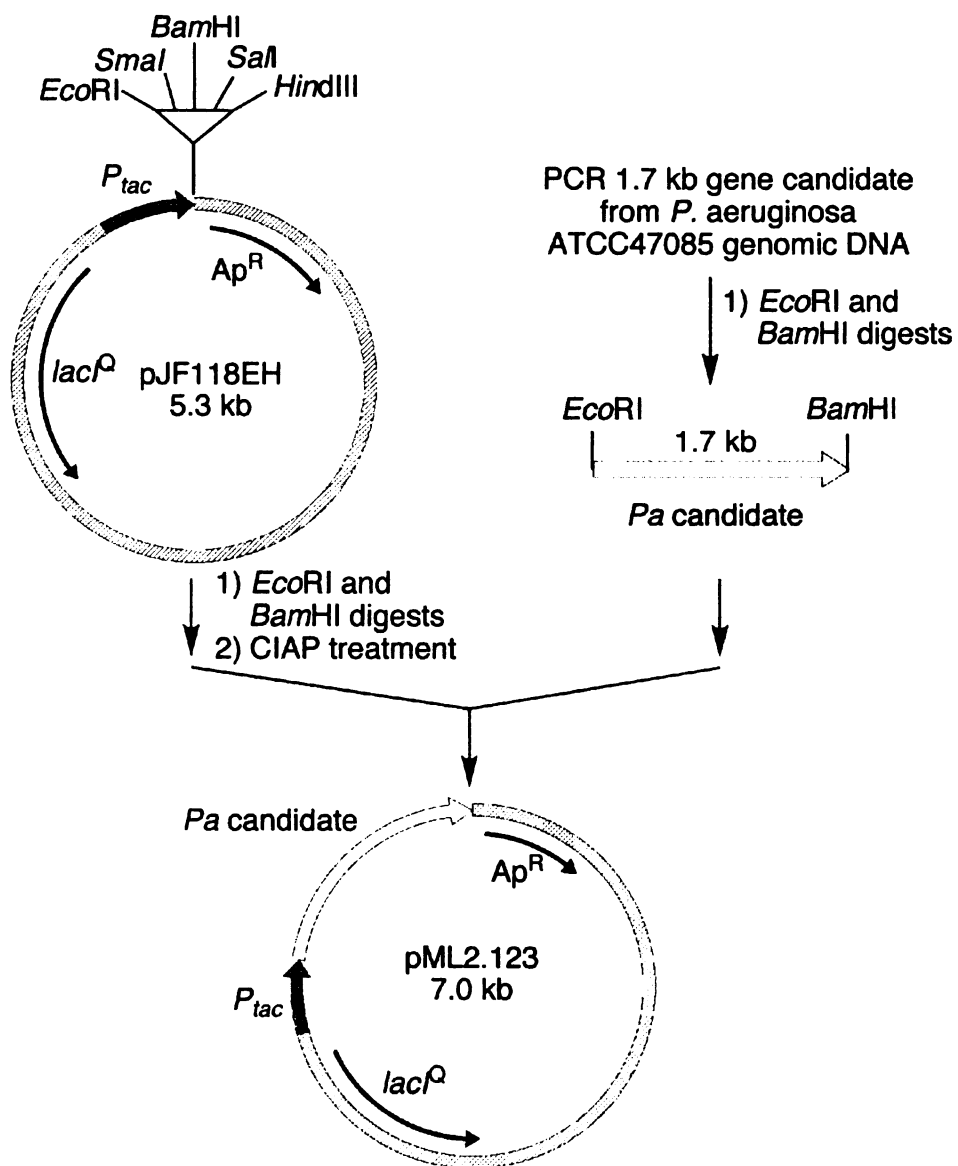


Figure 42. Construction of plasmid pML2.123.

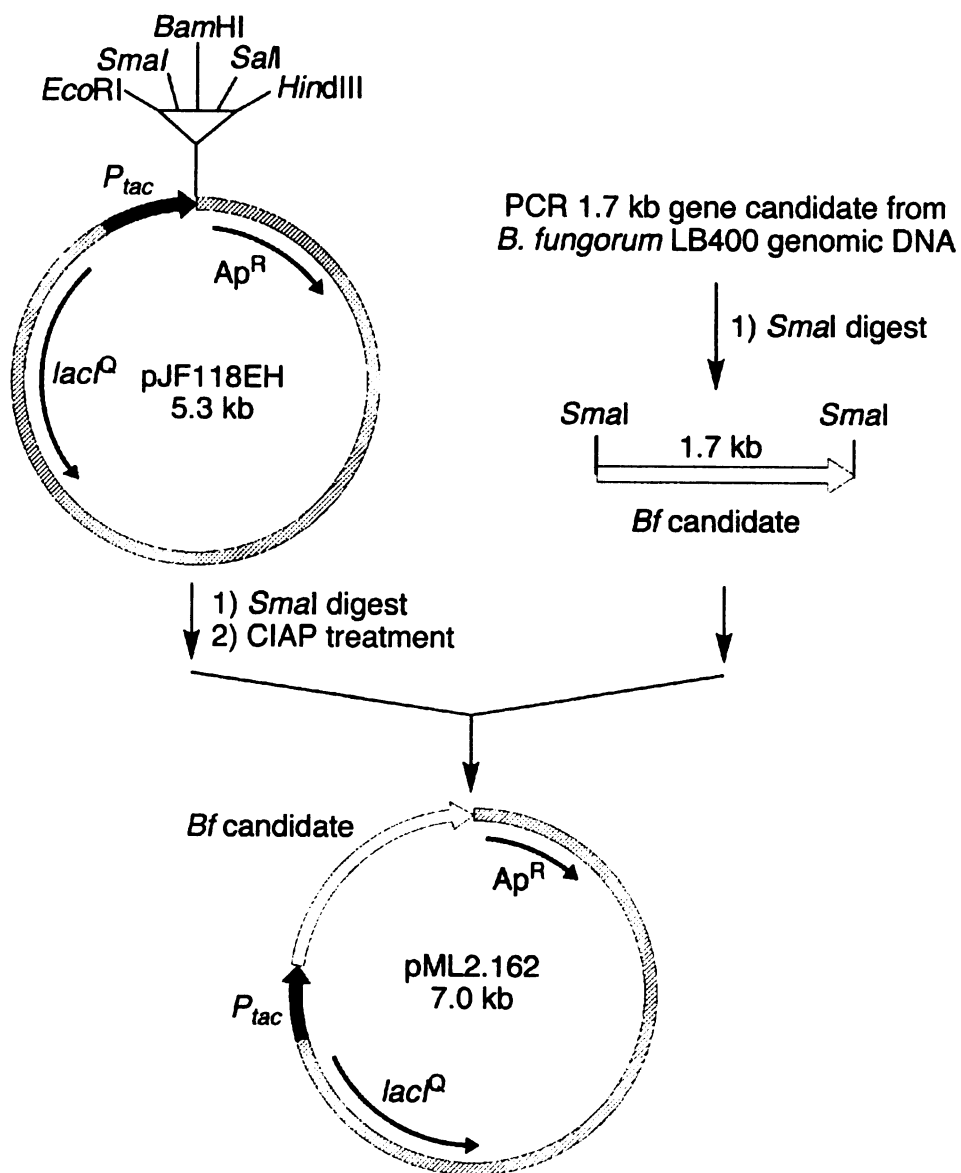


Figure 43. Construction of plasmid pML2.162.

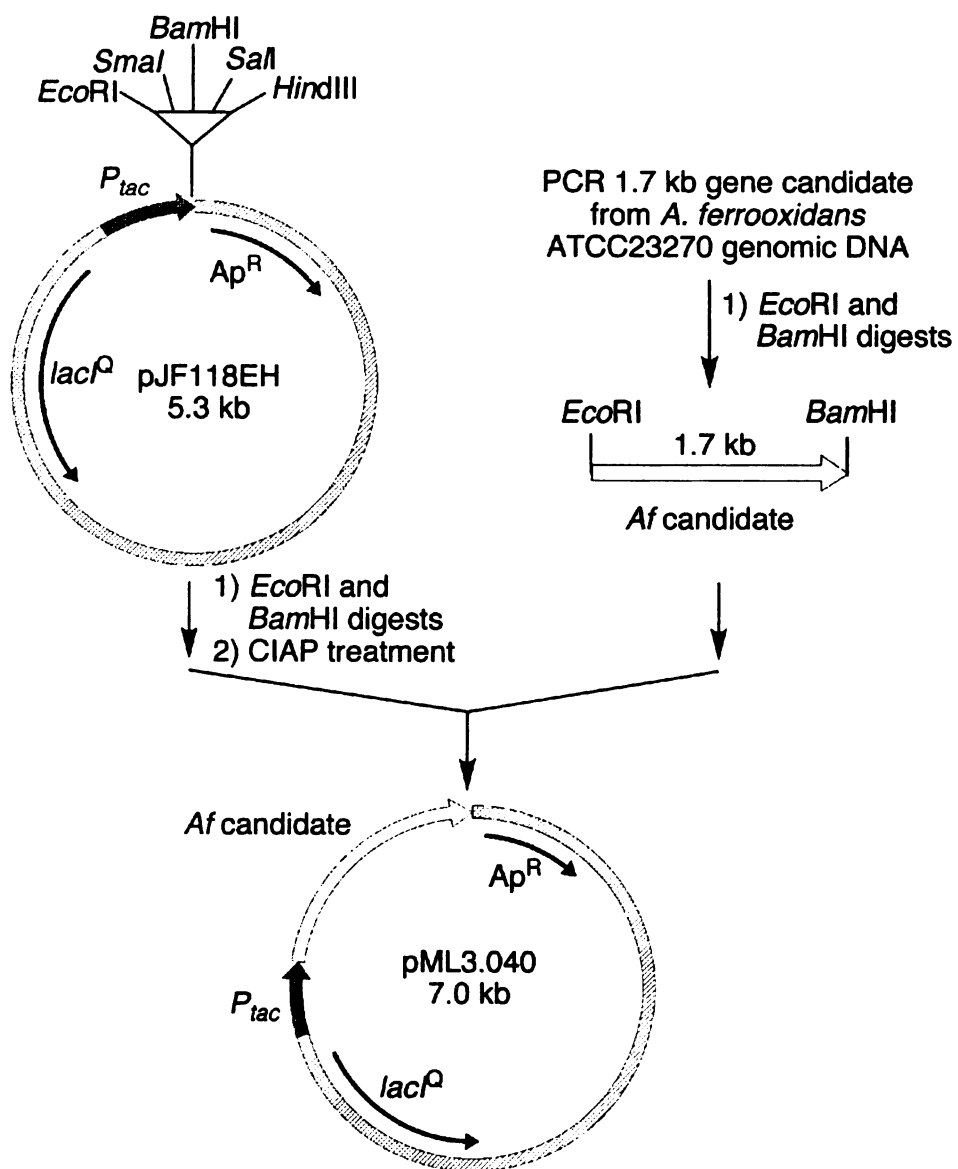


Figure 44. Construction of plasmid pML3.040.

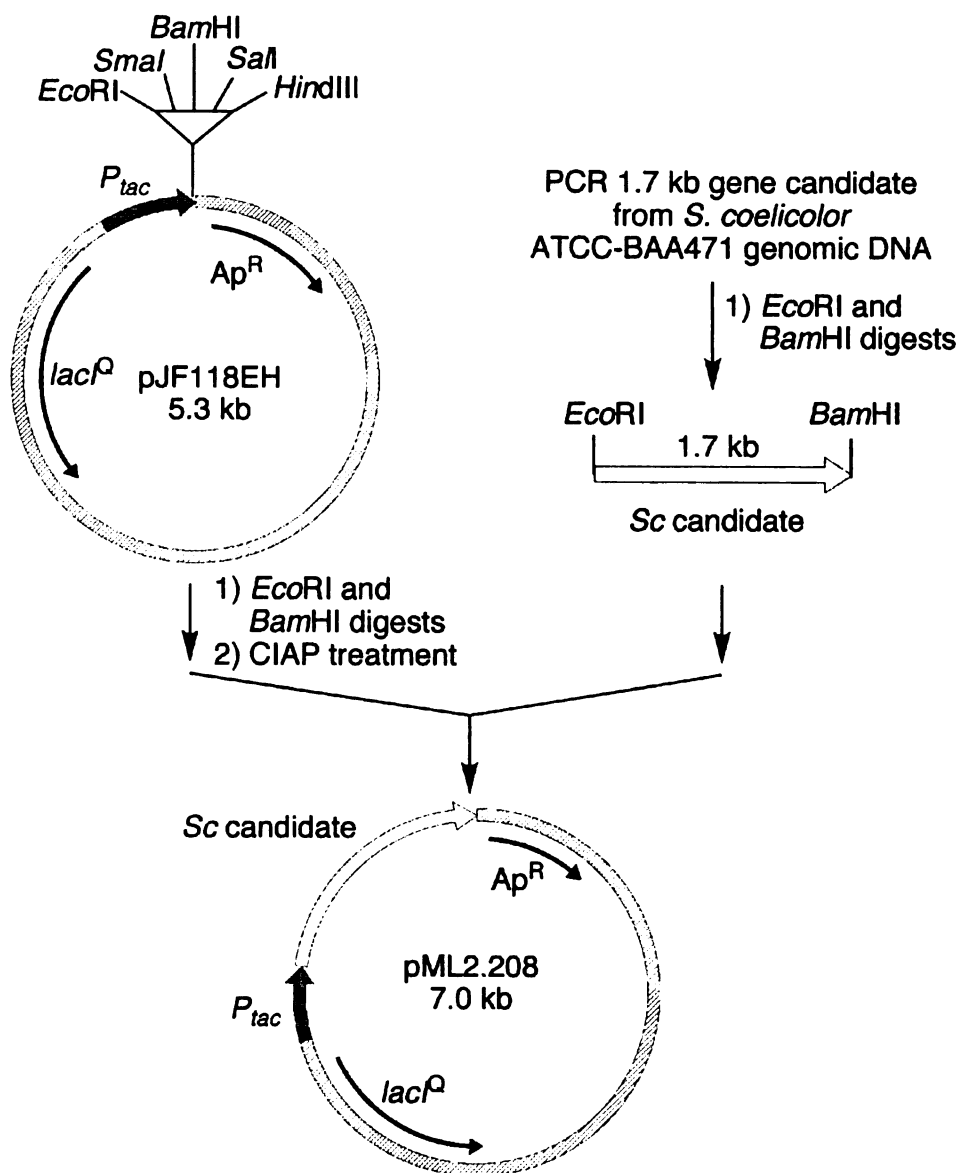


Figure 45. Construction of plasmid pML2.208.

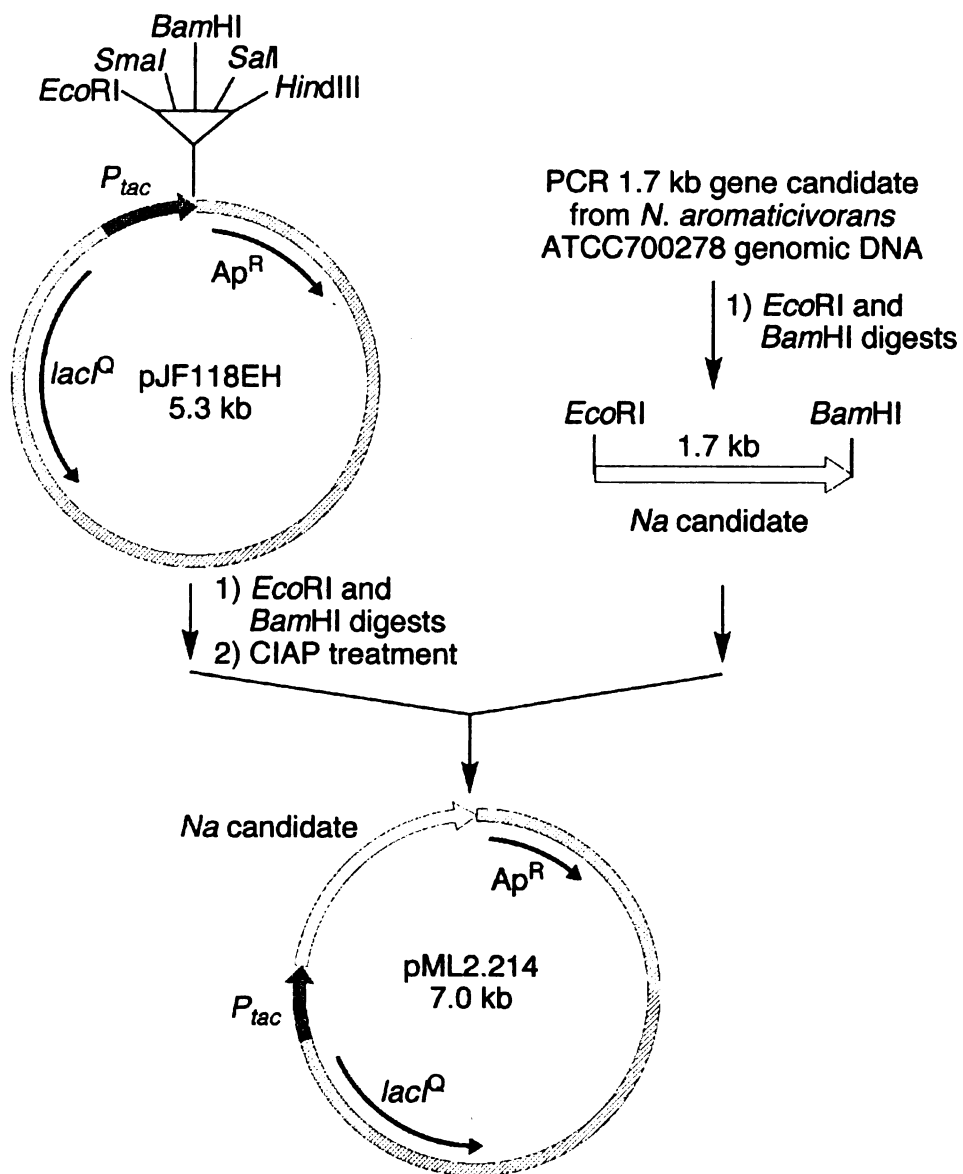


Figure 46. Construction of plasmid pML2.214.

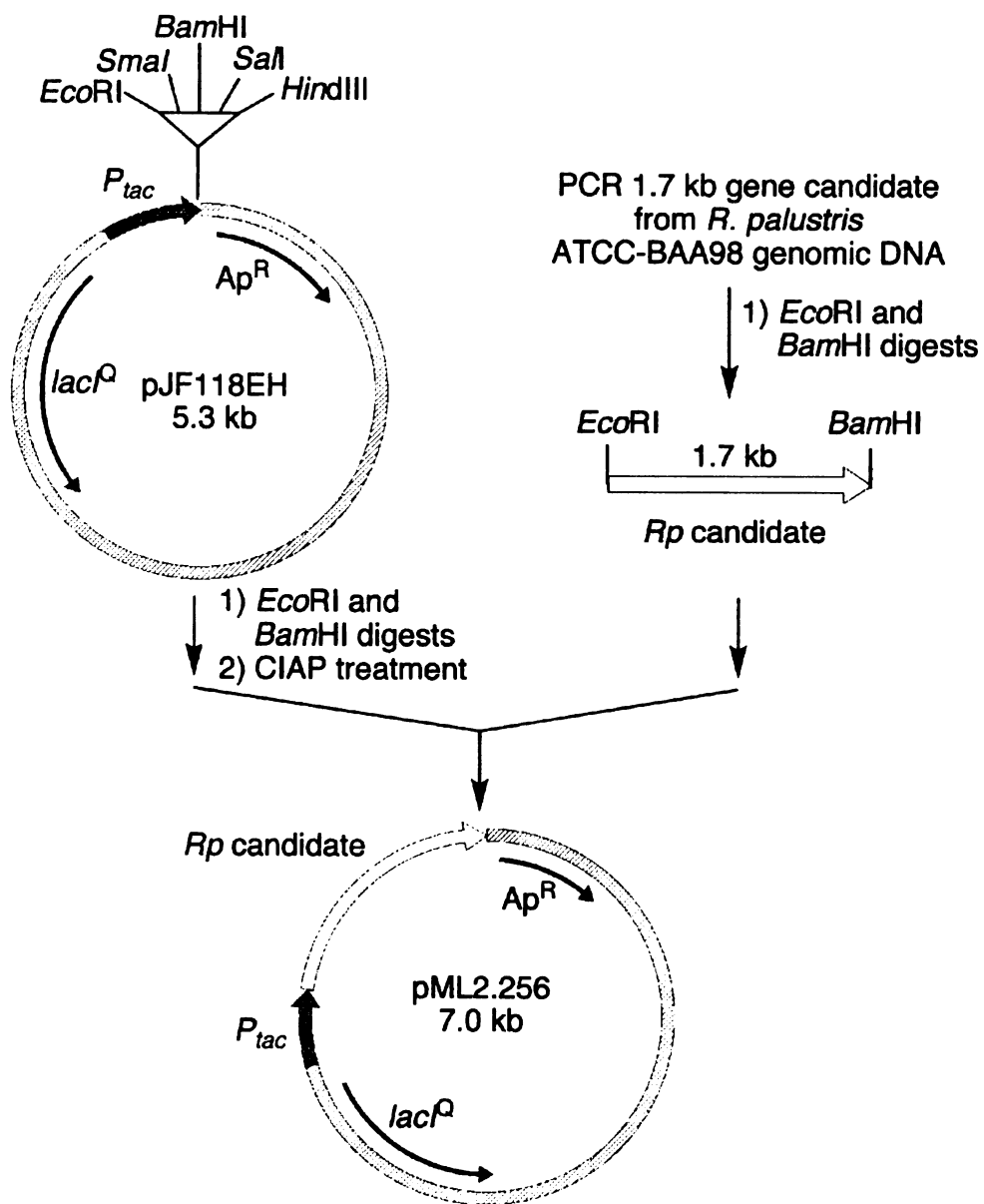


Figure 47. Construction of plasmid pML2.256.

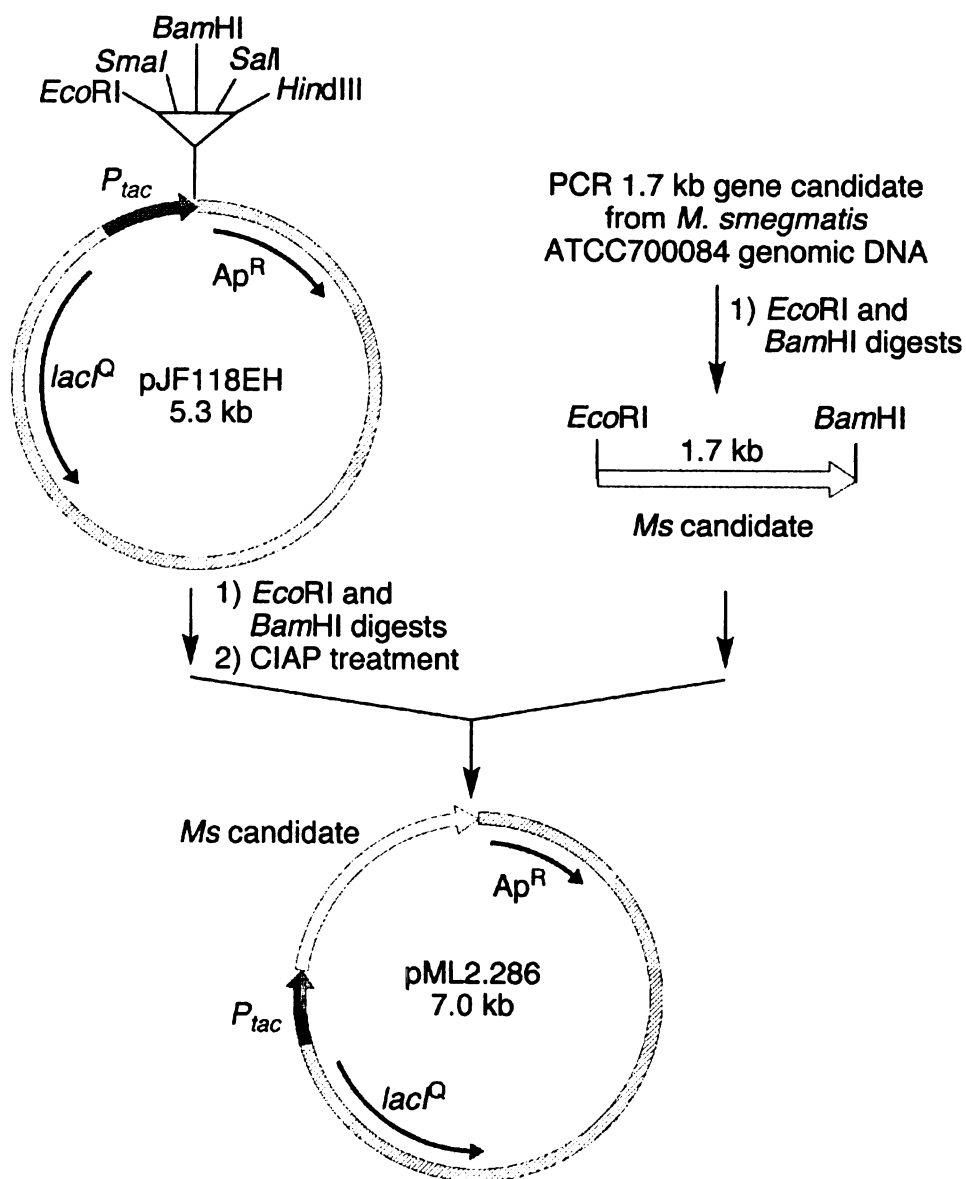


Figure 48. Construction of plasmid pML2.286.

Table 9. 3-Deoxy-glycero-pentulosonate decarboxylase candidates.

plasmid	strain	Identity to <i>P. putida mdIC</i> (%)
pWN5.238	<i>Pseudomonas putida</i> ATCC12633	100
pML2.118	<i>Pseudomonas fluorescens</i> ATCC-BAA477	60
pML2.123	<i>Pseudomonas aeruginosa</i> ATCC47085	59
pML2.162	<i>Burkholderia fungorum</i> LB400	61
pML3.040	<i>Acidithiobacillus ferrooxidans</i> ATCC23270	37
pML2.208	<i>Streptomyces coelicolor</i> ATCC-BAA471	49
pML2.214	<i>Novosphingobium aromaticivorans</i> ATCC700278	37
pML2.256	<i>Rhodopseudomonas palustris</i> ATCC-BAA98	40
pML2.286	<i>Mycobacterium smegmatis</i> ATCC700084	42

Site-directed Mutagenesis of benzoylformate and pyruvate decarboxylase gene

Benzoylformate decarboxylase from *P. putida* and pyruvate decarboxylase from *Z. mobilis* catalyze the decarboxylation of aromatic benzoylformic acid and pyruvic acid, respectively. In fact, benzoylformate decarboxylase poorly accepts D- and L-3-deoxy-glycero-pentulosonic acid as substrate. In a 2004 report, Kenyon and McLeish identified some of the amino acid residues potentially involved in substrate binding and thus in the specificity of these enzymes.⁵¹ Among various mutants involved in their study, benzoylformate decarboxylase mutant A460I (BFD A460I) and pyruvate decarboxylase mutant I472A (PDC I472A) are of particular interest and was generated. Both mutants were successful at increasing the catalytic activities in using C4 – C6 keto-acids as substrates by manipulating a single amino acid residue inside the active site. In addition to the two mutants mentioned above, pyruvate decarboxylase amino acid residue I472 were exchanged for glycine, serine and threonine in attempts to either open up the binding pocket (PDC I472G) or to increase the interactions between the protein and the hydroxyl groups on the 3-deoxy-glycero-pentulosonic acid substrate (PDC I472S and I472T) (Table

10). These mutants were prepared using *PfuTurbo* DNA polymerase and the QuikChange site-directed mutagenesis kit (Stratagene, LaJolla, CA) (Figure 49).⁵² Wild-type *P. putida mdlC* encoding benzoylformate decarboxylase and *Z. mobilis pdc* encoding pyruvate decarboxylase were available in the Frost group strain collection as plasmids pWN5.238A and pLOI276⁵³, respectively. These plasmids also served as templates for site-directed mutagenesis experiments without further genetic modifications. The forward primers used for the mutagenesis are shown in Table 10.

The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type. The QuikChange methodology (Figure 49) was performed using *PfuTurbo* DNA polymerase and a temperature cycler. The oligonucleotide primers containing the desired mutation, each complementary to opposite strands of the double-stranded DNA (dsDNA) plasmid template harboring the decarboxylase gene were extended during temperature cycling by the DNA polymerase. Incorporation of the oligonucleotide primers generated a mutated plasmid containing staggered nicks. Following the temperature cycling, the reaction mixture was treated with *DpnI*, which digested the parental methylated DNA template. The nicked plasmid DNA containing the desired mutation was then transformed into *E. coli* BL21. The transformants were plated on LB containing ampicillin and incubated at 37 °C overnight. Plasmid was purified from two single colonies and the mutation was verified by sequencing.

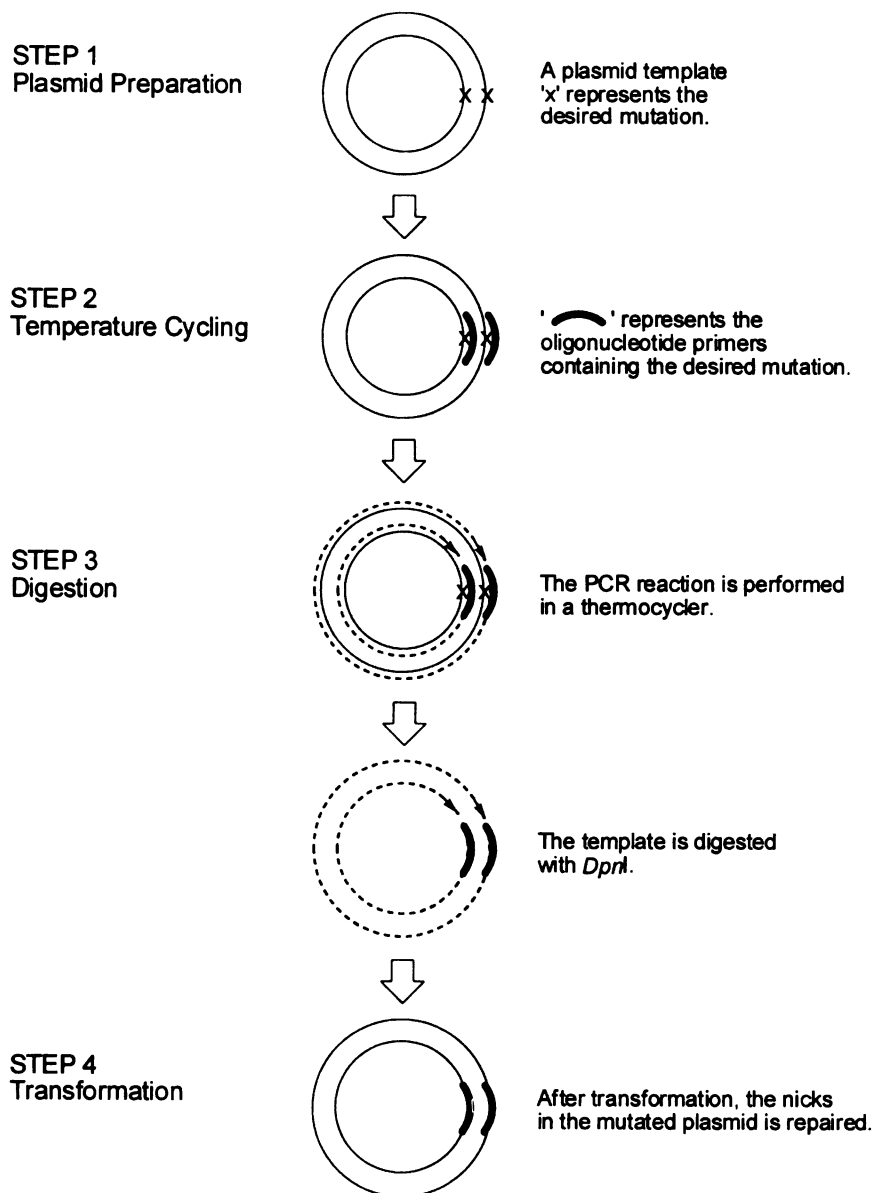


Figure 49. QuikChange Site directed mutagenesis.⁵²

Table 10. Forward primers used to generate mutations.

plasmid	mutation	forward primers
benzoylformate decarboxylase		
pWVN5.238A	wt ^a	
pML3.054	A460I	5'-ATGAACAACGGC <u>ACg</u> TACGGTatcTTGCGATGG
pyruvate decarboxylase		
pLOI276 ⁵³	wt ^b	
pML3.062	I472A	5'-GATCAATAACTATGGTTACACC <u>gC</u> CGAAGTTATGATCCATGATG
pML3.084	I472S	5'-TCAATAACTATGGTTACACCA <u>gC</u> CGAAGTTATGATCCATGATG
pML3.086	I472G	5'-GATCAATAACTATGGTTACACC <u>ggC</u> CGAAGTTATGATCCATGATG
pML3.104	I472T	5'-TCAATAACTATGGTTACACCA <u>AcC</u> CGAAGTTATGATCCATGATG

^awild-type *P. putida mdhC* encoding benzoylformate decarboxylase

^bwild-type *Z. mobilis pdc* encoding pyruvate decarboxylase

Enzyme activities for 3-deoxy-glycero-pentulosonate decarboxylase candidates

The catalytic activity of the eight wild-type candidates generated using bioinformatics and the five benzoylformate and pyruvate decarboxylase site-directed mutants were first examined by coupling the potential decarboxylation reaction with equine liver alcohol dehydrogenase-catalyzed reduction of 3,4-dihydroxybutanal (Figure 50). Racemic 3-deoxy-*glycero*-pentulosonic acid was synthesized using a method described earlier.³⁵ The reactions were followed by monitoring the consumption of NADH at 340 nm using a UV-vis spectrophotometer. Enzyme assays were carried out in the cell lysate of *E. coli* strains that expressed the enzyme candidates (Table 11). As a second line of evidence, each 3-deoxy-*glycero*-pentulosonate decarboxylase candidates was incubated with a reaction mixture of racemic 3-deoxy-*glycero*-pentulosonic acid, thiamin diphosphate, NADH and equine liver alcohol dehydrogenase for 24 h. After removal of the protein, the reactions were subject to BSTFA derivatization and analyzed by gas chromatography (Table 11). Based on the results of these two assays, eight novel

benzoylformate decarboxylases were identified with lower activity than the *P. putida mdhC* encoding gene product. Among these decarboxylases, four of them showed activities toward the non-native substrate 3-deoxy-*glycero*-pentulosonic acid. It is likely that BL21/pML2.118, BL21/pML2.123, BL21/pML2.162 and BL21/pML3.040 contain 3-deoxy-*glycero*-pentulosonate decarboxylase activities. In contrast, single amino acid mutant of benzoylformate decarboxylase BL21/pML3.054 (BFD A460I) showed more than a 25-fold activity decrease in using 3-deoxy-*glycero*-pentulosonic acid as substrate (Table 12). The ability of this enzyme to use benzoylformic acid as substrate also dropped compared to what is reported in the literature.⁵¹ All four pyruvate decarboxylase mutants, on the other hand, were able to synthesize a small amount of 1,2,4-butanetriol, as indicated by the *in vitro* enzymatic reaction experiments (Table 12). Although 3-deoxy-*glycero*-pentulosonate decarboxylase activities were successfully generated in these mutants, their activities were low when comparing to the *P. putida mdhC* encoding decarboxylase.

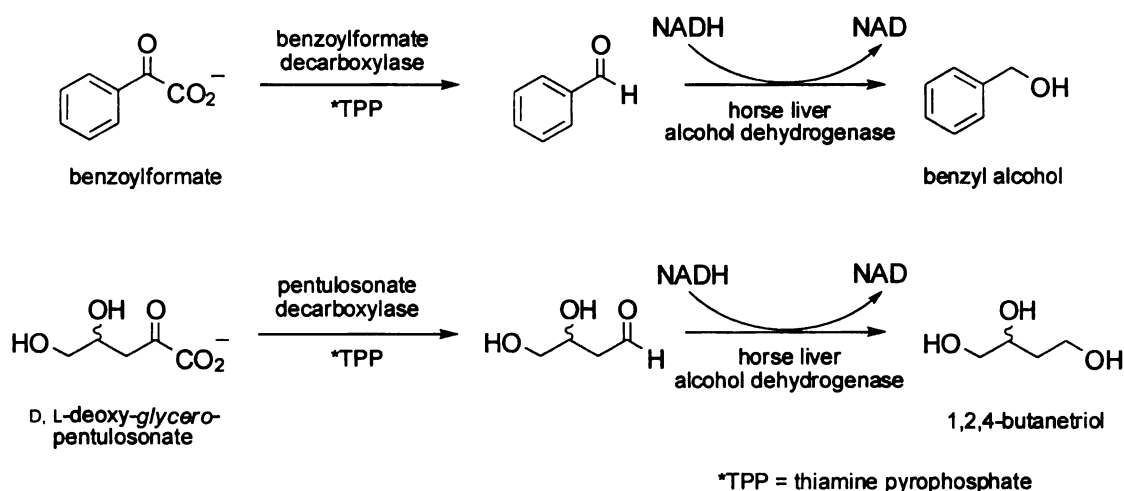


Figure 50. Coupled enzyme assay for decarboxylase activities.

Table 11. Activity table for wild-type pJF118EH clones.

construct	bacterial strain	specific activity (U/mg)		GC yield of BT ^b (%)
		benzoylformate	pentulosonate ^a	
DH5 α /pWN5.238A	<i>P. putida</i>	25.5	0.01	5.3
BL21/pML2.118	<i>P. fluorescens</i>	0.79	0.003	0.5
BL21/pML2.123	<i>P. aeruginosa</i>	0.26	<0.001	0.6
BL21/pML2.162	<i>B. fungorum</i>	0.51	0.009	0.15
BL21/pML3.040	<i>A. ferroxidans</i>	0.01	0.03	0.5
BL21/pML2.208	<i>S. coelicolor</i>	0.03	0.002	<0.01
BL21/pML2.214	<i>N. aromaticivorans</i>	0.003	<0.001	<0.01
BL21/pML2.256	<i>R. palustris</i>	0.005	0.008	<0.01
BL21/pML2.286	<i>M. smegmatis</i>	0.006	0.001	0

^a3-deoxy-glycero-pentulosonic acid (pentulosonate)^b1,2,4-butanetriol (BT)**Table 12. Activity table for benzoylformate and pyruvate decarboxylase mutants.**

construct	mutation	specific activity (U/mg)			GC yield of BT ^b (%)
		benzoylformate	2-oxo-hexanoate	pentulosonate ^a	
pWN5.238A	wt ^c	26.2	0.2	0.01	5.3
pML3.054	A460I	0.5	0.3	<0.01	0.2
pLOI276	wt ^d	0.03	0.08	<0.01	0
pML3.062	I472A	0.3	10.3	<0.01	0.03
pML3.084	I472S	<0.01	0.13	<0.01	0.08
pML3.086	I472G	0.01	1.26	<0.01	0.04
pML3.104	I472T	<0.01	0.14	<0.01	0.04

^a3-deoxy-glycero-pentulosonic acid (pentulosonate)^b1,2,4-butanetriol (BT)^cwild type *P. putida* mdIC encoding benzoylformate decarboxylase^dwild type *Z. mobilis* pdc encoding pyruvate decarboxylase

Table 13. Amino acid composition of the gene candidates.

bacteria	Arg		Leu		Ile		Pro	
	occur. ^a	%	occur. ^a	%	occur. ^a	%	occur. ^a	%
<i>P. putida</i>	23	4.4	50	9.5	25	4.7	40	7.6
<i>P. fluorescens</i>	29	5.5	54	10.2	22	4.2	47	8.9
<i>P. aeruginosa</i>	32	6.1	58	11.0	21	4.0	49	9.3
<i>B. fungorum</i>	29	5.5	51	9.7	28	5.3	41	7.8
<i>A. ferroxidans</i>	23	4.3	54	10.1	37	6.9	42	7.9
<i>S. coelicolor</i>	38	8.0	50	10.5	10	2.1	43	9.1
<i>N. aromaticivorans</i>	35	6.7	50	9.6	19	3.6	39	7.5
<i>R. palustris</i>	29	6.1	47	9.8	16	3.3	47	9.8
<i>M. smegmatis</i>	34	5.4	53	8.5	22	3.5	57	9.1

^anumber of occurrence (occur.)

Table 14. Rare codon analysis of the gene candidates.

bacteria	GC content (%)	Arg ^a	Leu ^b	Ile ^c	Pro ^d
<i>P. putida</i>	62.4	9	1	0	4
<i>P. fluorescens</i>	61.2	0	0	0	16
<i>P. aeruginosa</i>	67.1	1	0	0	13
<i>B. fungorum</i>	62.3	3	2	0	8
<i>A. ferroxidans</i>	58.3	6	3	9	7
<i>S. coelicolor</i>	72.3	0	0	0	19
<i>N. aromaticivorans</i>	62.5	3	12	0	6
<i>R. palustris</i>	64.7	1	0	0	7
<i>M. smegmatis</i>	67.4	9	2	2	23

^aAGG, AGA, CGA; ^bCTA; ^cATA; ^dCCC

Codon usage

In a nutshell, four novel 3-deoxy-*glycero*-pentulosonate decarboxylases were successfully identified in the previous experiment. Insight into these genes was gained using bioinformatics tools to evaluate codon usage (Table 13 – 10), compare sequence identities (Table 15) and search for conserved domains (Figure 51). Analysis of crude cell-free lysate by SDS-PAGE suggested that protein expression levels of the recombinant plasmids were low. Codon bias in *E. coli* is believed to be one of the factors leading to low heterologous gene expression. Codons that were suspected to be associated with translational inefficiency in *E. coli* were examined (Table 13 – 10). The differences in codon usage between *E. coli* and the bacterial strains in our study can account for the observed low protein expression level. For example, most gene candidates encode a large amount of proline often found at turns of folded proteins, thus errors in translation of the associated codons can result in protein misfolding and inclusion bodies.

Amino acid and nucleotide sequence identity

In light of our wild-type decarboxylase activities, use of sequence alignment tools is inadequate in searching for an active 3-deoxy-*glycero*-pentulosonate decarboxylase to replace the *P. putida* benzoylformate decarboxylase. Advanced protein engineering methods can provide an answer to those limitations. For example, gene family shuffling⁵⁴ during which chimeric progeny genes are generated to maximize sequence diversity by recombining a set of naturally occurring homologous genes is a powerful method. The resulting shuffled library can be screened for activity towards a particular substrate. However, such a method requires above 60% nucleotide sequence identities among shuffled genes. Nucleotide and amino acid sequence identities among the active

candidates point toward *P. fluorescens*, *P. aeruginosa* and *B. fungorum* as suitable for us in gene shuffling experiments (Table 15).

Table 15. Nucleotide and amino acid (in parentheses) sequence identity.

	<i>P. putida</i>	<i>P. fluorescens</i>	<i>P. aeruginosa</i>	<i>B. fungorum</i>	<i>A. ferroxidans</i>
<i>P. putida</i>		64% (62%)	66% (63%)	66% (64%)	30% (39%)
<i>P. fluorescens</i>	64% (62%)		81% (82%)	70% (67%)	33% (40%)
<i>P. aeruginosa</i>	66% (63%)	81% (82%)		72% (71%)	33% (40%)
<i>B. fungorum</i>	66% (64%)	70% (67%)	72% (71%)		31% (40%)
<i>A. ferroxidans</i>	30% (39%)	33% (40%)	33% (40%)	31% (40%)	

Identification of novel benzoylformate decarboxylases

It is noteworthy that the *P. putida mdhC* gene product is the only benzoylformate decarboxylase known to date. In the course of our investigation of 3-deoxy-glycero-pentulosonate decarboxylases with high identity with *mdhC*, we identified three new benzoylformate decarboxylases in the cloned *P. fluorescens*, *P. aeruginosa*, *B. fungorum* gene products (Table 11). This is all the more interesting as recent publications identified several important amino acid residues within the active site of *P. putida* BFD that bind to the substrate and co-factor thiamine pyrophosphate by a high resolution X-ray crystal structure.⁴¹ Figure 3 shows the multiple sequence alignments of *P. putida* and the other four active candidates. *P. fluorescens*, *P. aeruginosa*, *B. fungorum* genes show a perfect alignment of important residues, which suggests that these genes are likely to be encoding thiamine pyrophosphate dependent benzoylformate decarboxylases.

Pf	-MKTVHSASYDILRQQGLTTVFGNPGS NEL PFLKGFPEDFRYILGLHEGAVVGMADGFAL	59
Pa	-MKTVHSASYEILRRHGLTTVFGNPGS NEL PFLKDFPEDFRYILGLHEGAVVGMADGFAL	59
Bf	-MKTVQHAAYEILRRHGLTTIFGNPGS NEL PFLKHFPSPDFRYILGLHEGVVTGMADGYAQ	59
Pp	-MASVHGTTYELLRRQGIDTVFGNPGS NEL PFLKDFPEDFRYILALQEACVVGIADGYAQ	59
Af	MHVSVKTATFKLLRQLGMTWIVGNPGS TEL FLQDLPEDFTYIQLHEGSVVAIADGLAQ	60
Pf	ASGQPAFVNLHAAAGTGNGMGALTNAWYSHSPLVITAGQQVRSMIGVEAMLANVDAPQLP	119
Pa	ASGRPAFVNLHAAAGTGNGMGALTNAWYSHSPLVITAGQQVRSMIGVEAMLANVDAGQLP	119
Bf	ATGNPAFVNLHSAAGTGNAMGALANAWNSHTPLVVTSGQQVRSTIGMEPLLANVDANLP	119
Pp	ASRKPAFINLHSAAGTGNAMGALSNAWNSHSPVITAGQQTRAMIGVEALLTNVDAANLP	119
Af	GLRQAVLVNVHTGVGLGNAMGAILTAYQNKTPLIITSGQQTRDMLLFEPLLTNVNAVTMP	120
120 to 239		
Pf	CPFPTRHPSFRGVLPAAIAGISRCLADHDLILVVGAPVFRY HQ FAPGDYLPAGTELLHIT	299
Pa	CPFPTRHACFRGVLPAAIAGISRLLDGHDILVVGAPVFRY HQ FAPGDYLPAGAEVLQVT	299
Bf	CSFPTTHACFRGVLPAIASISRLLDGHDILVVGAPVFRY HQ YEPGALLPAGAEVLSIT	299
Pp	CPFPTRHPCFRGLMPAGIAAISQLLEGHDVVLVIGAPVFRY HQ YDPGQYLKPGTRLISVT	299
Af	AIFPCSNSLYAGILPAAIAPLADCLRGHDLAIVVGAPVFRY Y PYVAGHYLPEGTSLLQIT	297
300 to 418		
Pf	E---RRVIGVIGDGSANYGITALWTAQYQIPVVFIIILKNGTY GALRWF AGVLQVSDAP	474
Pa	R---RQVIGIIGDGSANYGITALWSAAQYRVPVFIILKNGTY GALRWF AGVLEVPDAP	474
Bf	G---RRVIGIIGDGSANYGITALWTAQYSIPTIFIIMKNGTY GALRWF AGVLGVEDVP	474
Pp	E---RQVIAVIGDGSANYSISALWTAQYNIPTIFVIMNNGTY GALRWF AGVLEAENVP	474
Af	ISGRHRPVIVFVGDAFNYPQCIYTGVRHHTHVFVVLQNHEYAIL KE FAIEEKLKNIP	477
Pf	GLDVPGLDFCIAIRGYGVHSVQANTREAFQALSEALAGDRPVLEIETPTLTIEP----	528
Pa	GLDVPGLDFCIAIRGYGVEALHAATREELEGALKHALAADRPVLEIETPTLTIEP----	528
Bf	GLYVPGIDFCALARGYGVHADSGASLTVALERALSSSRPTLTIEVETLA-----	525
Pp	GLDVPGIDFRALAKGYGVQALKADNLEQLKGSLEALSAGKGPVLEIETVSTVSPVK----	528
Af	GLELPGIAIADIGSAYGAHATVAITAIQLEAAFREAMGYQGVSVLQVPISSDLHLPLID	535

Figure 51. Multiple sequence alignment using ClustalW algorithm. Important amino acid residue are highlighted (Pp = *P. putida*, Pf = *P. fluorescens*, Pa = *P. aeruginosa*, Bf = *B. fungorum*, Af = *A. ferrooxidans*).

Directed evolution of 3-deoxy-*glycero*-pentulosonate decarboxylase for microbial 1,2,4-butanetriol synthesis

Background

Created biosyntheses of 1,2,4-butanetriol from pentoses involves four enzymatic conversions and were designed based on *Pseudomonas fragi* D-xylose, L-arabinose catabolism and *E. coli* D-xylonic acid catabolism. The enzymatic decarboxylation of 3-deoxy-*glycero*-pentulosonic acid to 3,4-dihydroxybutanal using *Pseudomonas putida* *mdlC* gene-encoded benzoylformate decarboxylase is known to play a critical role in the final 1,2,4-butanetriol yield and concentration. However, this decarboxylase showed unsatisfactory catalytic activity towards the non-native substrate 3-deoxy-*glycero*-pentulosonic acid. To this end, searching for a more active 3-deoxy-*glycero*-pentulosonate decarboxylase has been a central focus of efforts to improve the pentose-based microbial synthesis of 1,2,4-butanetriol. A variety of strategies had been employed targeting this goal. In the earlier section, bioinformatics searches on available wild-type bacterial genomes yielded eight decarboxylase candidates. Among them, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa* and *Burkholderia fungorum* were confirmed to encode 3-deoxy-*glycero*-pentulosonate decarboxylase activity. The enzyme activity of these novel decarboxylases was found to be lower than the *mdlC* gene encoding benzoylformate decarboxylase from *P. putida*. However, these genes appeared to be shared a high degree of nucleotide sequence identity (Table 15). An alternative strategy to take advantage of this fact is to evolve a better decarboxylase for microbial 1,2,4-butanetriol synthesis using DNA family gene shuffling.

Since its first application about a decade ago, directed evolution has rapidly become a powerful tool used in protein engineering research.⁵⁵ This methodology allows significant improvement of enzyme performance without prior knowledge of its structure or mechanism. In addition, by studying the mutants obtained from directed evolution experiments, one can gain a better understanding toward fundamental questions related to protein folding, substrate binding, and to identify important amino acid residues of an enzyme.⁵⁶ A previous attempt to evolve a better 3-deoxy-*glycero*-pentulosonate decarboxylase in the Frost group using error-prone PCR successfully isolated an enzyme mutant with a two-fold activity improvement.³⁵ It is generally believed that directed evolution of a single parent gene would be limited by inadequate sequence diversity.⁵⁴ To circumvent this issue, directed evolution employing single-stranded DNA (ssDNA) (Figure 52) and double-stranded DNA (dsDNA) family shuffling have been employed using *P. putida mdhC* and the newly identified 3-deoxy-*glycero*-pentulosonate decarboxylase genes from *P. fluorescens*, *P. aeruginosa* and *B. fungorum*. DNA family shuffling represents an alternative or additional approach for generating genetic diversity based on the mixing and concatenation of genes from a number of parent sequences with high nucleotide sequence identity.⁵⁴ As compared to random mutagenesis, DNA shuffling is advantageous in concentrating beneficial mutations that have arisen independently and subsequently translating into preferred enzyme mutant properties. These four parent genes were randomly fragmented and recombined to generate a chimera library. A 96-well format high-throughput screening method to assay *in vivo* activity of decarboxylase chimeras using intact *E. coli* microbes was also developed. In the end, 4,000 independent mutant clones were screened.

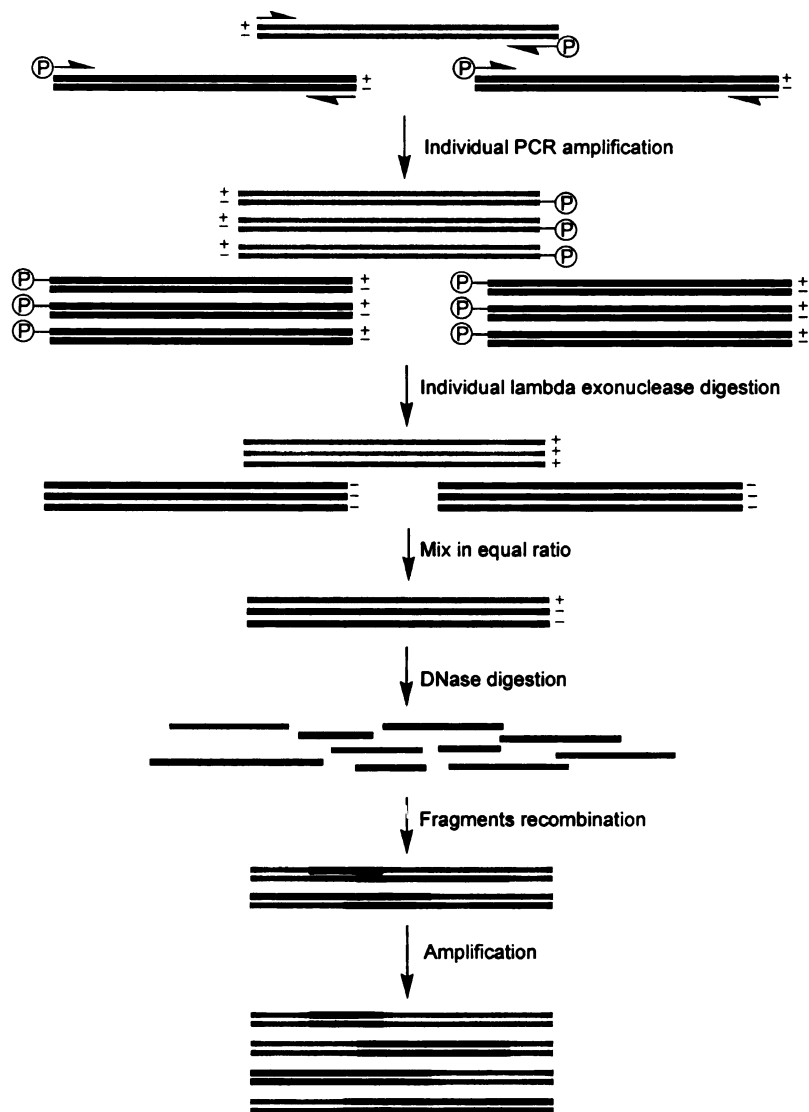


Figure 52. Schematic diagram for single-stranded DNA family shuffling.

Subcloning 3-deoxy-glycero-pentulose decarboxylase encoded genes

The *EcoRI* and *BamHI* restriction sites were eliminated from the *B. fungorum* decarboxylase gene through a QuikChange site-directed mutagenesis kit used to facilitate the chimera library. This double mutant was prepared using *PfuTurbo* DNA polymerase with the following oligonucleotide primers: 5'-CCAACGC-*GTGGA***ACT**CGCATACGCCG-3' and 5'-GCGGCGGCGGG**CAT**CCAGCTCGCG-3'. The *EcoRI* and *BamHI* sites are italicized and the mutations are in boldface. The mutations were designed such that the encoded amino acids remain unchanged. The plasmid that harbors this *B. fungorum* mutated gene was designated pML3.200.

The decarboxylase genes from *P. fluorescens*, *P. aeruginosa* and *B. fungorum* were amplified by PCR from plasmids pML2.118, pML2.123 and pML3.200, respectively. Only the ORFs were amplified from these new plasmids and they shared the same 20 nucleotides at both ends of the genes. The inclusion of a *HindIII* restriction enzyme site on the sense primer and an *EcoRI* restriction enzyme site on the anti-sense primer allowed digestion of the PCR product with both restriction enzymes and ligation of the digestion product into vector pJF118HE. These gene fragments were subcloned into pJF118HE to yield plasmids pML3.224 (*P. fluorescens*, Figure 53), pML3.225 (*P. aeruginosa*, Figure 54) and pML3.226 (*B. fungorum*, Figure 55). As a consequence, the inserted DNA could only be located into the vector in a defined orientation. This directional cloning strategy ensured that the insert gene would be expressed off a *tac* promoter located upstream from the *HindIII* recognition site on pJF118HE. Due to the expression of a plasmid-encoded Lac repressor protein, the *tac* promoter initiated transcription of mutant genes in different clones could be regulated by IPTG addition.

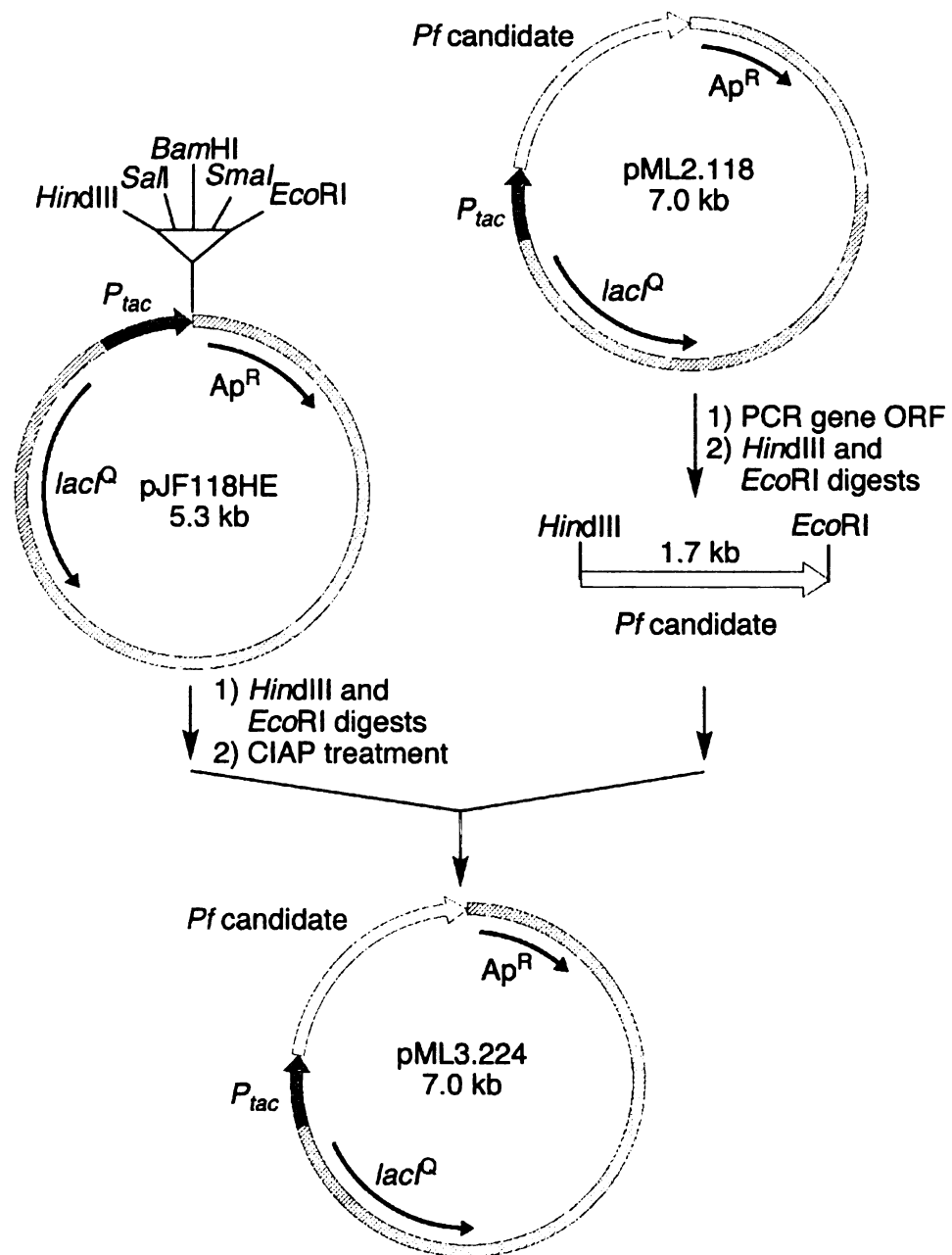


Figure 53. Construction of plasmid pML3.224.

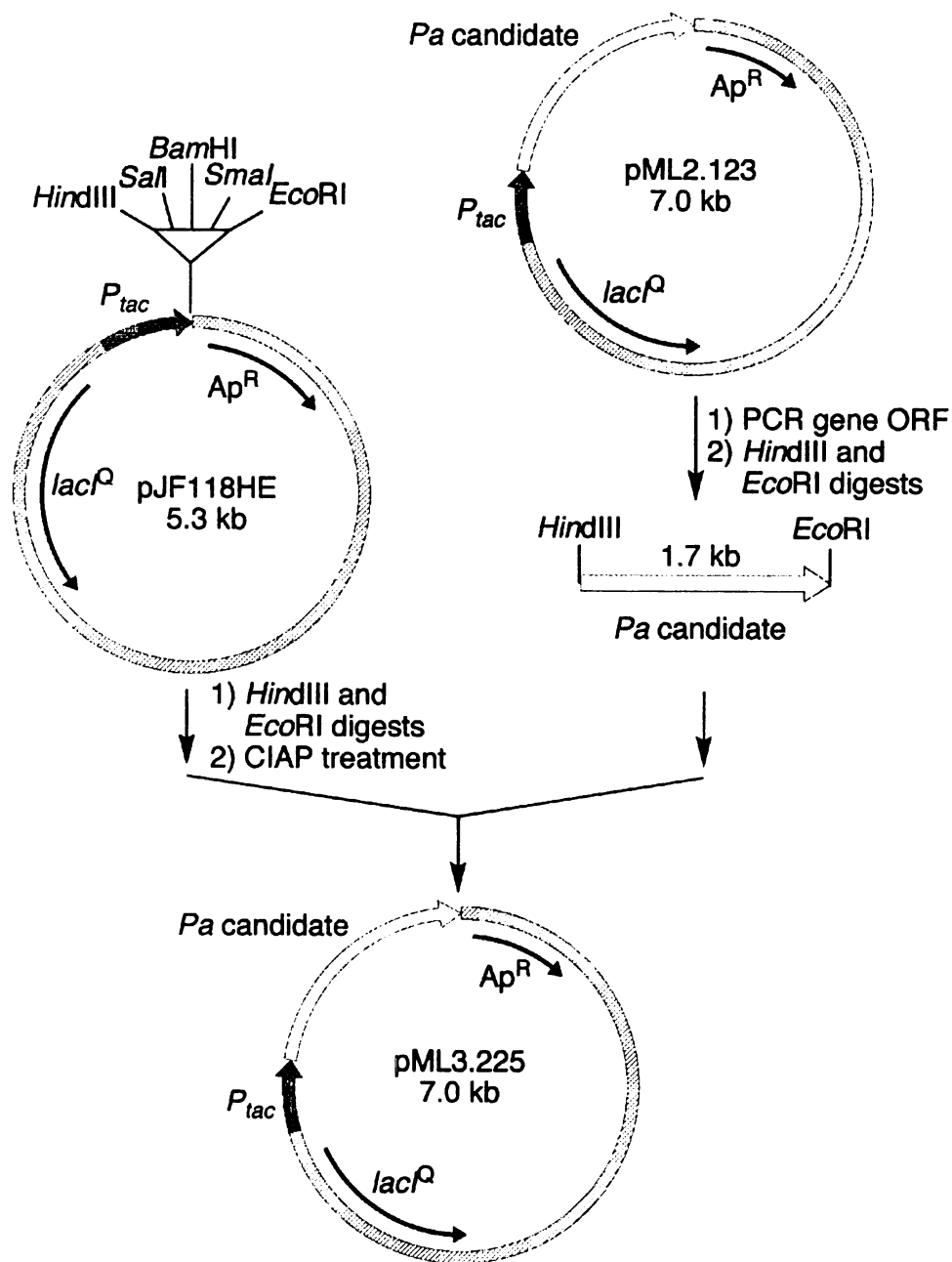


Figure 54. Construction of plasmid pML3.225.

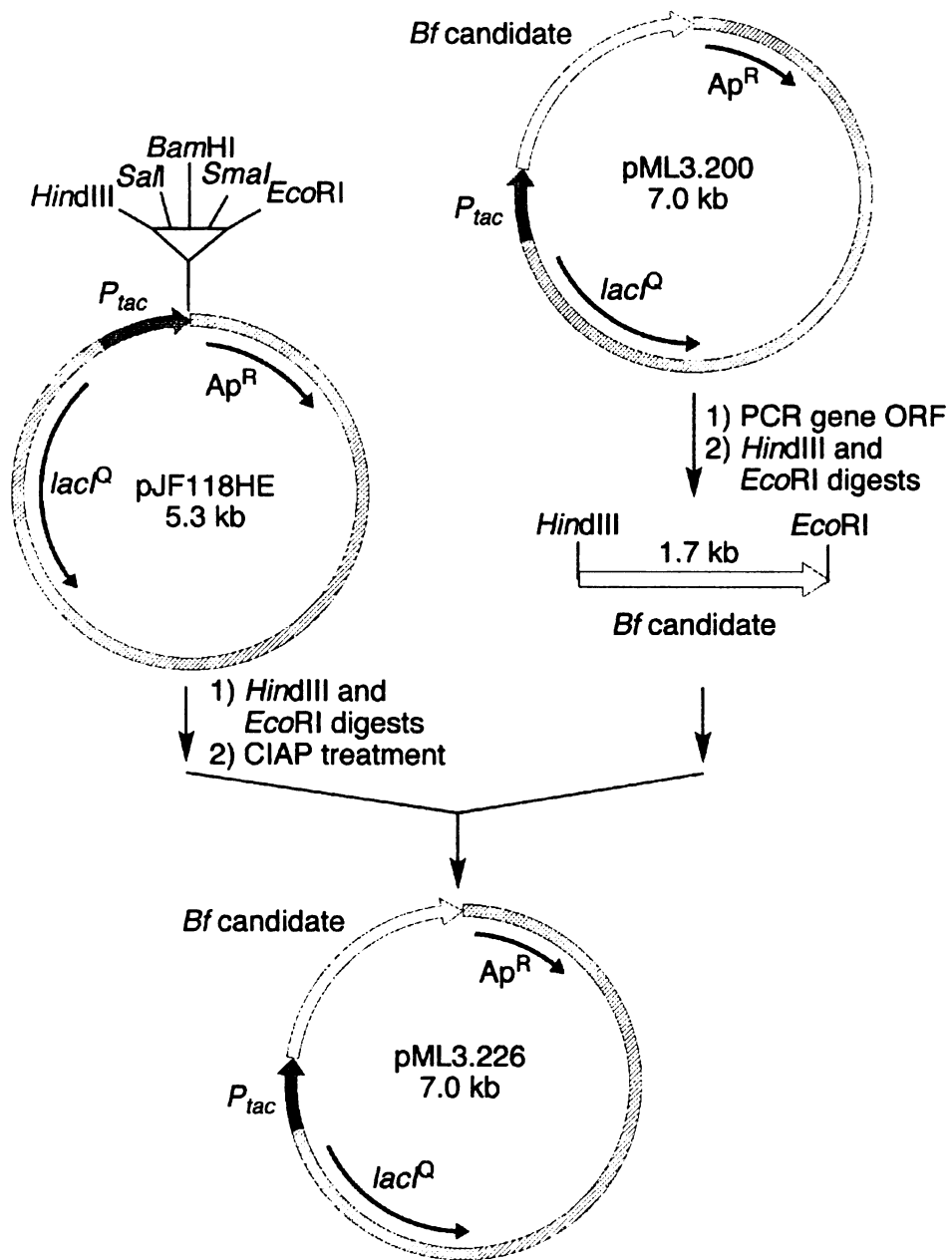


Figure 55. Construction of plasmid pML5.226.

Chimera library construction for ssDNA family shuffling

The decarboxylase parent genes were amplified from plasmids pML3.224, pML3.225 and pML3.226, respectively using suitable phosphorylated primers (Figure 52). 5 µg of each purified PCR product was subject to lambda exonuclease digestion to yield the desired upper or lower ssDNA. 0.5 µg of each purified ssDNA was then mixed and digested with DNase I for 30 sec. The 500 – 1,000 bps DNA fragments were gel-purified and recovered using spin columns. The gel purified fragments were subsequently subjected to primerless DNA fragment recombination using a thermocycler. This reassembly reaction mixture served as a DNA template for the subsequent normal PCR reaction. As a final step, the full length PCR product was purified, digested by *HindIII* and *EcoRI* and then cloned into pJF118HE to generate the chimera library. This chimera library was transformed into *E. coli* W3110 for sequence analysis.

Twenty clones were randomly chosen from a 96 well plate and submitted for high-throughput sequence analysis at the Genomics Technology Support Facility at Michigan State University, using the following primers: 5'-CGGCTCGTATAATGTGTGGA-3', 5'-TCCATTCCCTACGACGACTG-3', 5'-GATGCGCTGGTGGGAGACAT-3', and 5'-TCGGCCAACTACGGTATCAC-3'. These primers were designed such that the sequencing results would cover the entire gene. Comparison of the nucleotide sequence of 20 shuffled pentulosonate decarboxylase genes showed the unsuccessful recombination among *P. fluorescens*, *P. aeruginosa* and *B. fungorum* genes (Figure 56). These chimeras had an average number of crossover equal to one, where *B. fungorum* gene fragments rarely ended up into the hybrid genes. In addition, the chimera library was heavily biased towards *P. fluorescens* at the 5'-end, and *P. aeruginosa* at the 3'-end. There were several

possible explanation for these results. Family shuffling with large parent gene sizes (>1.5 kb) with low nucleotide homology (< 70%) was generally agreed upon to be problematic. The use of large ssDNA fragments (500 – 1000 bps) in the recombination step was certainly another factor. Attempts to generate the chimera library using smaller fragments was unsuccessful, leading us to explore dsDNA family shuffling as an alternative.

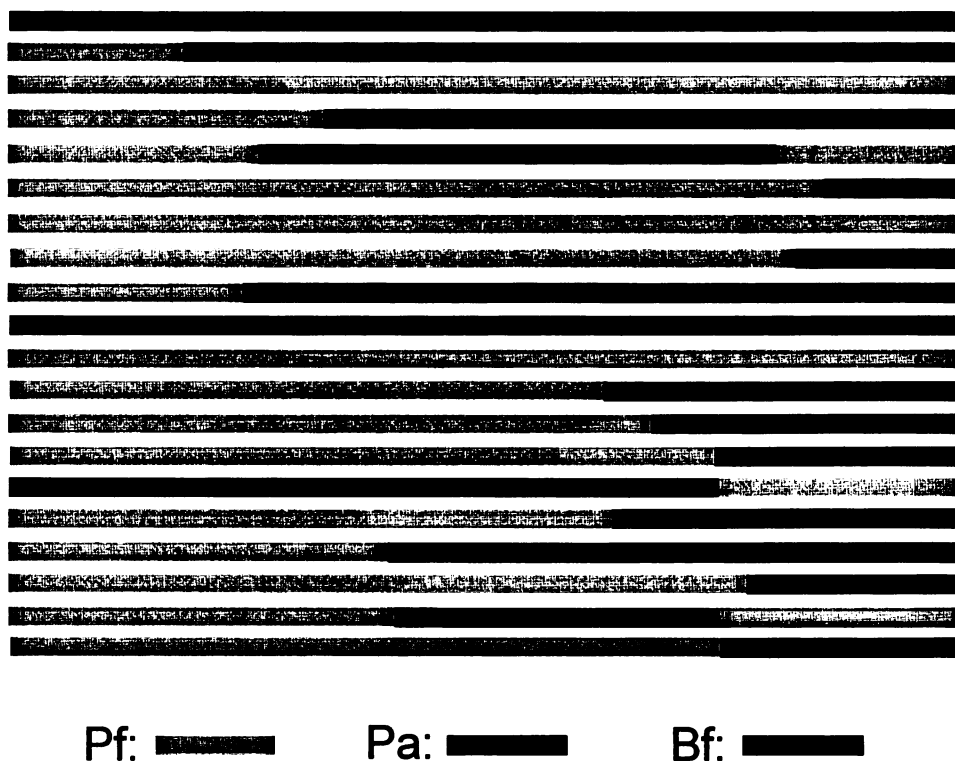


Figure 56. Sequencing results of 20 randomly chosen clones from ssDNA family shuffling. Pf = *P. fluorescens*; Pa = *P. aeruginosa*; Bf = *B. fungorum*

Chimera library construction for dsDNA family shuffling

dsDNA family shuffling is highly similar to the ssDNA family shuffling method, however, the use of phosphorylated primers and lambda exonuclease digestion are not required. Joern, Arnold and coworkers reported the protocol to generate a chimera library

for dsDNA family shuffling.⁵⁷ A 2.1 kb gene was successfully recombined with an average crossover of 3.7. Given their success, a chimera library of decarboxylase genes from *P. putida*, *P. fluorescens*, *P. aeruginosa* and *B. fungorum* was generated.

The four parent genes were amplified using a pair of non-phosphorylated primers. Purified PCR products (1 µg) from each microorganism was mixed and digested with DNase I for 4 min. The 100 – 300 bps DNA fragments were gel-purified and recovered using spin columns. These fragments were subsequently recombined in a thermocycler (Figure 57, left). This reassembly reaction mixture (5 µL) was employed as a DNA template for the subsequent PCR reaction (Figure 57, right). The full length 1.6 kb PCR product was purified, digested by *Hind*III and *Eco*RI, and ligated into plasmid pJF118HE. This chimera library was then transformed into *E. coli* W3110 for sequence analysis and *in vivo* enzyme activity screening.

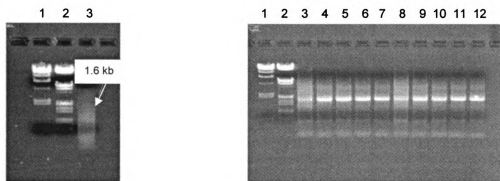


Figure 57. DNA agarose gels for dsDNA family shuffling. (left) primerless recombination reaction, lane 1 and 2: DNA size markers, lane 3: recombination reaction sample; (right) normal PCR for full length gene, lane 1 and 2: DNA size marker, lanes 3 – 7: PCR reaction at different annealing temp. 51.2 °C, 53.6 °C, 56.4 °C, 58.2 °C, 60.0 °C, respectively, lanes 8 – 12: replicates of lane 3 – 7.

22 clones were randomly chosen from a 96-well plate and submitted for high-throughput sequence analysis. The results are summarized in Figure 58. In these 22

sequenced clones, the highest crossover obtained was 7 with an average of 3.4. The diversity of a chimera library depends on the number of crossovers, and more crossovers are expected when participating genes share a high level of sequence identity. This explains the gene fragment distribution shown in Figure 58. *P. fluorescens* and *P. aeruginosa* genes share up to 82% sequence identity, the fraction of crossover occurring between them were expected to be more frequent. On the other hand, the *P. putida* gene only shares 64% identity to the rest, leading to the fact that the fraction of crossovers was relatively low.



Figure 58. Sequencing results of 22 randomly chosen clones from dsDNA family shuffling. Pp = *P. putida*; Pf = *P. fluorescens*; Pa = *P. aeruginosa*; Bf = *B. fungorum*.

Chimera library screening for 3-deoxy-glycero-pentulosonate decarboxylase

Designing a reliable and efficient methodology for the identification of active decarboxylase mutants possessing improved catalytic activity towards 3-deoxy-glycero-pentulosonic acid is of premier importance in this study. Decarboxylase activities of the mutants were screened *in vivo* by cultivating *E. coli* W3110 transformants of the chimera library in minimal salts medium containing D-xylonic acid. D-Xylonic acid was synthesized according to literature procedure⁵⁸ and served both as a carbon source for growth as well as substrate for the potential decarboxylase. The decarboxylase activity was measured by the amount of D-1,2,4-butanetriol accumulated in the culture medium based upon a response factor using gas chromatography.

Screening of the 3-deoxy-glycero-pentulosonate decarboxylase chimera library was carried out in a 96-well format. Single colonies from the chimera library were inoculated and cultured in a 96-well growth block. After 36 h of incubation at 37 °C, a glycerol freeze plate was made and stored at -80 °C. Simultaneously, a fresh growth block was inoculated using a 10% seed culture from the first block and incubated for another 24 h. Protein expression was induced with IPTG, and the cells were harvested by centrifugation after 12 h. 100 µL of clarified culture medium was sampled from each well and evaporated to dryness in a CentriVap (Labconco) under reduced pressure. The residue was derivatized with a reagent mixture containing bis(trimethylsilyl)trifluoroacetamide (BSTFA). 1,2,4-Butanetriol was quantified by gas chromatography with dodecane as an internal standard.

Chimera library screening of 3,960 independent mutants was carried out in a 96-well format using the method previously described. In general, around 90% of the

mutants were found to be inactive, while the rest of the mutants carried genes with 3-deoxy-*glycero*-pentulose decarboxylase activity lower than our benchmark, benzoylformate decarboxylase from *P. putida*. In order to determine whether the failure to isolate better mutants was due to possible experimental errors from our screening method, a secondary screening using the same methodology was carried out and the results were compared to the primary screening. Five active mutants were picked arbitrarily based on the results obtained in the primary screening and were screened again (Table 16). The activity fold between the two screenings indicated a maximum error of ± 0.3 , so therefore it was unlikely that potent mutants would have been missed due to poor assay design. Since over 90% of the mutants were assayed as inactive clones, testing for the possibility of false-negative results was also of particular importance. Seven mutants were randomly screened accordingly (Table 17). This experiment showed zero activity for all mutants toward D-1,2,4-butanetriol synthesis, which was consistent to the results obtained in the primary screening.

Consistency of cell growth inside a 96-well block was found to be problematic, which was believed to be the consequence of inadequate aeration inside the incubated shaker. This became more obvious during screening experiments involving a prolonged incubation period of six days. Putting spacers between blocks and leaving space between stacks of blocks were measures taken to facilitate aeration. At the end of the screening, eight mutants showed relatively high D-1,2,4-butanetriol formation with low cell density and were subject to a second screen. All eight candidates showed normal cell growth in this second screening experiment and the results are shown in Table 18. Except 21G5, all other mutants showed lower activity relative to *P. putida* Md1C.

Table 16. Data reproducibility using *in vivo* GC screening assay.

entry	mutant	cell density OD ₆₀₀	activity fold ^a	
			1 st screen	2 nd screen
1	16D10	1.67	0.9	0.9
2	18A5	2.67	0.4	0.5
3	19F3	3.39	0.9	0.7
4	28D2	2.64	0.9	0.6
5	29D6	2.17	0.9	1.1

^aactivity fold = mol of 1,2,4-butanetriol produced with mutant/mol of 1,2,4-butanetriol produced with W3110/pWN5.238A (control)

Table 17. Test for false-negative results.

entry	mutant	cell density OD ₆₀₀	activity fold ^a	
			1 st screen	2 nd screen
1	29A3	2.53	0	0
2	29E6	2.41	0	0
3	30A2	3.05	0	0
4	30E9	1.70	0	0
5	32A2	1.49	0	0
6	32B5	2.75	0	0
7	32E12	3.07	0	0

^aactivity fold = mol of 1,2,4-butanetriol produced with mutant/mol of 1,2,4-butanetriol produced with W3110/pWN5.238A (control)

Table 18. Test for mutants with growth issues.

entry	mutant	cell density OD ₆₀₀	activity fold ^a	
			1 st screen	2 nd screen
1	21G5	2.76	3.5	1.3
2	22G2	2.50	1.7	1.0
3	26H5	3.04	2.8	0.5
4	27A9	2.18	1.5	0.4
5	27D9	2.16	2.0	0.8
6	28B7	2.63	1.1	0.9
7	29H11	2.06	2.2	0.6
8	31G8	2.40	1.5	0.4

^aactivity fold = mol of 1,2,4-butanetriol produced with mutant/mol of 1,2,4-butanetriol produced with W3110/pWN5.238A (control)

Mutants 29D6 and 21G5 showed activity fold increases of 1.1 and 1.3, respectively. They were subject to a third round of screening for D-1,2,4-butanetriol production in baffled flasks. Two culture tubes of 5 mL M9/D-xylonate/Ap cultures were inoculated with a single colony of each mutant and allowed to grow overnight at 37 °C. They were used as seed culture to inoculate 100 mL M9/xylonite/Ap medium in 500 mL Erlenmeyer flasks. The experiment was conducted at 30 °C with agitation at 250 rpm. After 24 h, 1 mL of culture was sampled from each flask. Clarified culture medium was subsequently evaporated to dryness, subject to BSTFA derivatization and finally assayed by gas chromatography. Control constructs *E. coli* W3110/pWN5.238A, W3110/pML3.224, W3110/pML3.225 and W3110/pML3.226 showed reproducible activities (Table 19). The activity fold for mutants 29D6 and 21G5 were found to be 0.4 and 0.8, respectively. To understand the structure of these chimera decarboxylase genes, all the mutant candidates that entered the secondary screening were sequenced.

Table 19. Shake flask experiments for improved decarboxylase activities.

entry	construct	cell density OD ₆₀₀	activity fold ^a
1	W3110/pWN5.238A (Pp)	4.2	1.0
2	W3110/pML3.224 (Pf)	4.0	0.4
3	W3110/pML3.225 (Pa)	4.1	0.5
4	W3110/pML3.226 (Bf)	4.0	0.2
5	W3110/21G5	4.0	0.8
6	W3110/29D6	4.1	0.4

^aactivity fold = mol of 1,2,4-butanetriol produced with mutant/mol of 1,2,4-butanetriol produced with W3110/pWN5.238A (control)

Sequence alignment of mutants from the secondary screening is shown in Figure 59. The average numbers of crossovers for active and inactive clones are 4.5 and 6.2, respectively. Gene fragments from *P. putida* were not incorporated in any of these mutants, presumably due to the relatively low nucleotide identity (around 64%) to the other candidates. *B. fungorum* fragments were found in most inactive mutants, suggesting that these fragments might be responsible to the high percentage of inactive clones generated (>90%). Mutants 29D6 and 21G5 were found to have the same amount of crossover (equals 2) and share a similar gene recombination pattern. Both of them contain a small *P. fluorescens* gene fragment in the middle, flanked by *P. aeruginosa* fragments. Interestingly, the absence of *P. putida* gene fragments in mutant 21G5 suggested its decarboxylase activity was solely inherited from *P. fluorescens* and *P. aeruginosa* genes, which in fact resulted in a 1.6-fold activity improvement relative to either parent genes. Nonetheless, this improvement only generated another decarboxylase with activity comparable to the benchmark *P. putida* MdlC.

In summary, *P. putida*, *P. fluorescens*, *P. aeruginosa* and *B. fungorum* were identified to bear pentulosonate decarboxylase encoded genes that were suitable for DNA family gene shuffling. A DNA chimera library of these decarboxylase genes were successfully generated with high numbers of crossover (average crossover = 4.5) using double-stranded DNA family shuffling techniques. Library screening of 3,960 mutants were conducted. Identification of the mutant 21G5 with at least 1.6 fold activity improvement compared to its parent genes was achieved. However, the activity of mutant 21G5 decarboxylase was similar to the *P. putida* MdlC. Nevertheless, the success in

generating a diverse mutant library and the valuable experience gained from this study allows us to revisit the task in the future once high-throughput robotics become available.

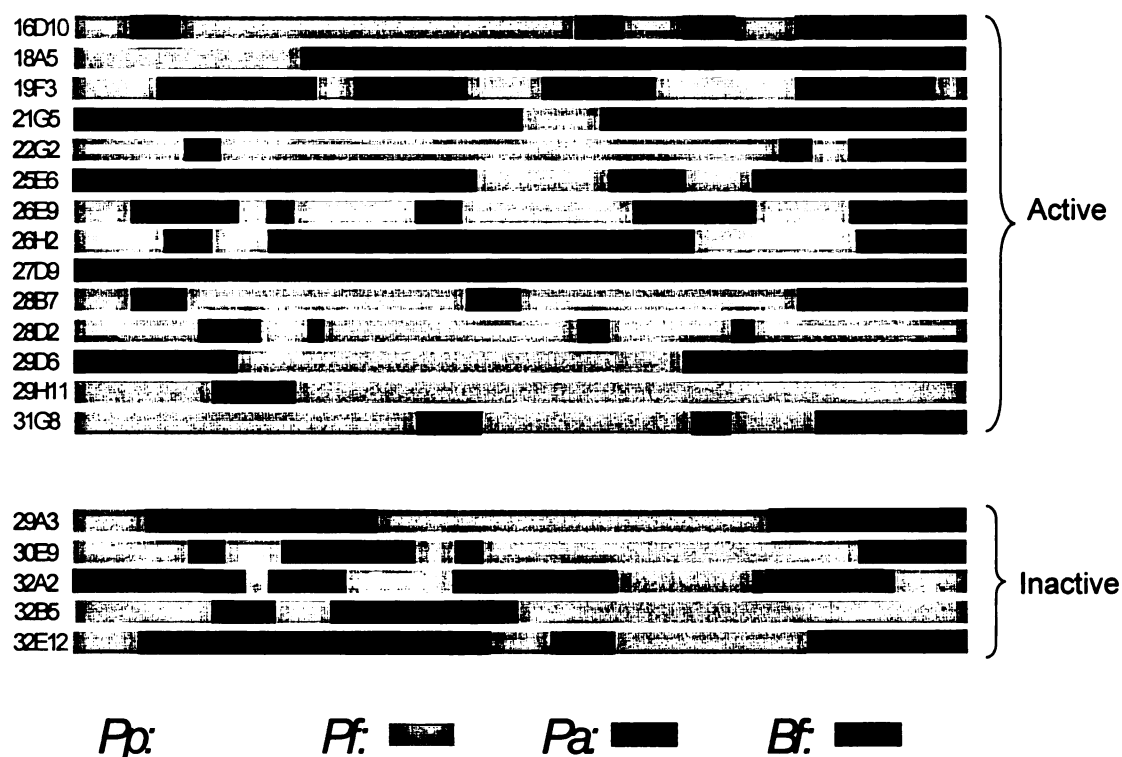


Figure 59. Sequence alignment of selected mutants. Pp = *P. putida*; Pf = *P. fluorescens*; Pa = *P. aeruginosa*; Bf = *B. fungorum*.

Bacteriophage T7 promoter and *E. coli* chaperones for *mdlC* expression

Background

Fermentor-controlled synthesis of D-1,2,4-butanetriol using a single recombinant *E. coli* W3110 derivative WN13/pWN7.126B was scaled up to 100 L at MBI International.⁵⁹ A 500 g quantity of D-1,2,4-butanetriol was successfully isolated as a clear oily liquid using this process. Byproduct formation during D-1,2,4-butanetriol biosynthesis provided a guide for what aspects of this microbial catalyst needed to be improved. As shown in Figure 10, although 6 g/L in 28% yield (mol/mol) of D-1,2,4-butanetriol was synthesized at 1 L working volumes by *E. coli* WN13/pWN7.126B under fermentor-controlled conditions, substantial concentrations of other byproducts are also synthesized. These byproducts included 3-deoxy-D-*glycero*-pentanoic acid (5 g/L), 4,5-dihydroxy-*threo*-L-norvaline (4 g/L), 3-deoxy-D-*glycero*-pentulosonic acid (1 g/L), and 3,4-dihydroxy-D-butanoic acid (3 g/L). One of these byproducts (3-deoxy-D-*glycero*-pentulosonic acid) was likely the precursor to two of the other products (3-deoxy-D-*glycero*-pentanoic acid and 4,5-dihydroxy-*threo*-L-norvaline).³⁶ Reduction of byproduct formation would not only increase the yield and concentration of D-1,2,4-butanetriol production, but would also substantially lower the cost currently associated with downstream purification of product from the fermentation broth.

In the previous section, a large effort towards the discovery of novel 3-deoxy-*glycero*-pentulosonate decarboxylases was discussed. By increasing the rate of conversion of 3-deoxy-D-*glycero*-pentulosonic acid into 3,4-dihydroxy-D-butanal, three of the four byproducts observed during the synthesis of D-1,2,4-butanetriol from D-xylose would

likely be eliminated. While this approach would undoubtedly be effective, alternative approaches will be discussed in this section.

The plasmid-borne *P. putida mdnC* gene encoding benzoylformate decarboxylase in our benchmark *E. coli* microbe WN13/pWN7.126B is transcribed based on a *tac* promoter system. Another commonly used system for achieving high level of protein production depends on the extremely selective nature of the bacteriophage T7 RNA polymerase for specific promoters.^{60,61} To investigate the effect on using a strong T7 promoter, the *P. putida mdnC* gene was cloned into pET28c(+). The resulting plasmid was introduced into a DE3 lysogen of *E. coli* WN13 to generate *E. coli* KIT15/pML7.135. However, culturing this *E. coli* microbe resulted in reduced D-1,2,4-butanetriol production. Co-expression of *E. coli* chaperonins GroEL-GroES in this *E. coli* strain was subsequently pursued in an attempt to restore proper protein folding and assembly of this heterologously expressed *P. putida* benzoylformate decarboxylase. *E. coli* strains WN13/pWN7.180 and WN13/pWN7.202 were generated and evaluated.

Expression of plasmid-borne mdnC gene under phage T7 promoter

Often the first step to improve plasmid-localized enzyme expression in *E. coli* is the use of a strong and controllable expression vector system. Protein expression vectors that utilize the bacteriophage T7 polymerase/promoter system are capable of very high levels of protein production. In the pET vector system (Novagen), the gene encoding T7 RNA polymerase is usually supplied by the host bacterial cell in the form of a λ lysogen that expresses the polymerase gene under control of the *lacUV5* promoter.⁶⁰

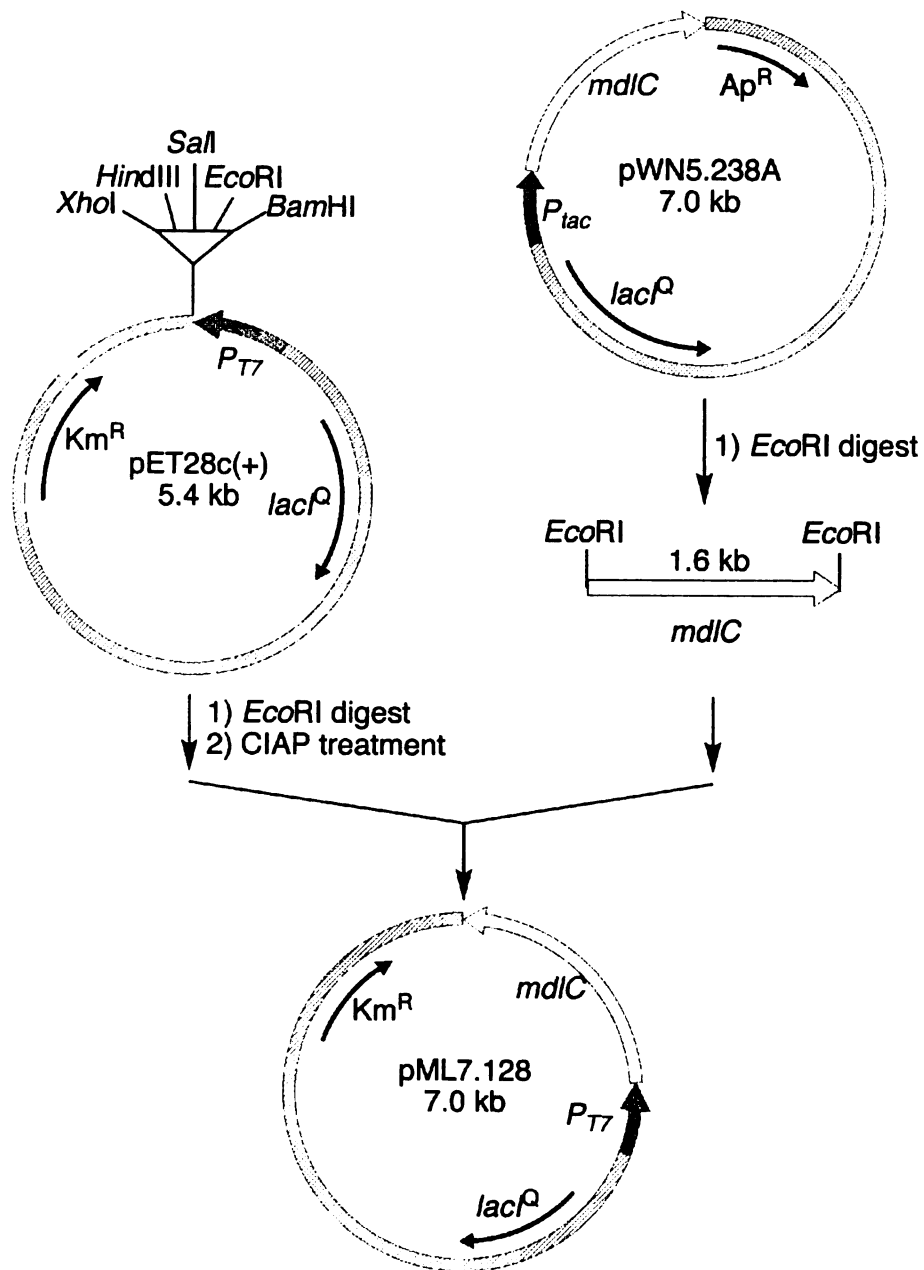


Figure 60. Construction of plasmid pML7.128.

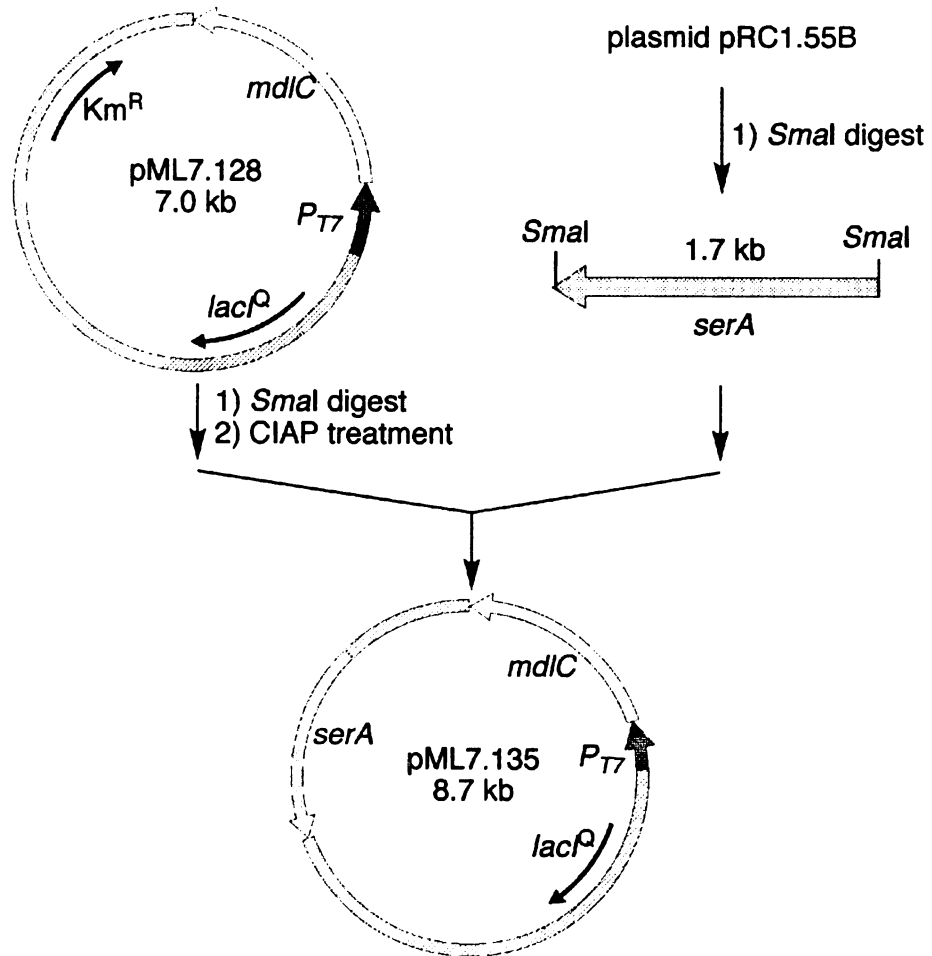


Figure 61. Construction of plasmid pML7.135.

The target gene, *P. putida mdxC* gene was encoded on plasmid pET28c(+) so that transcription could be driven by the T7 RNA polymerase. A 1.6 kb fragment encoding the *P. putida mdxC* gene was excised from plasmid pWN5.238A by digestion with *EcoRI* and ligated to the 5.4 kb plasmid pET28c(+), which had been previously treated with *EcoRI* to afford a 7.0 kb plasmid pML7.128 (Figure 60). A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *SmaI* and ligated to the *SmaI* digested plasmid pML7.128. The ligation mixture was transformed into *E. coli* WN13. Transformants carrying the *serA* insert were selected on M9 medium plates. In the resulting 8.7 kb plasmid pML7.135, the *serA* and *mdxC* genes are convergently transcribed (Figure 61). The DE3 lysogen *E. coli* KIT2 was prepared using the DE3 lysogenation kit of Novagen according to the manufacturer's protocol.⁶² *E. coli* host strain KIT15 was constructed and retransformed with pML7.135. The resulting *E. coli* KIT15/pML7.135 was evaluated under fermentor-controlled conditions.

Fed-batch fermentations were performed in a 2.0 L working capacity Biostat MD B. Braun fermentor equipped with a DCU-3 system and a Dell Optiplex Gs⁺ 5166M personal computer utilizing B. Braun MFCS/Win software (v.2.0) for data acquisition and automatic process monitoring. Temperature, pH, and glucose feeding were controlled by PID control loops. Temperature was maintained at 36 °C by temperature adjusted water flow through a jacket surrounding the fermentor vessel. The pH was maintained at 7.0 by the addition of concentrated ammonium hydroxide or 2 N H₂SO₄. Dissolved oxygen (D.O.) was measured using a Mettler-Toledo 12 mm O₂ sensor fitted with an Ingold A-type O₂ permeable membrane. D.O. level was maintained at 10% of air saturation.

Antifoam (Sigma 204) was added manually as needed. Fed-batch fermentations were run in duplicate, and reported results represent an average of the two runs.

Inoculants were started by introduction of a single colony picked from an M9 agar plate containing glucose into 5 mL of M9 glucose medium. Cultures were grown at 37 °C with agitation at 250 rpm until they were turbid and subsequently transferred to 100 mL of fresh M9 glucose medium. After culturing at 37 °C and 250 rpm for an additional 10 h, the inoculants ($OD_{600} = 3.0-4.0$) were transferred into the fermentation vessel and the batch fermentation was initiated ($t = 0$ h). The initial glucose concentration in the fermentation medium ranged from 24 to 28 g/L according to the growth requirement of different constructs.

Cultivation under fermentor-controlled conditions was divided into three stages with each stage corresponding to a different method for controlling D.O. In the first stage, the airflow was kept constant at 0.06 L/L/min and the impeller speed was increased from 50 rpm to 950 rpm to maintain the D.O. at 10%. Once the impeller speed reached its preset maximum of 950 rpm, the mass flow controller started to maintain the D.O. by increasing the airflow from 0.06 L/L/min to 1.0 L/L/min. Depending on the construct under examination, the two stages took a total of 12 – 15 h for completion. At constant impeller speed and constant airflow rate, the D.O. level was maintained at 10% for the remainder of the fermentation by use of an O_2 sensor to control glycerol feeding. At the beginning of the third stage, the D.O concentration fell below 10% due to the residual glucose initially added to the medium. This lasted from approximately 10 – 30 min before glucose feeding (650 g/L) started. Fermentation cultures typically entered the stationary phase between 24 – 30 h after inoculation of the fermentor culture medium. IPTG stock

solution was added to the culture medium at 18 h to a final concentration of 0.5 mM. The solution of D-xylose was added to the fermentation medium at 24 h, 30 h, 36 h, and 42 h.

Fermentations were run for 48 h. Samples (5 – 10 mL) of the fermentation culture were removed every 6 h after the fermentor was inoculated. Cell densities were determined by dilution of fermentation broth with water (1:100) followed by measurement of OD₆₀₀. Dry cell weight of *E. coli* cells (g/L) was calculated using a conversion coefficient of 0.43 g/L/OD₆₀₀. The remaining fermentation culture was centrifuged to obtain cell-free broth and analyzed by ¹H NMR and gas chromatography. The cell pellets were used for enzyme assays.

Compared with our benchmark *E. coli* WN13/pWN7.126B, which produced D-1,2,4-butanetriol at a concentration of 10.5 g/L at 50% yield after modification of various culturing parameters, KIT15/pML7.135 produced only 1.7 g/L of D-1,2,4-butanetriol in 8% yield under the same culture conditions. Interestingly, this construct yielded 23 g/L 3-deoxy-D-*glycero*-pentulosonic acid, which accounted to 80% of the total D-xylose in the system. Later on, *E. coli* KIT15/pRC1.55B was constructed. Plasmid pRC1.55B carries the *serA* gene locus in order to complement the *serA* mutation of the *E. coli* KIT15. The absence of *mdlC* gene encoding benzoylformate decarboxylase in this construct was expected to give a similar yield and titer of 3-deoxy-D-*glycero*-pentulosonic acid as observed in *E. coli* KIT15/pML7.135. The cultivation of this new construct, however, only produced 11.8 g/L 3-deoxy-D-*glycero*-pentulosonic acid at 40% yield. The fact that nearly half of the D-xylose was recovered at the end of the fermentation indicated that the T7 promoter is not suitable for producing functional decarboxylase. The translation of genes from GC rich organisms is inefficient in *E. coli*.

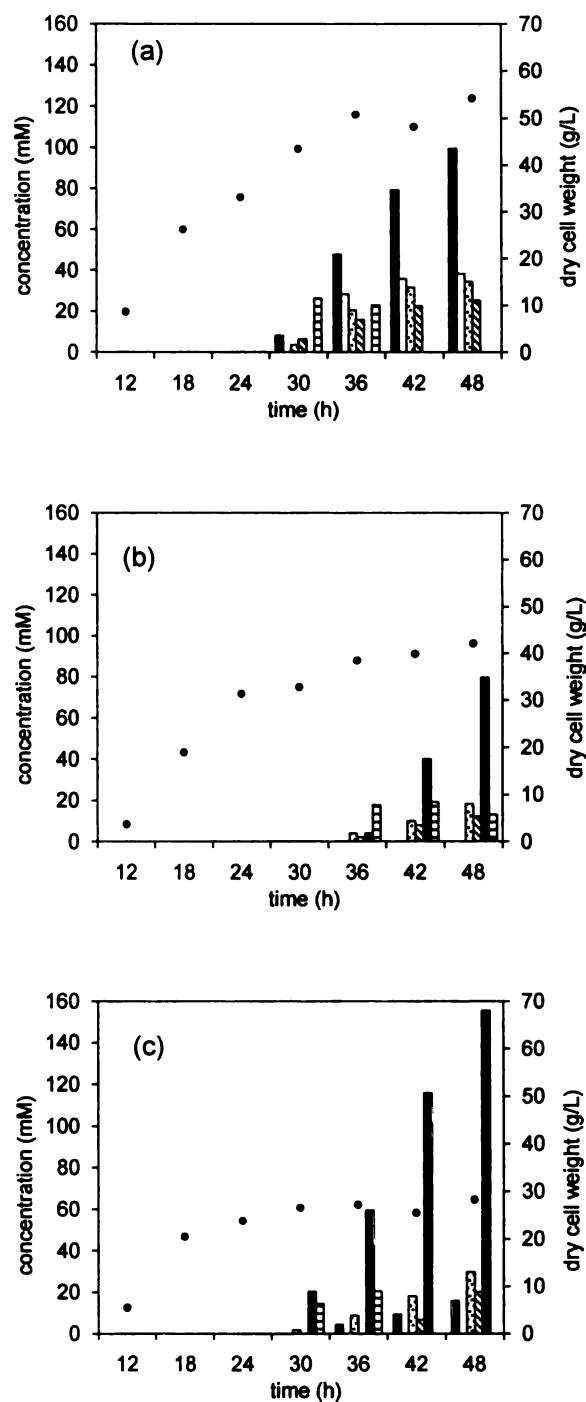


Figure 62. Biosynthesis of 3-deoxy-D-glycero-pentulosonic acid under fermentor-controlled conditions. (a) benchmark construct *E. coli* WN13/pWN7.126B; (b) *E. coli* KIT15/ pRC1.55B; (c) *E. coli* KIT15/pML7.135; ■ D-1,2,4-butanetriol; □ 3,4-dihydroxy-D-butanoic acid; ▤ 3-deoxy-D-glycero-pentanoic acid; ▨ 4,5-dihydroxy-threo-L-norvaline; ▬ 3-deoxy-D-glycero-pentulosonic acid; ▮ D-xylonic acid.

E. coli chaperonins GroES-GroEL for *mdlC* gene expression

E. coli is commonly used as a host for heterologous protein expression. However, there are issues associated with this widely used system. For instances, inclusion bodies and protease degradation are often encountered by researchers as a consequence of improper protein folding. This phenomenon was observed in our previous attempt to express *P. putida mdlC* under a strong T7 promoter. Molecular chaperones have been demonstrated to be involved in the protein folding process *in vivo*,⁶³ and the overexpression of these proteins proved to enhance active protein production.⁶⁴ The *E. coli* chaperonin GroEL first binds ATP and GroES co-chaperonin to create a large binding cavity. The resulting complex then encapsulates partially folded or misfolded proteins and unfolds it. As a final step in this GroEL-GroES mode of action, the partially unfolded protein will undergo conformational changes yielding a hydrophobic core that compels the misfolded protein to refold and obtain its correct structure.⁶⁵ In a recent report, Jaeger and coworkers utilized a commercially available chaperone plasmid system (Takara Bio USA) with success.⁶⁶ *P. putida* MdlC protein was co-expressed with chaperone proteins in *E. coli* and increased *in vitro* enzyme activity was reported.⁶⁶

To construct a plasmid harboring the *P_{tac}-groEL-groES* gene locus, a 3.4 kb fragment carrying the *groEL-groES* chaperones and the trigger factor (TF) genes was amplified from the plasmid pG-Tf2 (Takara Bio USA). The PCR product was digested with *Hind*III and ligated with *Hind*III digested pJF118EH to afford the 8.8 kb plasmid pML7.166. This subcloning step replaced the tetracycline induced promoter Pzt1 in the original plasmid with a *tac* promoter, thus avoided the use of light sensitive tetracycline for induction during fermentation.

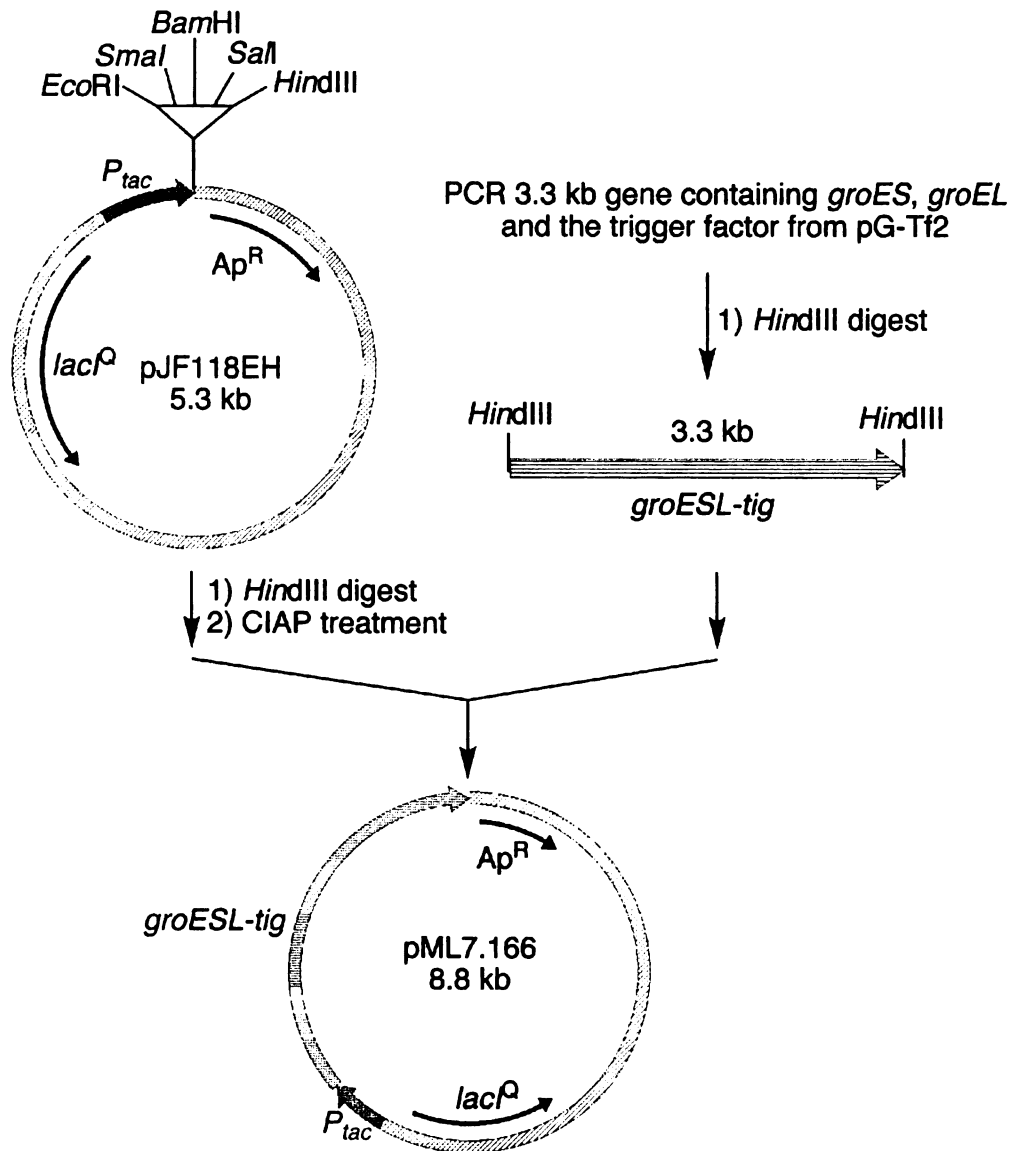


Figure 63. Construction of plasmid pML7.166.

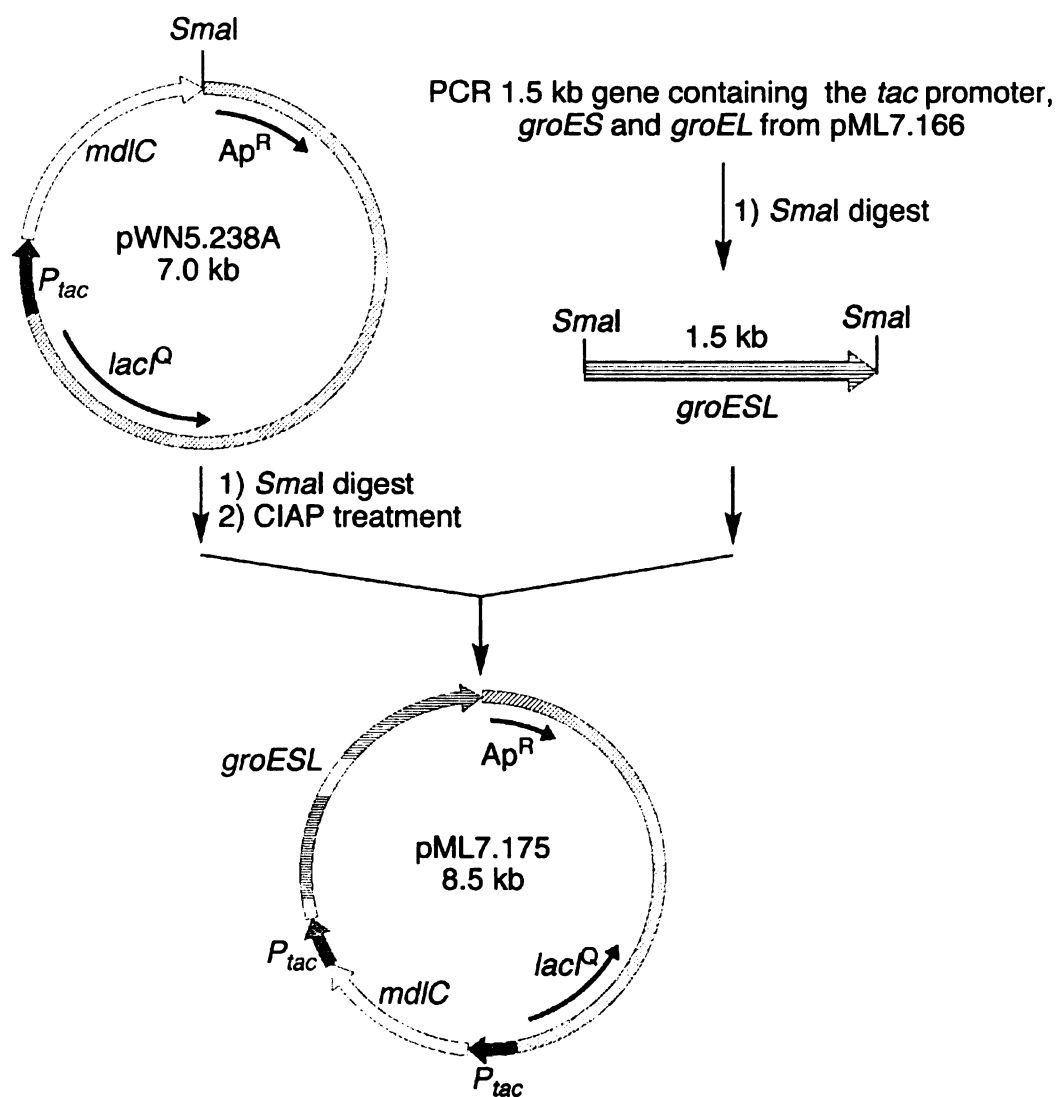


Figure 64. Construction of plasmid pML7.175.

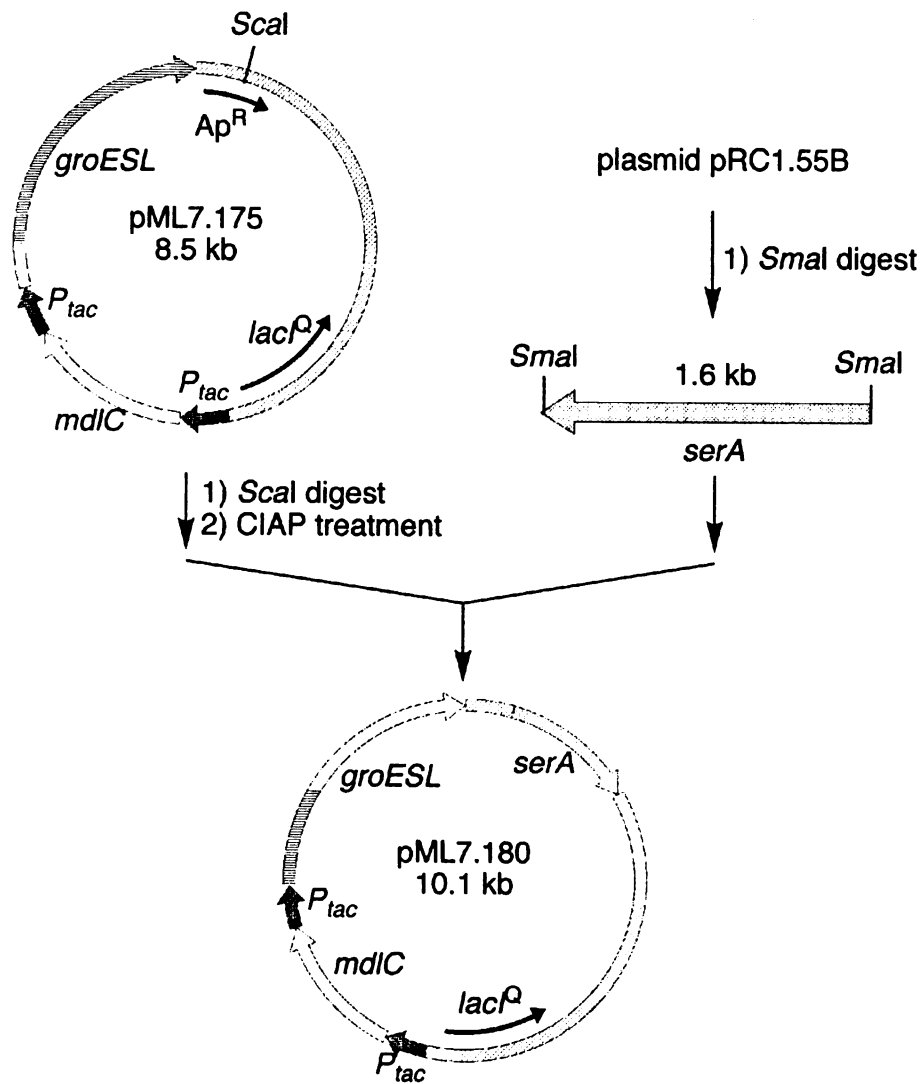


Figure 65. Construction of plasmid pML7.180.

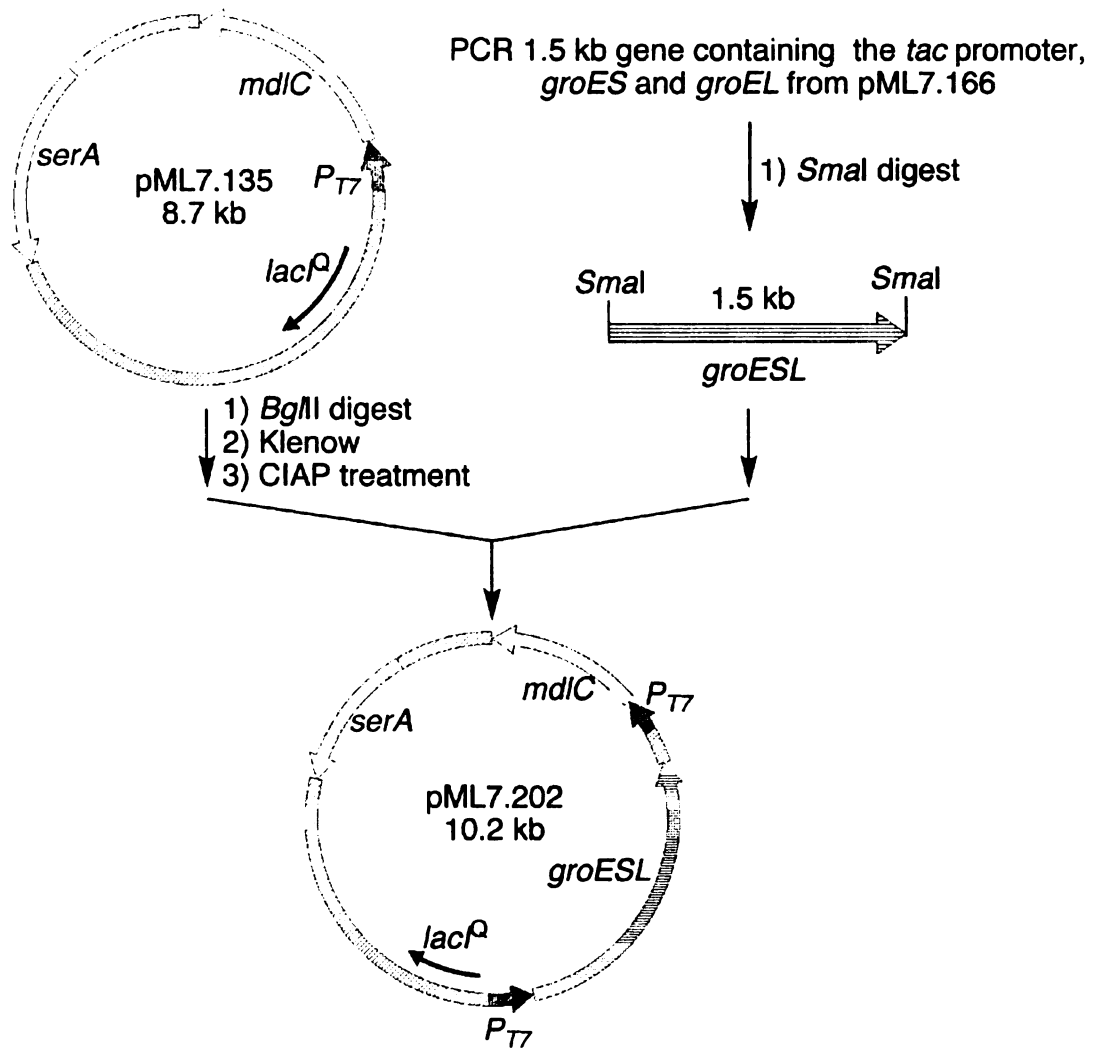


Figure 66. Construction of plasmid pML7.202.

The 1.5 kb *P_{lac}-groEL-groES* locus was amplified by PCR from the plasmid pML7.166. The PCR product was digested with *Sma*I and ligated to *Sma*I digested pWN5.238A to yield plasmid pML7.175. A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the *Sca*I digested plasmid pML7.175 to afford the 10.1 kb plasmid pML7.180. A similar strategy was used to insert the chaperone locus into a plasmid expressing *mdlC* under a phage T7 promoter. In this case, plasmid pML7.135 was used as the parent plasmid for cloning. The same 1.5 kb fragment carrying the *P_{lac}-groEL-groES* fragment obtained from PCR was digested with *Sma*I. This fragment was subsequently ligated to pML7.135, which had been treated sequentially with *Bgl*II, Klenow and CIAP, to afford upon ligation the 10.2 kb plasmid pML7.202. Plasmids pML7.180 and pML7.202 were then transformed into host strains *E. coli* strains KIT2 and KIT15, respectively, and evaluated under the same fermentation controlled condition as discussed earlier. However, similar to the case of *E. coli* KIT15/pML7.135, all D-xylose was consumed and these *E. coli* microbes synthesized 21 g/L and 20 g/L 3-deoxy-D-*glycero*-pentulosonic acid as product, respectively.

In a nutshell, neither using a strong phage T7 promoter or co-expressing *E. coli* chaperonins GroES and GroEL proteins was a successful approach in yielding a better D-1,2,4-butanetriol producing construct. Cultivation of these constructs under fermentor-controlled conditions resulted in a significant decrease in D-1,2,4-butanetriol production. The expression of *mdlC* under the T7 promoter vector system resulted in insufficient 3-deoxy-*glycero*-pentulosonate decarboxylase activity. Metabolic burden associated with using a strong promoter had led to a significant drop in cell mass compared to the microbial synthesis using our benchmark microbe WN13/pWN7.126B. Nevertheless, *E.*

coli microbes KIT15/pML7.135, KIT2/pML7.180 and KIT15/pML7.202 were able to synthesize 3-deoxy-D-*glycero*-pentulosonic acid as the major product (Figure 62).

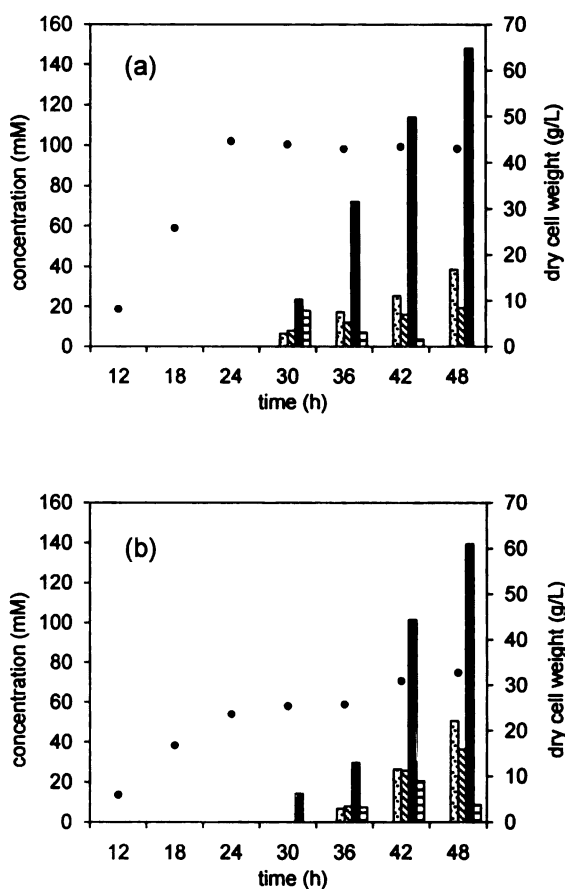


Figure 67. Cultivation of chaperonin *groEL-groES* constructs under fermentor-controlled conditions. (a) *E. coli* KIT2/pML7.180; (b) *E. coli* KIT15/ ML7.202; D-1,2,4-butanetriol; 3,4-dihydroxy-D-butanoic acid; 3-deoxy-D-*glycero*-pentanoic acid; 4,5-dihydroxy-*threo*-L-norvaline; 3-deoxy-D-*glycero*-pentulosonic acid; D-xylonic acid.

Discovery of alcohol dehydrogenases for D-1,2,4-butanetriol biosynthesis

Background

The remaining byproduct formed during microbial D-1,2,4-butanetriol synthesis was 3,4-dihydroxy-D-butanoic acid, which was derived from intermediate 3,4-dihydroxy-D-butanal. Elimination of the formation of this byproduct was pursued by identification of the enzyme and encoding gene or genes responsible for the conversion of 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol. Overexpression of the identified dehydrogenase was anticipated to reduce or eliminate formation of byproduct 3,4-dihydroxy-D-butanoic acid. Screening of NAD- and NADP- dependent 1,2,4-butanetriol dehydrogenases in *E. coli* was therefore pursued. *E. coli adhP*, *adhE* and *yiaY* encoding substrate-nonspecific alcohol dehydrogenases were found active towards using 3,4-dihydroxy-D-butanal as substrate. Among these dehydrogenases, the importance of native *E. coli* AdhP in the 1,2,4-butanetriol biosynthesis was demonstrated using a chromosomal *adhP* inactivation of *E. coli* WN13/pWN7.126B. Overexpression of *adhP* gene in an *E. coli* D-1,2,4-butanetriol producing host strain with additional 2-ketoacid dehydrogenase gene inactivations led to a second generation microbe *E. coli* KIT4/pWN7.126B with improved D-1,2,4-butanetriol production and reduced organic acid byproduct formation. Recently, varying culture parameters for this new construct resulted in the synthesis of 18 g/L D-1,2,4-butanetriol in 51% yield.

Cloning and screening for D-1,2,4-butanetriol dehydrogenase candidates

To characterize the last biosynthetic step in WN13/pWN7.126B that converts 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol, five *E. coli* NAD- or NADP- dependent

enzymes that are commonly annotated as alcohol dehydrogenases or putative dehydrogenase were screened. These enzymes are designated as AdhP, AdhE, YhdH, YiaY and YdjO in the ERGO bacterial genome database. AdhE is an iron-dependent bi-functional enzyme with both dehydrogenase and pyruvate formate lyase activities. YiaY is believed to be another iron-dependent dehydrogenase and it shares 63% amino acid sequence identity to *Zymomonas mobilis adhB* encoding alcohol dehydrogenase. YdjO is annotated as an alcohol dehydrogenase in ERGO with no significant similarity to any other known enzymes. Among these candidates, zinc-dependent AdhP and YhdH alcohol dehydrogenases shared over 80% nucleotide sequence identity with *Z. mobilis adhA* and were of particular interest due to their ability to utilize propanal as substrate. These gene candidates were independently amplified by PCR using Platinum *Taq* DNA polymerase (Invitrogen) and cloned into pKK223-3 (*adhP*, Figure 68 and *adhE*, Figure 69) and pJF118EH (*yhdH*, *yiaY* and *ydjO*, Figure 70 – 50) for protein expression. The resulting recombinant plasmids were transformed into an *E. coli* DH5 α expression host. The cell-free extract of these constructs was used for *in vitro* enzymatic assay studies. The possible reduction reaction was assayed by monitoring the depletion of NADH at 340 nm using acetaldehyde and racemic 3,4-dihydroxybutanal as substrates. No measurable enzyme activities were observed with YhdH and YdjO using both acetaldehyde and racemic 3,4-dihydroxybutanal as substrates (Table 20). AdhP, AdhE and YiaY, however, successfully catalyzed the conversion of these substrates to form ethanol and 1,2,4-butanetriol, respectively (Table 20). These results suggested AdhP, AdhE and YiaY indeed had 1,2,4-butanetriol dehydrogenase activities. Among them, enzyme AdhP showed 20 times higher activity over the other dehydrogenases and was subjected to further study.

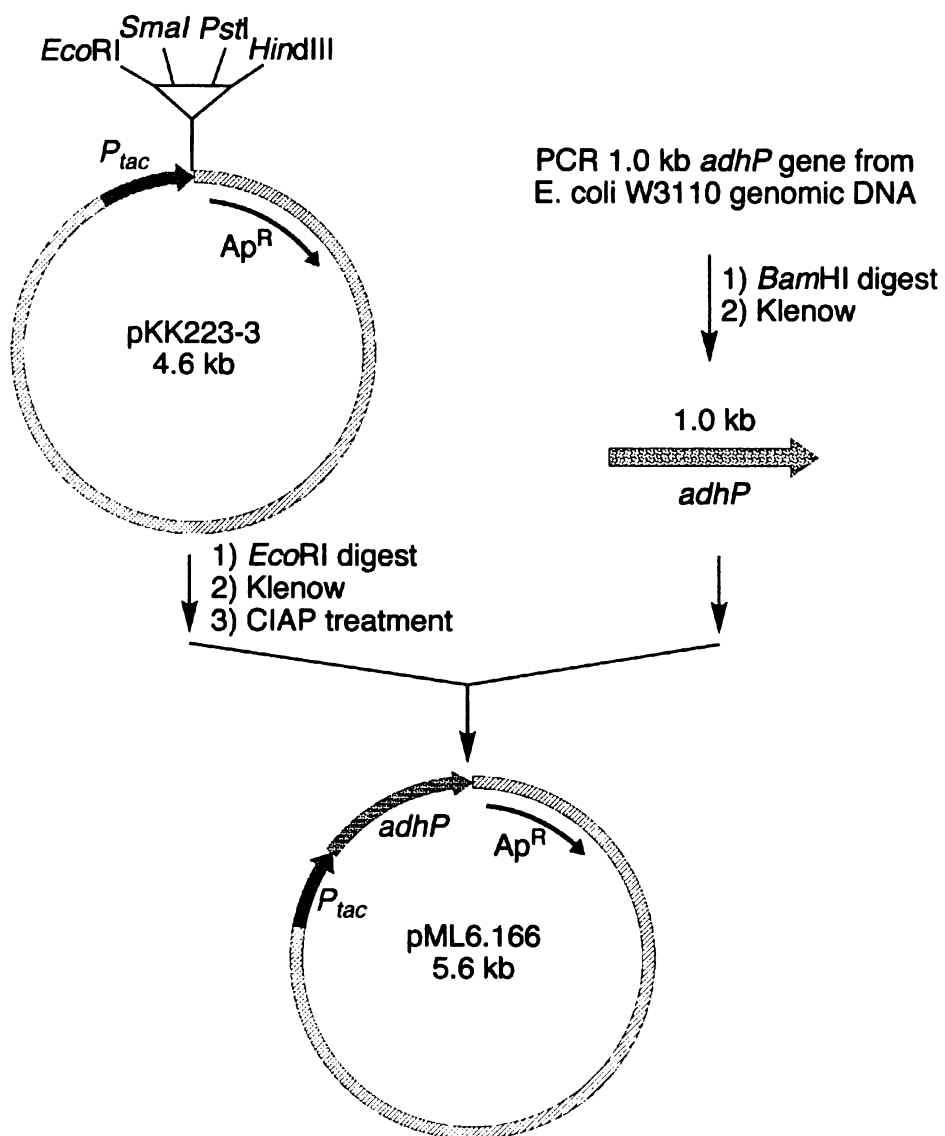


Figure 68. Construction of plasmid pML6.166.

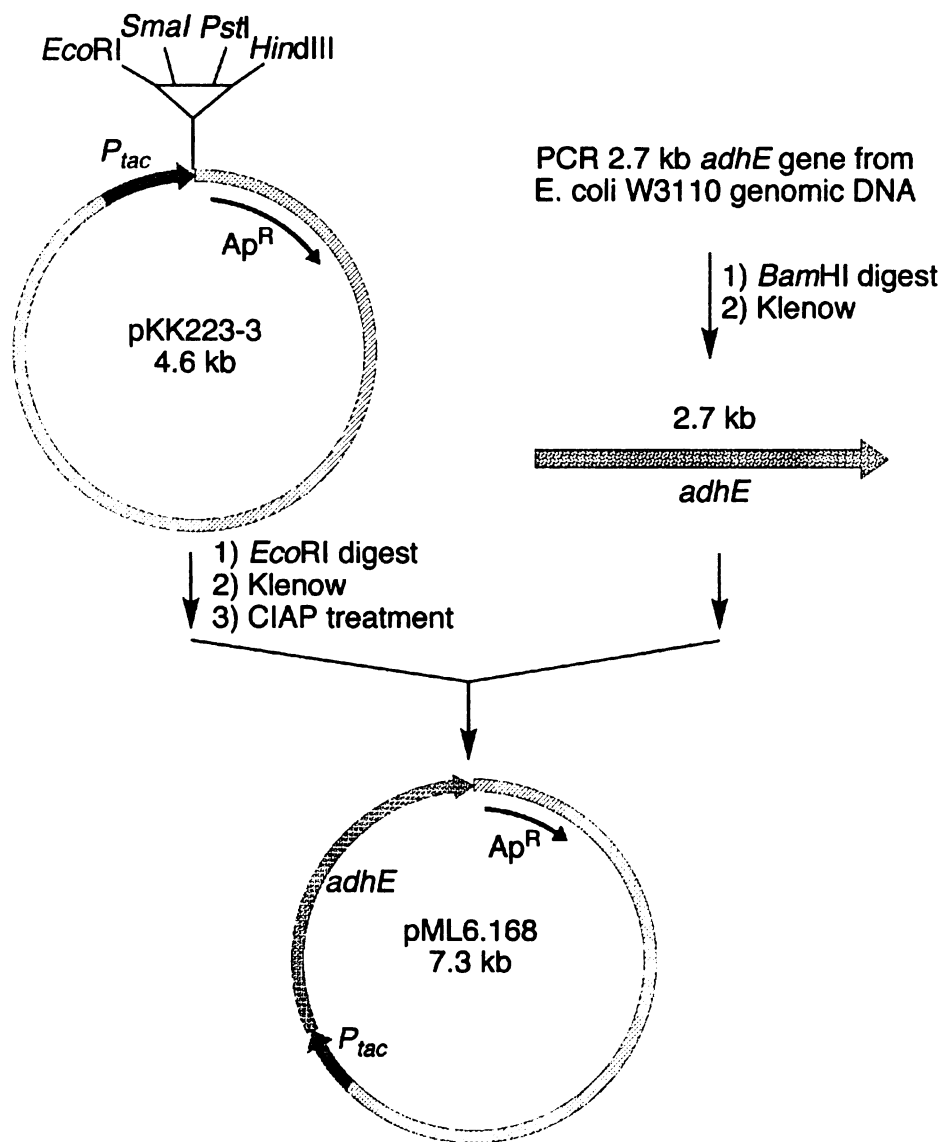


Figure 69. Construction of plasmid pML6.168.

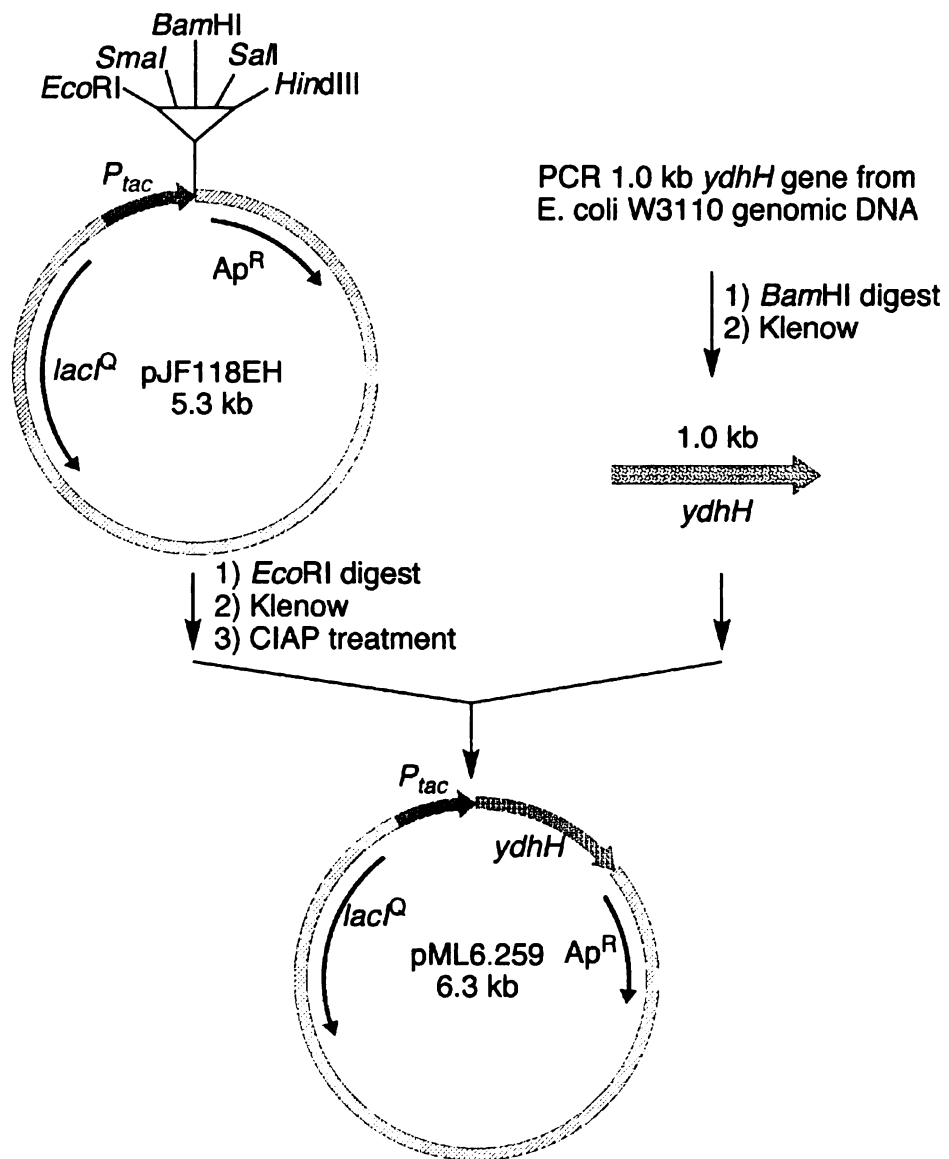


Figure 70. Construction of plasmid pML6.259.

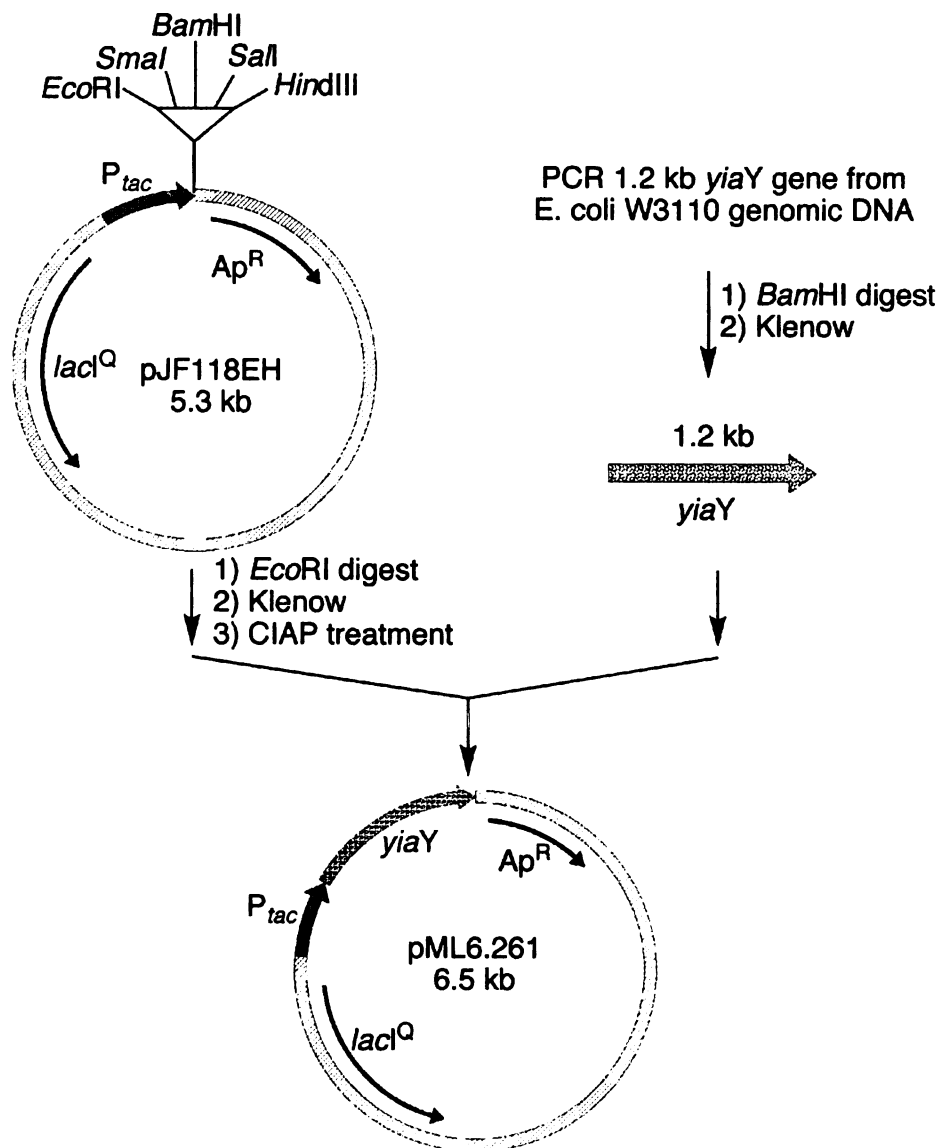


Figure 71. Construction of plasmid pML6.261.

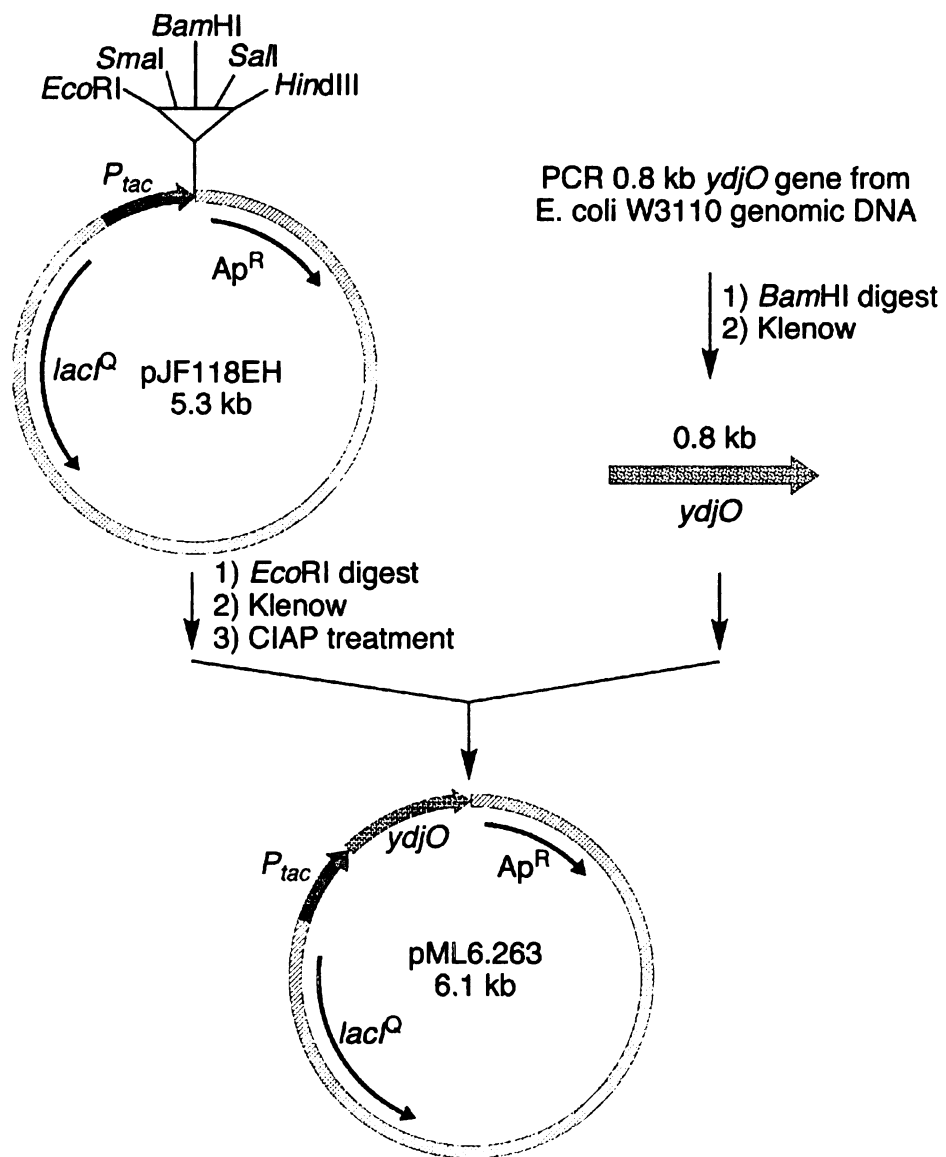


Figure 72. Construction of plasmid pML2.263.

Table 20. *E. coli* alcohol dehydrogenase candidates.

entry	gene	size (kb)	construct	crude lysate activity (U/mg)	
				acetaldehyde	3,4-dihydroxybutanal
1	<i>adhP</i>	1.0	DH5 α /pML6.166	9.6	0.8
2	<i>adhE</i>	2.7	DH5 α /pML6.168	0.06	0.02
3	<i>yhdH</i>	1.0	DH5 α /pML6.259	0	0
4	<i>yiaY</i>	1.2	DH5 α /pML6.261	0.13	0.03
5	<i>ydjO</i>	0.8	DH5 α /pML6.263	0	0

Chromosomal adhP gene inactivation

To further examine whether AdhP was responsible for *in vivo* conversion of 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol, the *E. coli* strain WN13/pWN7.126B with a chromosomal *adhP* gene inactivation was prepared. A typical deletion was introduced into *E. coli* using the Datsenko-Wanner methodology.⁶⁷ PCR primers for the *adhP* gene inactivation were designed based on 40 nt homologous sequences for this gene and 20 nt for the priming site on the template plasmid pKD3. Plasmid pKD3 contains a chloramphenicol resistant gene flanked by Flp recognition target (FRT). *E. coli* wild-type W3110 carrying plasmid pKD46 encoding λ Red recombinase was made competent and subsequently transformed with the FRT-flanked chloramphenicol linear PCR gene fragment. The transformants were selected on an LB solid agar plate containing chloramphenicol. Genomic DNA was purified from three single colonies that appeared on this plate using the CTAB method as described in the earlier chapter. Disruption of chromosomal *adhP* was confirmed by PCR using a pair of primers flanking the *adhP* locus. The resulting *E. coli* strain W3110*adhP*::FRT-Cm^R-FRT was designated *E. coli* KIT8.

This mutation was then transferred to the *E. coli* WN13/pWN7.126B by P1 phage-mediated transduction into *E. coli* KIT2 using KIT8 as a donor strain. The resulting strain *E. coli* KIT9 was selected on an LB plate containing chloramphenicol. Subsequent removal of the drug marker was achieved by thermal induction of Flp recombinase in *E. coli* KIT9/pCP20.⁶⁸ The resulting marker-free strain was designated *E. coli* KIT10. Both *E. coli* strains KIT9 and KIT10 were transformed with the *mdlC* carrying plasmid pWN7.126B and their catalytic efficiencies were evaluated under fermentor-controlled conditions.

Table 21. *E. coli* strains for chromosomal *adhP* inactivation study.

strain	genotype
WN7	W3110yjhH::FRTyagE::FRTserA
WN13	W3110yjhH::FRTyagE::FRTserAxyLAB::xdh-FRT-Cm ^R -FRT
KIT2	W3110yjhH::FRTyagE::FRTserAxyLAB:: xdh-FRT
KIT8	W3110adhP::FRT-Cm ^R -FRT
KIT9	W3110yjhH::FRTyagE::FRTserAxyLAB:: xdh-FRT adhP::FRT-Cm ^R -FRT
KIT10	W3110yjhH::FRTyagE::FRTserAxyLAB:: xdh-FRT adhP::FRT

E. coli KIT10/pWN7.126B produced only 6.5 g/L D-1,2,4-butanetriol in 31% yield under fermentor-controlled conditions. D-3,4-Dihydroxybutanoic acid accumulated to 5.3 g/L in the culture medium, which accounted for a 15% increase in concentration compared to its parent microbe *E. coli* WN13/pWN7.126B. This result suggests that *E. coli* AdhP is indeed an *in vivo* D-1,2,4-butanetriol dehydrogenase. Since KIT10/pWN7.126B was still able to synthesize D-1,2,4-butanetriol, it therefore becomes tempting to speculate that one or more other native *E. coli* dehydrogenases may reduce 3,4-dihydroxy-D-butanal to D-1,2,4-butanetriol *in vivo*.

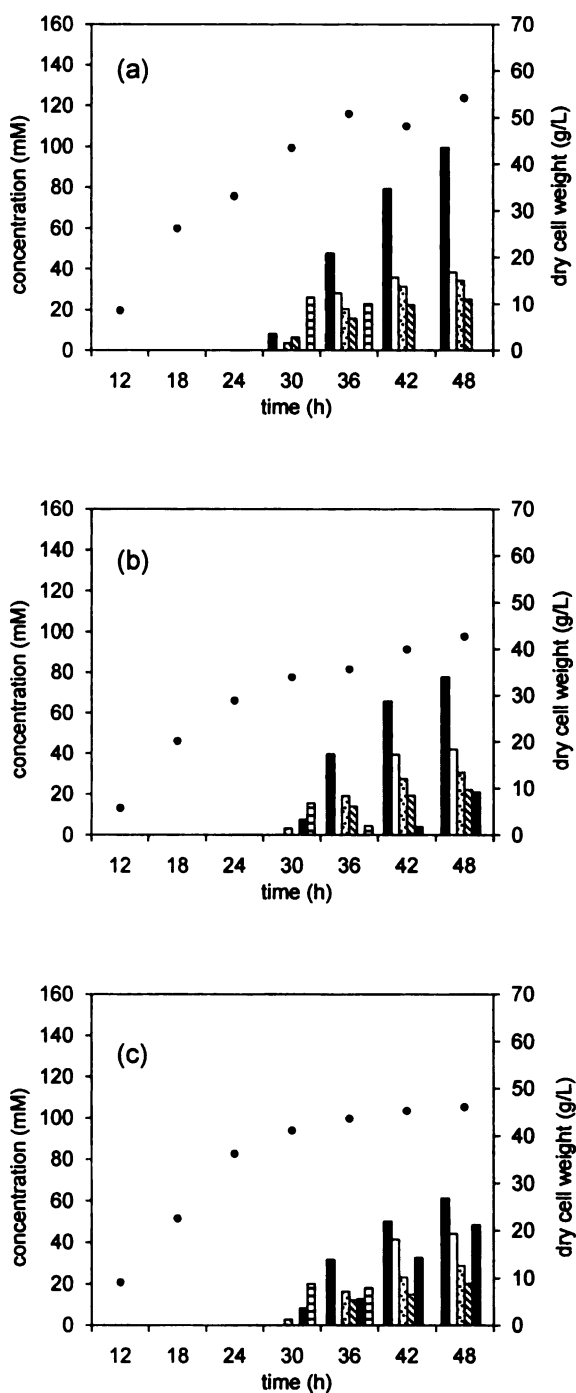


Figure 73. Cultivation of *E. coli* strains with *adhP* gene inactivation under fermentor-controlled conditions. (a) *E. coli* WN13/pWN7.126B; (b) *E. coli* KIT9/pWN7.126B; (c) *E. coli* KIT10/pWN7.126B; ■ D-1,2,4-butanetriol; □ 3,4-dihydroxy-D-butanoic acid; ▤ 3-deoxy-D-glycero-pentanoic acid; ▨ 4,5-dihydroxy-threo-L-norvaline; ▩ 3-deoxy-D-glycero-pentulosonic acid; ▧ D-xylonic acid.

Plasmid-based expression of *adhP* gene

Further understanding of the dehydrogenation reaction in the D-1,2,4-butanetriol biosynthetic pathway offered us an opportunity to improve the performance of our benchmark *E. coli* WN13/pWN7.126B. Three additional alcohol dehydrogenase candidates identified in other microorganisms were screened in an attempt to obtain a better wild-type D-1,2,4-butanetriol dehydrogenase. *Z. mobilis* AdhA and AdhB shares 80% and 63% amino acid identity to *E. coli* AdhP and YiaY, respectively. *Klebsiella pneumoniae* DhaT uses 1,3-propanediol as its native substrate and is of particular interest because of the structural similarity between 1,3-propanediol and 1,2,4-butanetriol. *Z. mobilis* alcohol dehydrogenase candidates *adhA* and *adhB* were localized in plasmids pLOI135 and pLOI295, respectively. *K. pneumoniae dhaT* was amplified from *K. pneumoniae* ATCC25955 genomic DNA and subsequently cloned into pJF118EH to yield pML7.179. These plasmid-localized genes were expressed in *E. coli* DH5 α and the enzyme activities were determined using the same spectrophotometric assay as described earlier (Table 22). Crude lysate of AdhA and DhaT showed activities of 0.06 and 0.003 U/mg, respectively, using 3,4-dihydroxybutanal as substrate. AdhB showed no detectable activity. The enzyme activities toward their native substrates were also determined.

Table 22. Other alcohol dehydrogenase candidates

entry	gene	size (kb)	construct	crude lysate activity (U/mg)	
				native substrate	3,4-dihydroxybutanal
1	<i>adhA</i>	1.2	DH5 α /pLOI135	73 ^a	0.06
2	<i>adhB</i>	1.8	DH5 α /pLOI295	50 ^a	<0.001
3	<i>dhaT</i>	1.2	DH5 α /pML5.179	4.3 ^b	0.003

^aEthanol was used as substrate, NAD was used as cofactor.

^b1,3-propanediol was used as substrate, NAD was used as cofactor.

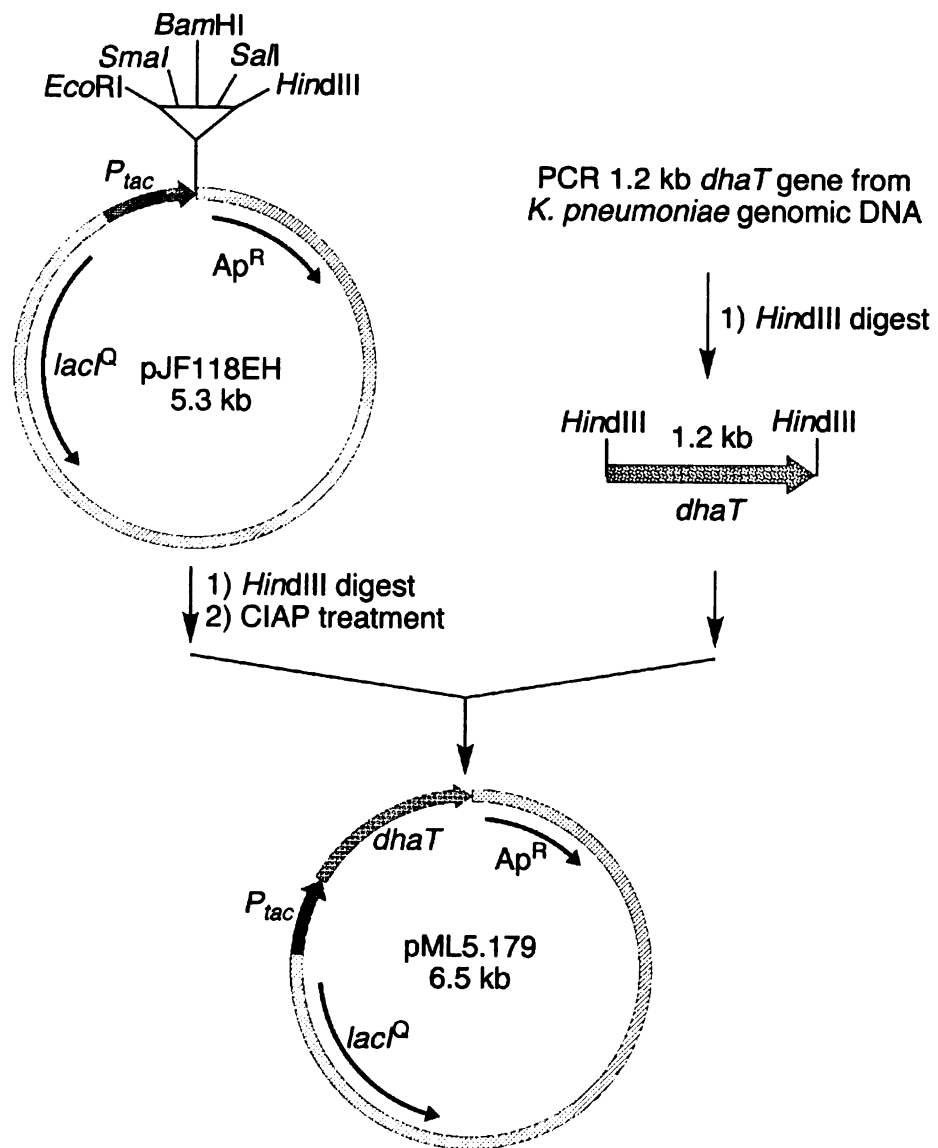


Figure 74. Construction of plasmid pML5.179.

E. coli adhP, *Z. mobilis adhA* and *K. pneumoniae dhaT* were subsequently cloned into plasmid pWN7.126B. The effect of expressing these alcohol dehydrogenases during the biosynthesis of D-1,2,4-butanetriol were examined under fermentor-controlled conditions. A 1.3 kb fragment carrying the *tac* promoter region and the *adhP* gene was excised from plasmid pML6.166 with *Bam*HI and ligated to *Bam*HI digested pWN5.238A to afford plasmid pML6.185. A 1.6 kb *serA* locus excised from plasmid pRC1.55B with *Sma*I was ligated with *Sca*I treated pML6.185 to afford plasmid pML6.195. On the other hand, plasmid pWN6.096B containing *Z. mobilis adhA* and *mdlC* gene fragments was obtained from the Frost group strain collection. The same *serA* locus was ligated to *Sca*I digested pWN6.096B to yield plasmid pML6.133. Finally, the *K. pneumoniae dhaT* gene was excised from plasmid pML5.179 by treating with *Bam*HI and Klenow. This gene fragment was subsequently ligated to pKK223-3, which was sequentially treated with *Eco*RI and Klenow to yield plasmid pML6.090. The *P_{tac}-dhaT* locus was excised with *Bam*HI and ligated to *Bam*HI treated pWN5.238A to afford plasmid pML6.128. Inserting the *Sma*I digested *serA* gene into the *Sca*I site of this plasmid yielded plasmid pML6.135. The resulting *mdlC* genes and dehydrogenase candidates in these three plasmids were transcribed in the opposite direction under a separate *tac* promoter. Transcription of both genes was controlled by the LacI^Q repressor protein encoded in the *lacI*^Q gene and initiated by addition of IPTG. With transformation of these plasmids into *E. coli* host WN13, constructs WN13/pML6.195, WN13/pML6.133 and WN13/pML6.135 were obtained and evaluated in fermentors using the same cultivation conditions as described earlier.

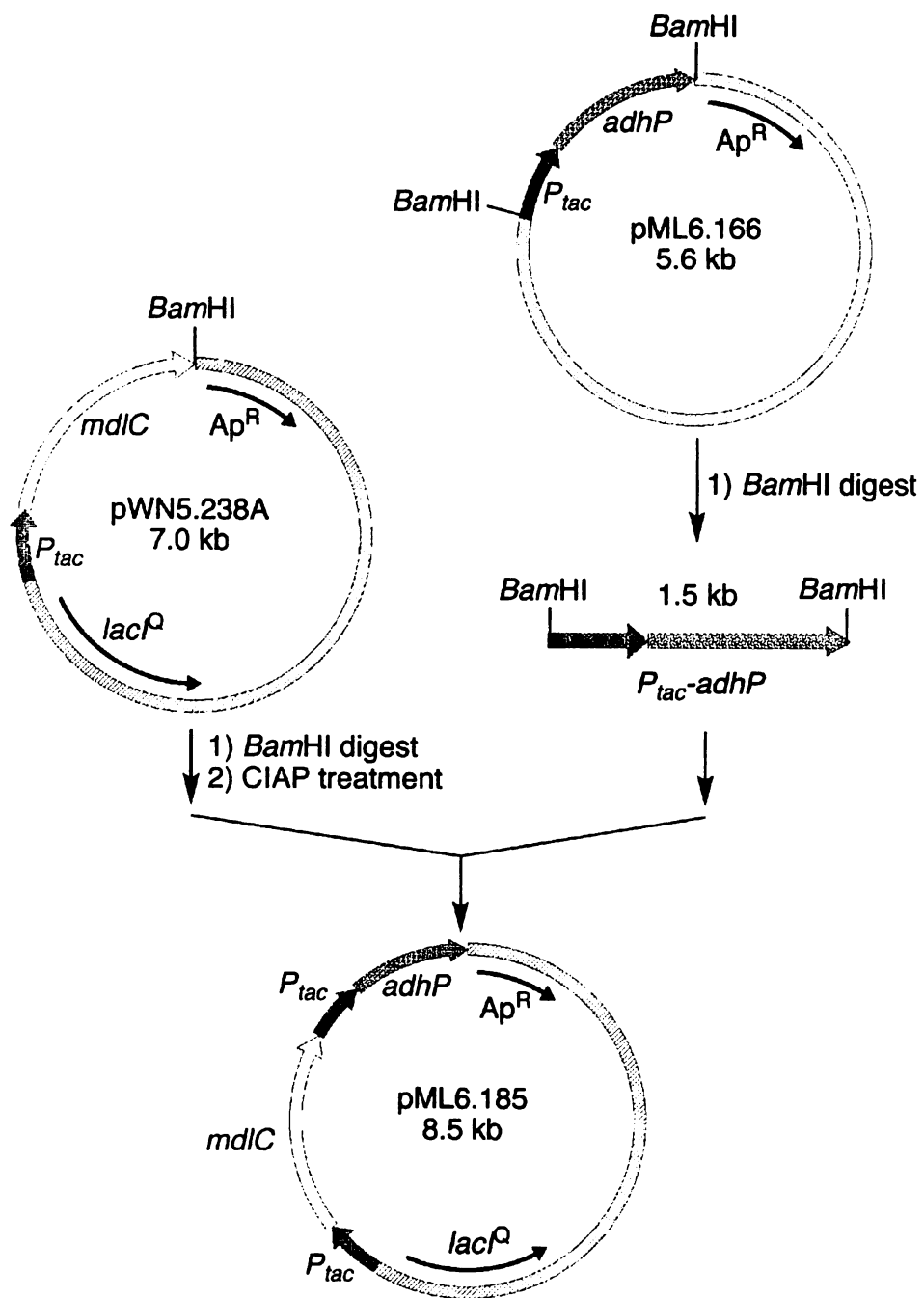


Figure 75. Construction of plasmid pML6.185.

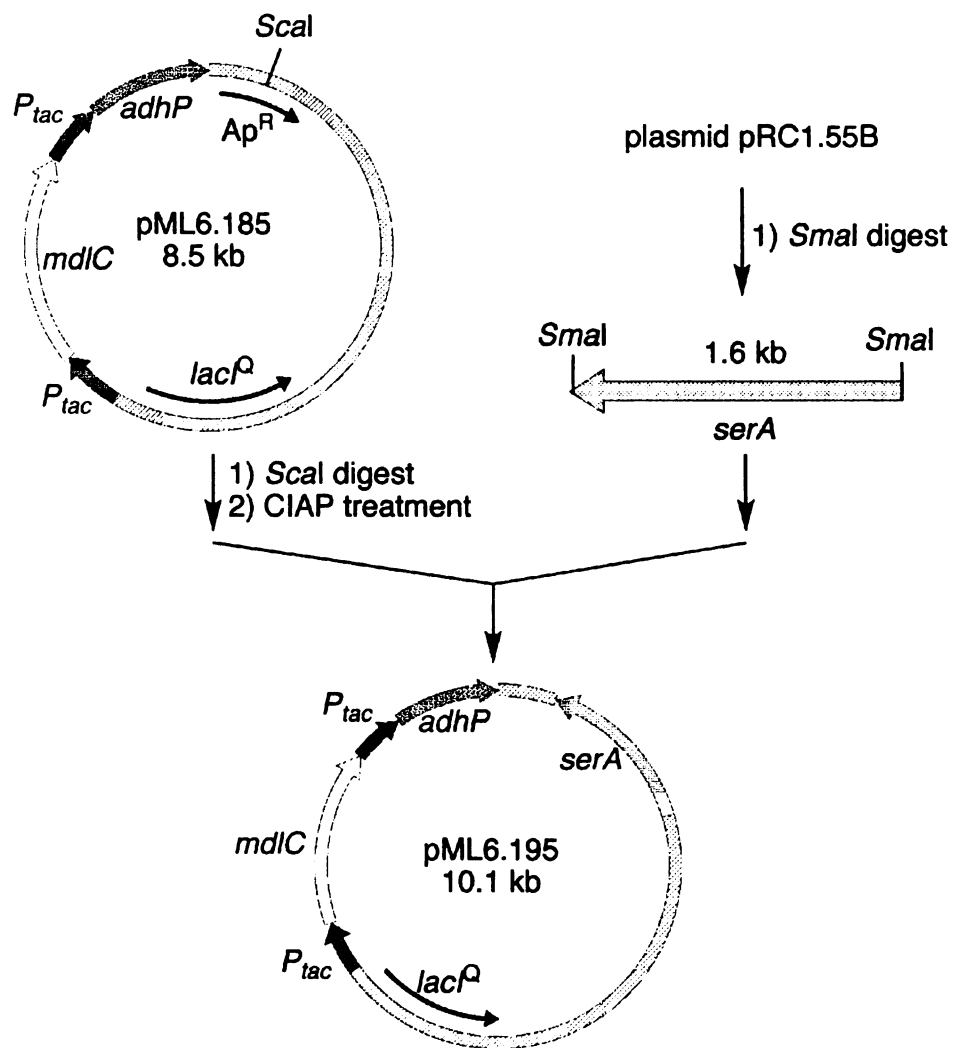


Figure 76. Construction of plasmid pML6.195.

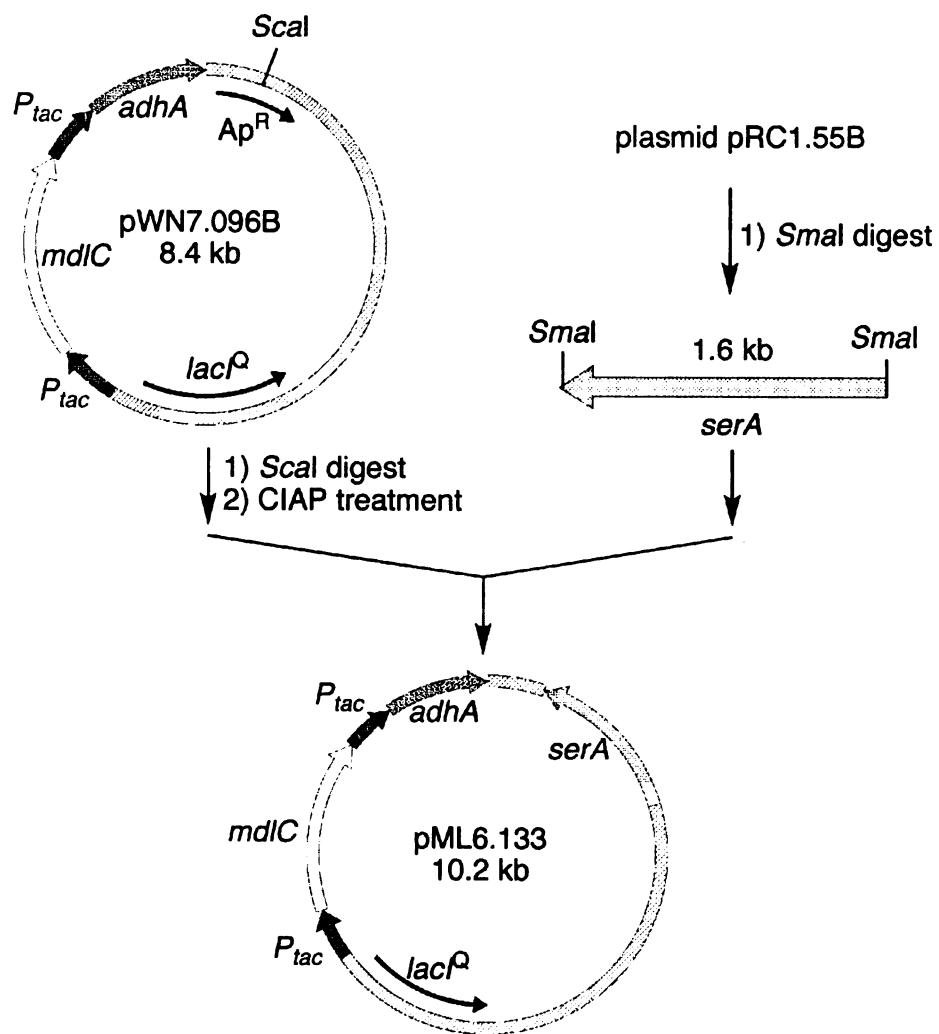


Figure 77. Construction of plasmid pML6.133.

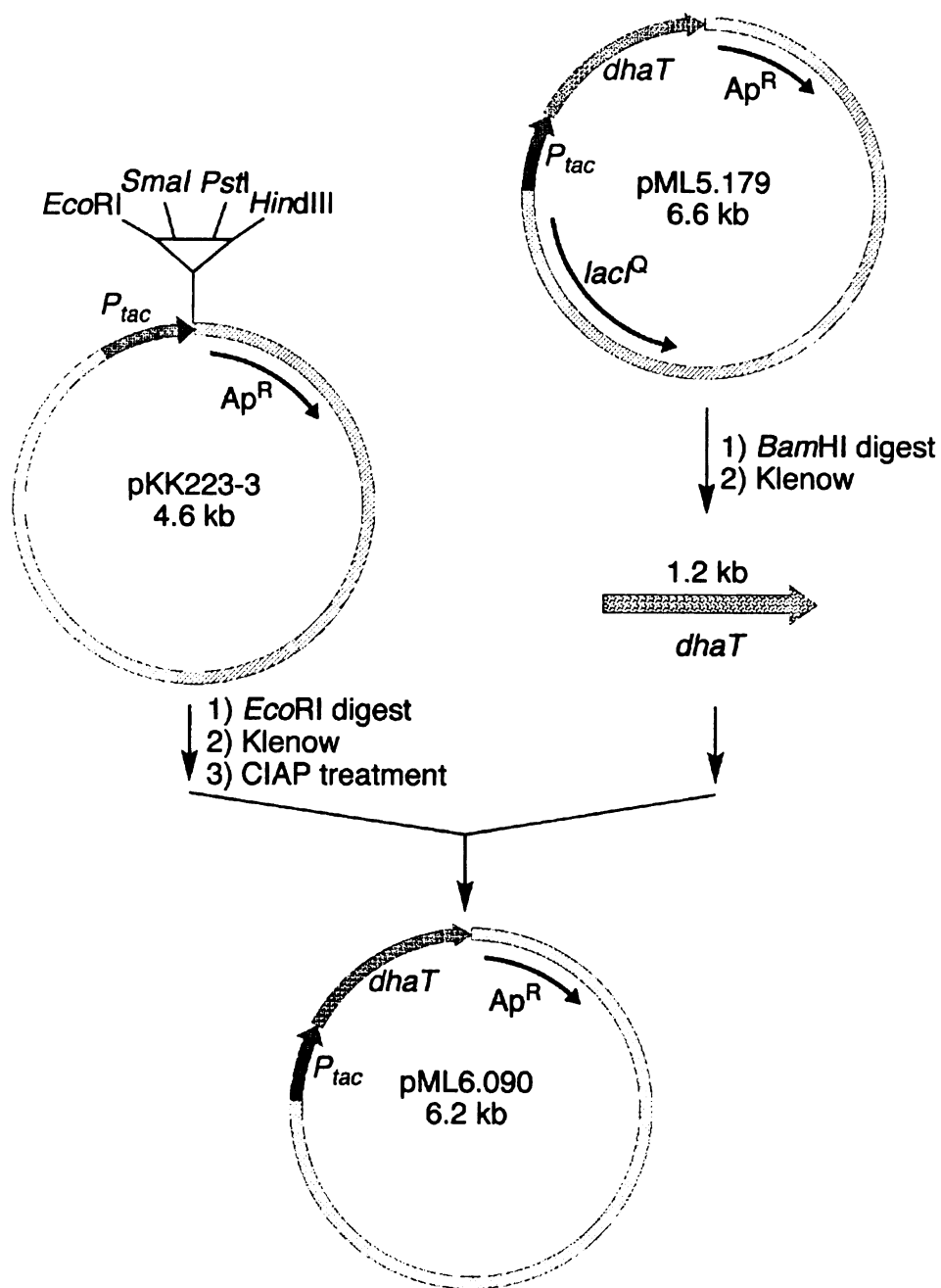


Figure 78. Construction of plasmid pML6.090.

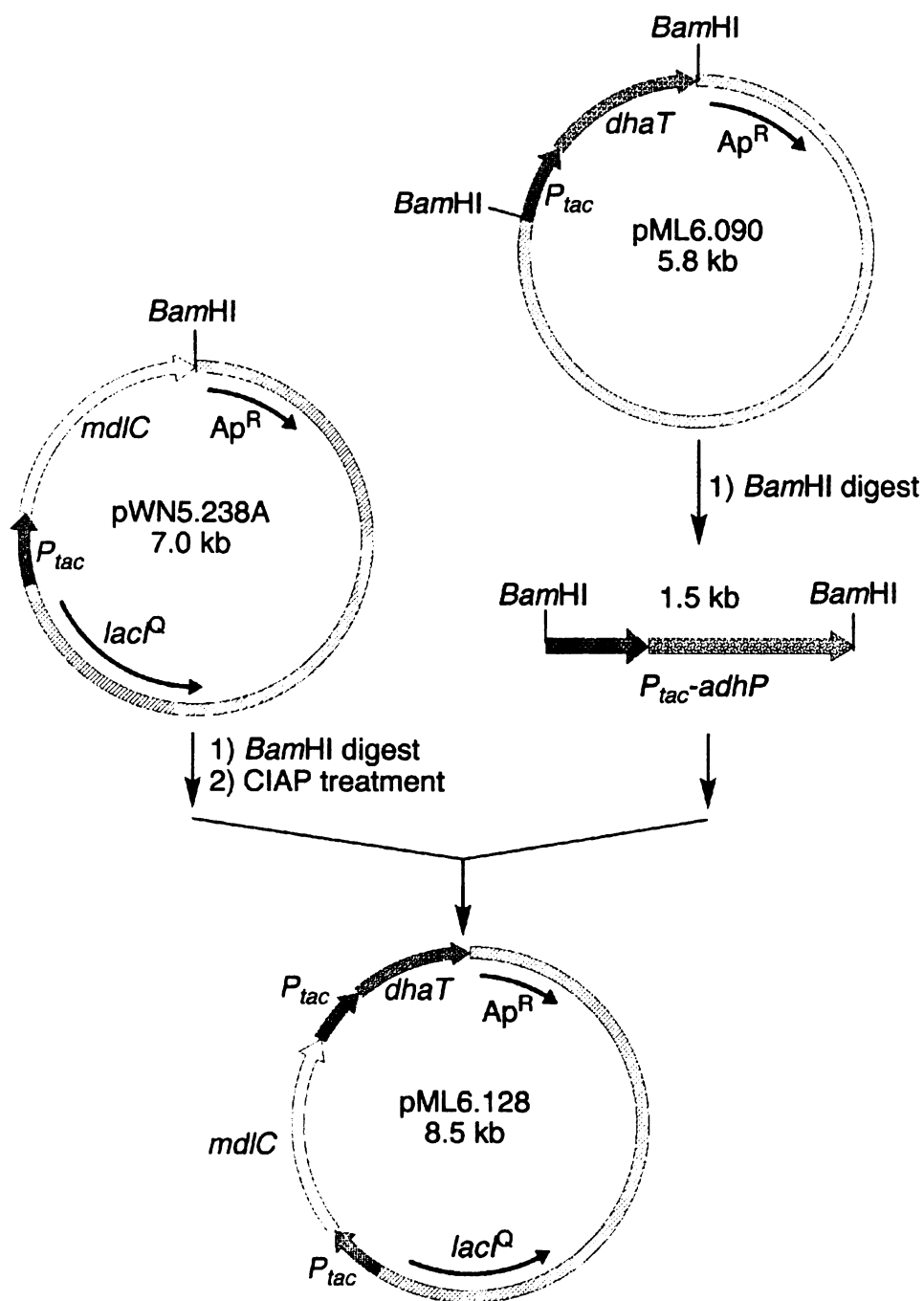


Figure 79. Construction of plasmid pML6.128.

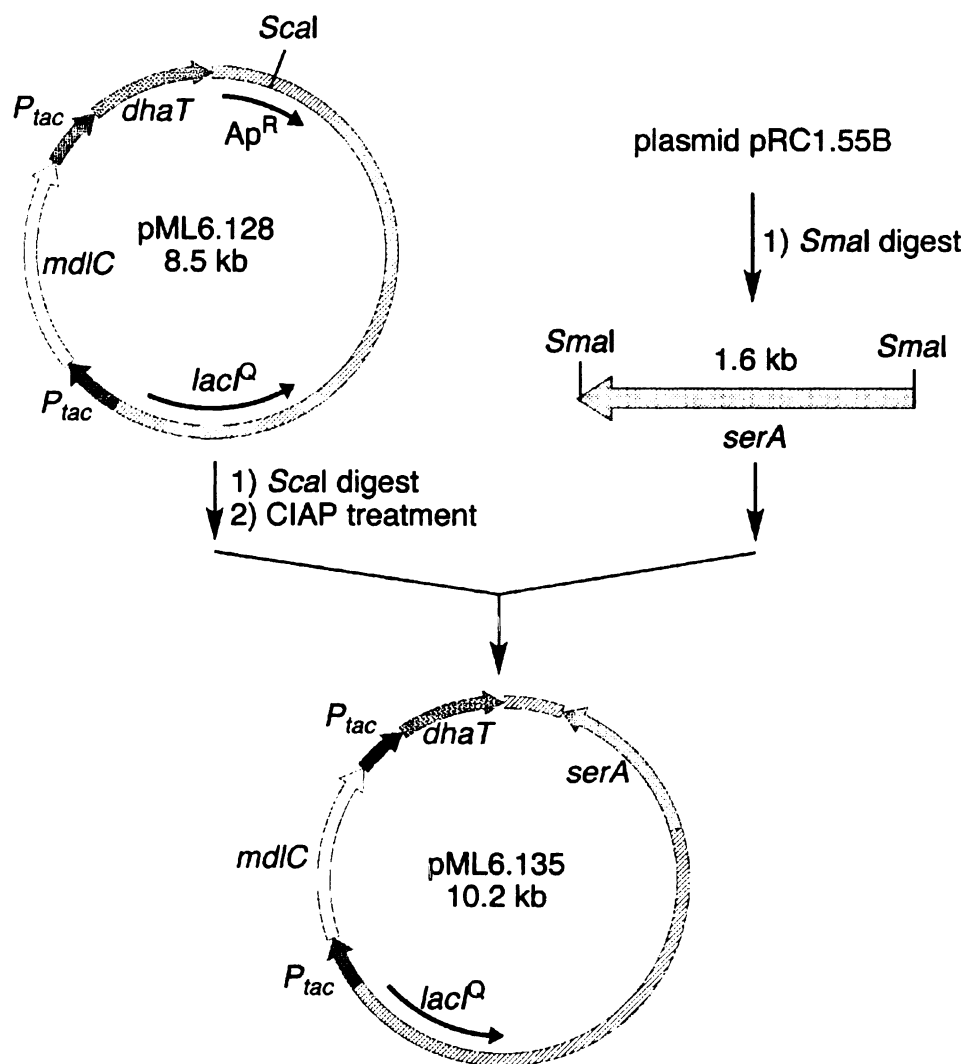


Figure 80. Construction of plasmid pML6.135.

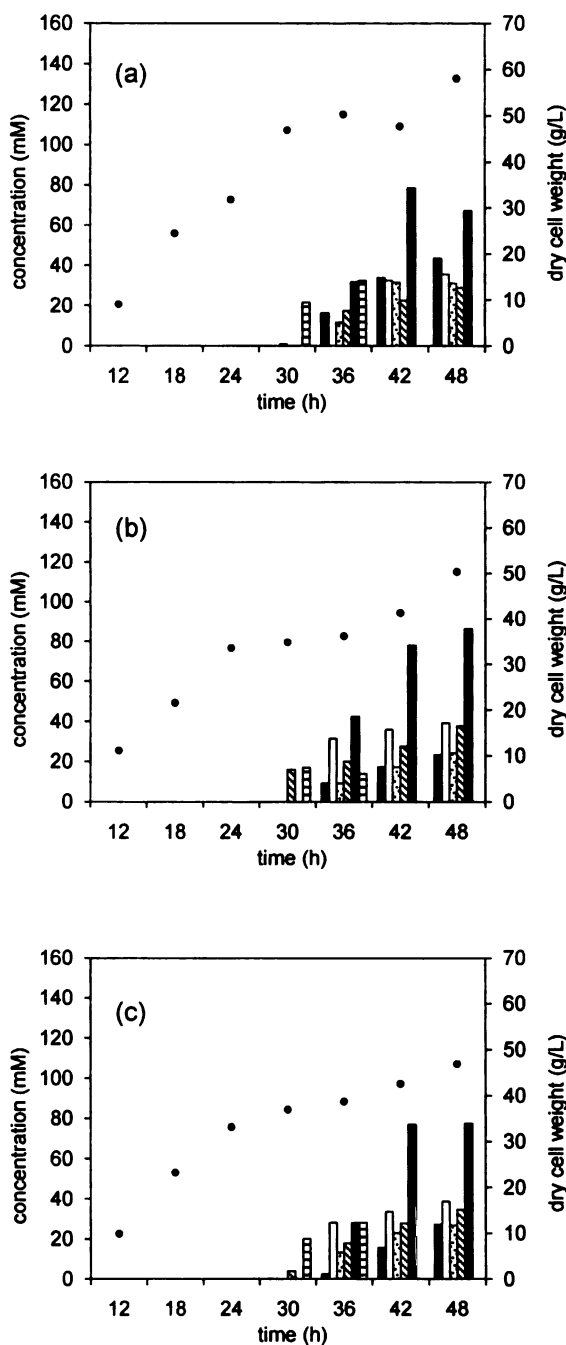


Figure 81. Cultivation of *E. coli* constructs expressing plasmid-borne alcohol dehydrogenases under fermentor-controlled conditions. (a) *E. coli* WN13/pML6.195; (b) *E. coli* WN13/pML6.133; (c) *E. coli* WN13/pML6.135; ■ D-1,2,4-butanetriol; □ 3,4-dihydroxy-D-butanoic acid; ▨ 3-deoxy-D-glycero-pentanoic acid; ▩ 4,5-dihydroxy-threo-L-norvaline; ▤ 3-deoxy-D-glycero-pentulosonic acid; ▥ D-xylonic acid.

As shown in Figure 81, *E. coli* microbes WN13/pML6.195, WN13/pML6.133 and WN13/pML6.135 synthesized D-1,2,4-butanetriol at a concentration of 4.6 g/L (22% yield, mol/mol), 2.5 g/L (12% yield, mol/mol) and 2.9 g/L (14% yield, mol/mol), respectively. Overexpressing plasmid-localized *adhP* in these constructs led to poor cell growth and decreased benzoylformate decarboxylase activities as indicated by the accumulation of 3-deoxy-D-*glycero*-pentulosonic acid. In addition, a high glutamic acid concentration was observed in ¹H NMR spectra of clarified culture medium at different time intervals. Glutamic acid accumulation has long been known to associate with hyperosmotically stressed *E. coli* and *Streptomyces typhimurium*.⁶⁹ According to the literature, it is advantageous in terms of stability, genetic regulation, and metabolic burden to modulate expression of relevant genes on the chromosome rather than relying on overexpressing these genes on a multi-copy plasmid.⁷⁰ To strive for a balance between getting higher enzyme activity and avoiding stress to cells, chromosomal insertion of an alcohol dehydrogenase on the chromosome was pursued. Only the *E. coli adhP* gene was employed in this study to avoid potential complication associated with heterologously expressing *Z. mobilis adhA* and *K. pneumoniae dhaT* in *E. coli*.

Chromosomal expression of E. coli adhP

Two strategies were pursued to increase expression of AdhP activity in *E. coli* chromosomally. *E. coli* KIT3 was derived from *E. coli* WN7 by replacing the genomic copy of the *xylA**xylB* gene cluster with an *xdh*(*Caulobacter crescentus*)-*adhP*-*P_{tac}*-FRT-Cm^R-FRT gene cassette. The *xylA* gene encodes D-xylose isomerase, while the *xylB* gene encodes D-xylulose kinase. These are the two enzymes essential for *E. coli* catabolism of D-xylose. *C. crescentus xdh* gene encoding D-xylose dehydrogenase was identified in a

previous study to convert D-xylose into D-xylonic acid.³⁶ This gene was successfully incorporated in our benchmark construct *E. coli* WN13/pWN7.126B to synthesize D-1,2,4-butanetriol from D-xylose. The proposed chromosomal modification in *E. coli* KIT3 therefore abolished its ability to utilize D-xylose as a sole carbon source for growth. As a second consequence, *E. coli* KIT3 could express a D-xylose dehydrogenase activity under the control of the *xylA* promoter, while the alcohol dehydrogenase encoding gene *adhP* was expressed under an independent *tac* promoter supplied in the cassette. The *xdh* and *adhP* genes would transcribe in opposite directions.

A typical chromosomal gene insertion in *E. coli* was achieved using a modified Datsenko-Wanner methodology (Figure 82). First, the *P_{tac}-adhP* gene locus was amplified from plasmid pML6.166 using Platinum *Taq* DNA polymerase (Invitrogen) and subsequently cloned into the *Afl*III site of plasmid pKD3 to yield the plasmid pML6.255 (Figure 83). The template plasmid containing both *adhP* and *xdh* genes was obtained by cloning the *xdh* gene into the *Sph*I site of pML6.255. The resulting plasmid was designated pML6.272 (Figure 84). PCR primers for the gene insertion experiment were then designed based on 40 nt homologous sequence for this gene and 20 nt for the priming site on the template plasmid pML6.272. *E. coli* wild-type W3110 carrying plasmid pKD46 encoding the λ Red recombinase was made competent and transformed with this gene cassette. The transformants were selected on an LB solid agar plate containing chloramphenicol. Chromosomal gene insertion was verified by PCR using a pair of primers flanking the *xylA**xylB* gene locus. The resulting *E. coli* strain was designated *E. coli* KIT1. The mutation in *E. coli* KIT1 was then transferred to *E. coli* WN7 by P1 phage-mediated transduction. The resulting *E. coli* KIT3 was selected on an LB plate

containing chloramphenicol. Subsequent removal of the drug marker was achieved by thermal induction of Flp recombinase in KIT3/pCP20 to yield *E. coli* strain KIT4. Both KIT3 and KIT4 were transformed with the plasmid pWN7.126B carrying the *mdlC* gene and evaluated under fermentor-controlled conditions (Table 23 and Figure 86).

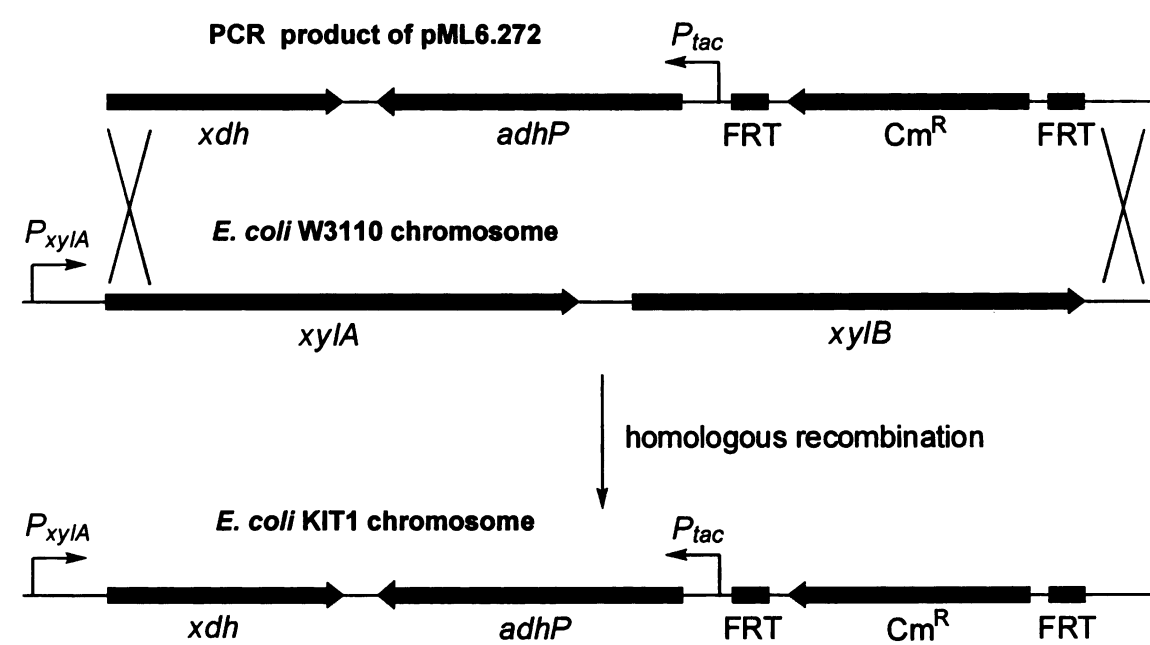


Figure 82. Strategy to integrate *xdh-adhP-P_{tac}* gene cassette into *E. coli* W3110.

Table 23. *E. coli* strains for chromosomal *adhP* expression study.

strain	genotype
WN7	W3110 <i>yjhH::FRTyagE::FRTserA</i>
WN13	W3110 <i>yjhH::FRTyagE::FRTserAxylAB::xdh-FRT-Cm^R-FRT</i>
KIT1	W3110 <i>xylAB::xdh-adhP-P_{tac}-FRT-Cm^R-FRT</i>
KIT3	W3110 <i>yjhH::FRTyagE::FRTserAxylAB::xdh-adhP-P_{tac}-FRT-Cm^R-FRT</i>
KIT4	W3110 <i>yjhH::FRTyagE::FRTserAxylAB::xdh-adhP-P_{tac}-FRT</i>
KIT5	W3110 <i>P_{T5}-FRT-Cm^R-FRT</i>
KIT6	W3110 <i>yjhH::FRTyagE::FRTserAxylAB::xdh-FRT P_{T5}-FRT-Cm^R-FRT</i>
KIT7	W3110 <i>yjhH::FRTyagE::FRTserAxylAB::xdh-FRT P_{T5}-FRT</i>

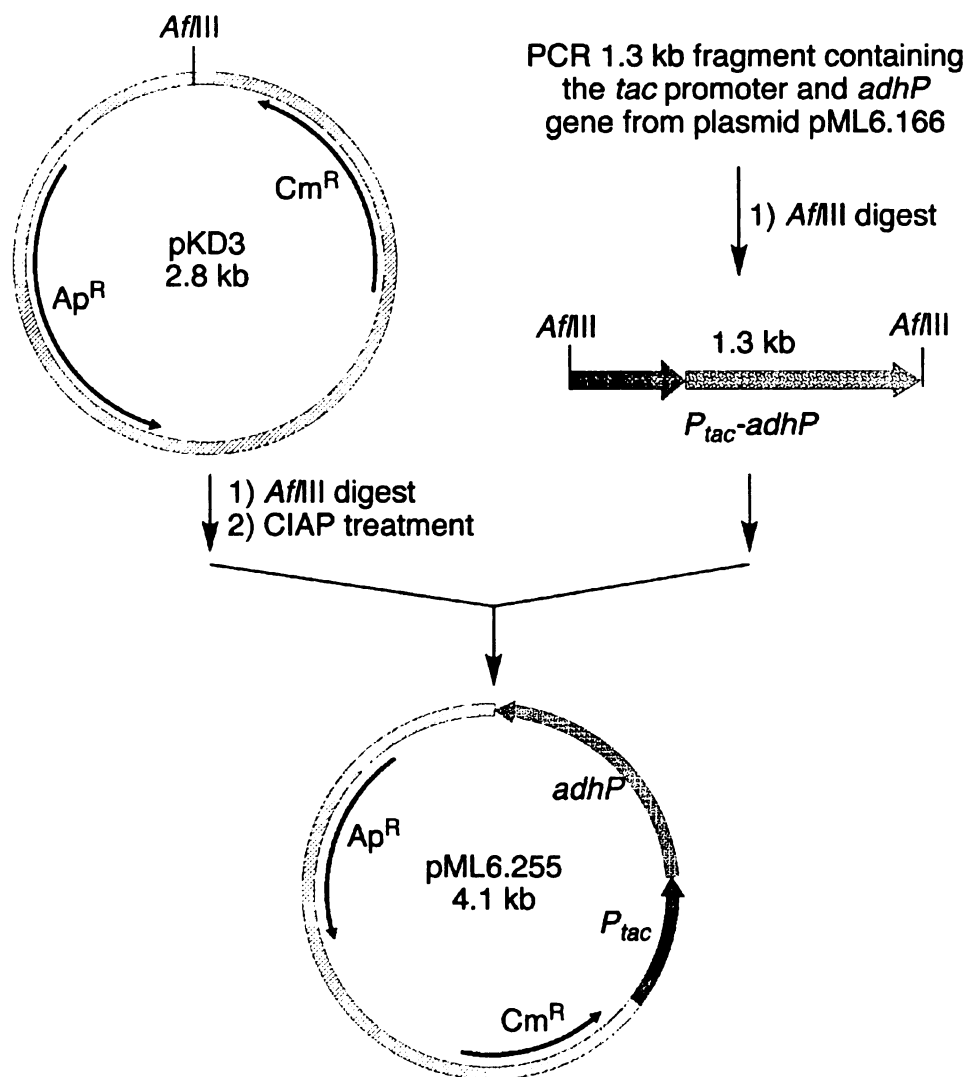


Figure 83. Construction of plasmid pML6.255.

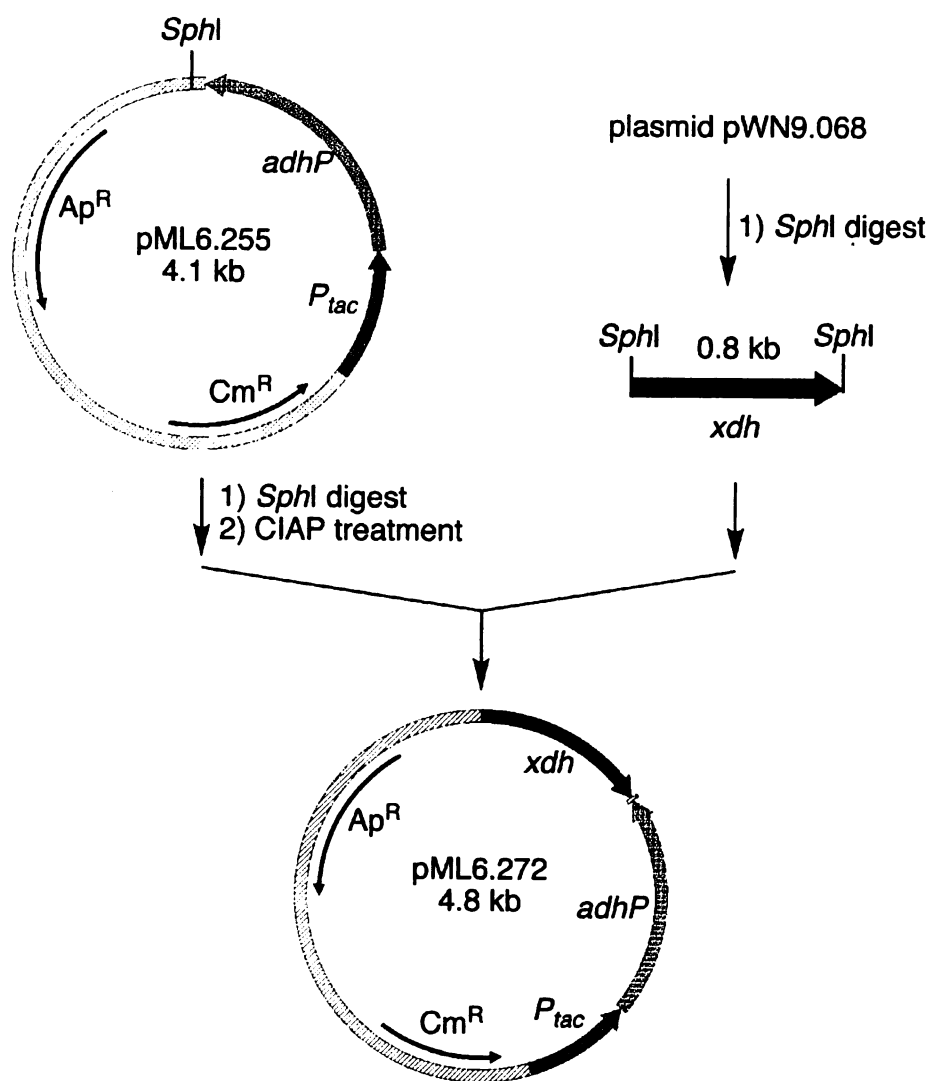


Figure 84. Construction of plasmid pML6.272.

The second *E. coli* strain was derived from *E. coli* WN13. Similarly, the modified Datsenko-Wanner methodology was employed to replace the native *adhP* promoter with the gene locus containing the phage T5 promoter region, the *lacO* gene and the ribosome binding site obtained from plasmid pQE30. This strategy was particularly appealing because only a minimal chromosomal modification was required to improve native *adhP* expression. A gene cassette P_{T5} -FRT-Cm^R-FRT was constructed by cloning the P_{T5} locus into the *Nde*I site of plasmid pKD3 to yield plasmid pML7.042 (Figure 85). The resulting chloramphenicol resistant gene is transcribed in an opposite direction as the phage T5 promoter. This plasmid pML7.042 also delivers a ‘universal’ phage T5 cassette for other native gene expression experiments along the *E. coli* genome.

To replace the native *E. coli adhP* promoter with this phage T5 cassette along with a chloramphenicol marker, the phage T5 cassette was amplified by PCR and subsequently transformed into electrocompetent *E. coli* W3110. The transformants were subject to selection on LB plates containing chloramphenicol. The resulting strain was verified by PCR and designated as *E. coli* KIT5. Strain KIT6 was prepared from *E. coli* WN13 by P1 phage-mediated transduction using KIT5 as the donor strain. This *E. coli* KIT6 was selected on the same LB/Cm plate and its genotype was verified by PCR. Using the same recombination method, thermal induction of Flp recombinase of KIT6/pCP20 excised the chloramphenicol marker and afforded strain KIT7. *E. coli* KIT7 contains only a 0.2 kb genomic insert. Both *E. coli* KIT6 and KIT7 were transformed with plasmid pWN7.126B and evaluated under fermentor-controlled conditions (Table 23 and Figure 87).

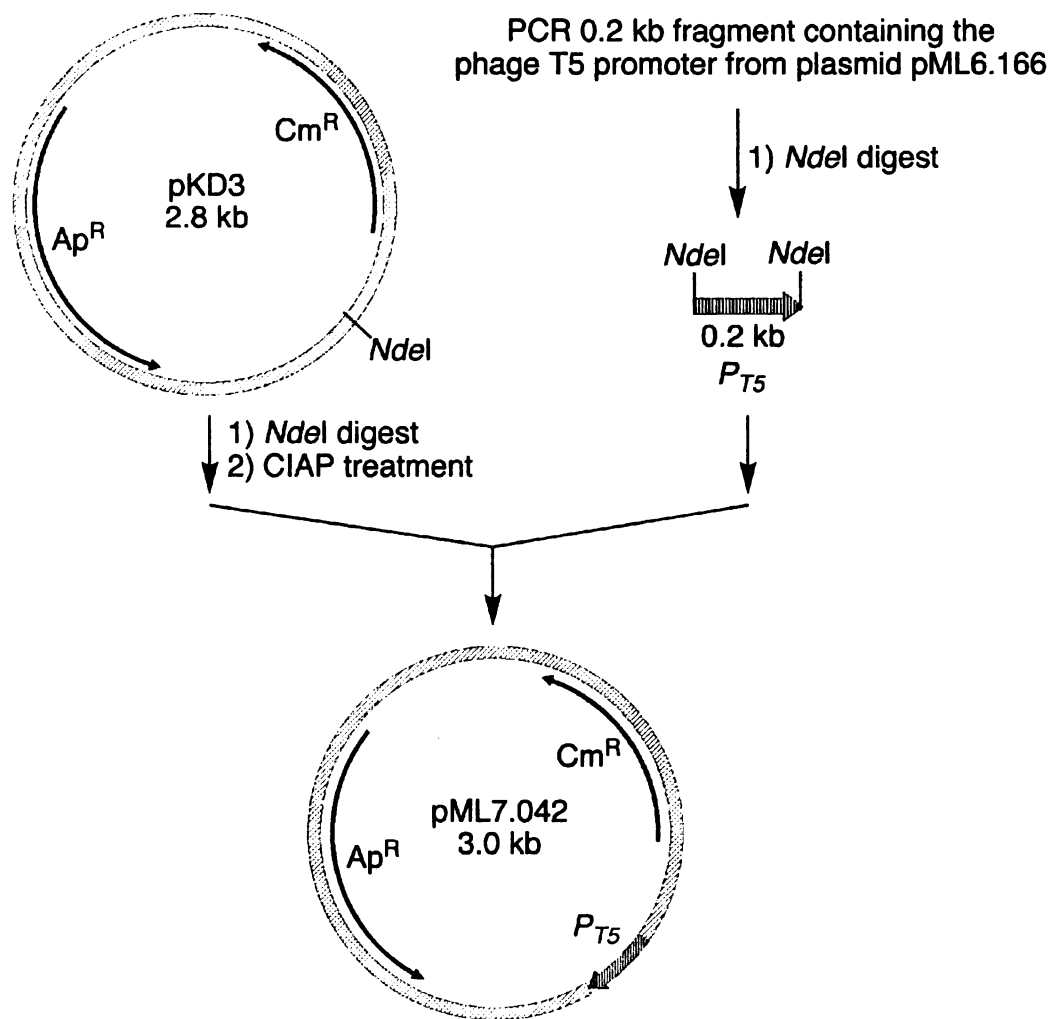


Figure 85. Construction of plasmid pML7.042.

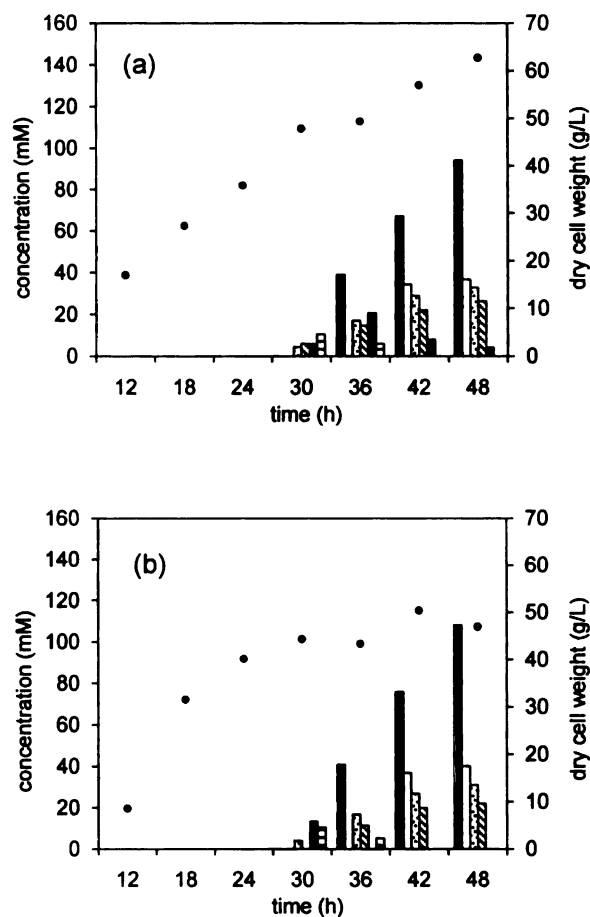


Figure 86. Cultivation of *E. coli* constructs having a chromosomal P_{tac} -*adhP* gene insertion under fermentor-controlled conditions. (a) *E. coli* KIT3/pWN7.126B; (b) *E. coli* KIT4/pWN7.126B; ■ D-1,2,4-butanetriol; □ 3,4-dihydroxy-D-butanoic acid; ▤ 3-deoxy-D-glycero-pentanoic acid; ▨ 4,5-dihydroxy-threo-L-norvaline; ▩ 3-deoxy-D-glycero-pentulosonic acid; ▪ D-xylonic acid.

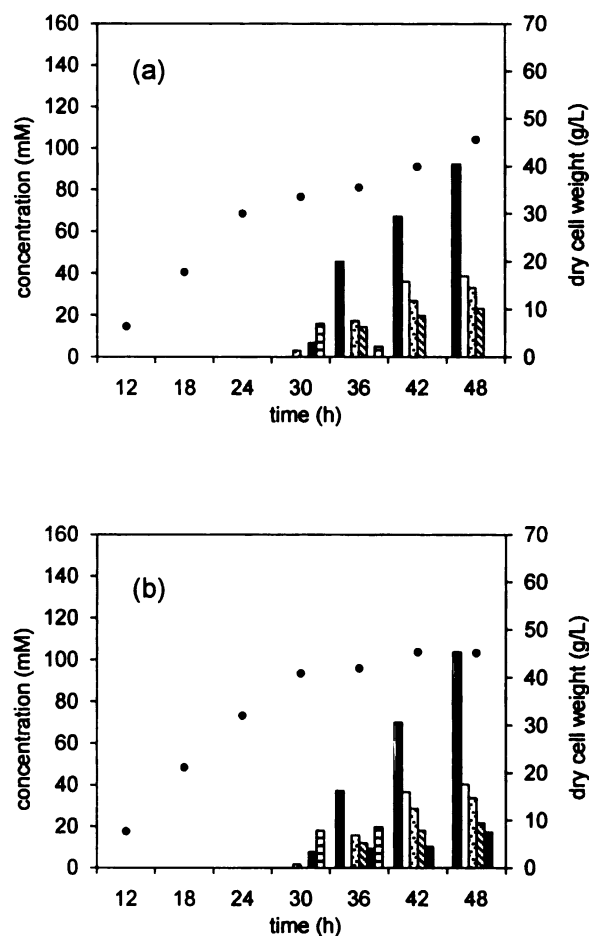


Figure 87. Cultivation of *E. coli* constructs having a chromosomal T5 promoter insertion under fermentor-controlled conditions. (a) *E. coli* KIT6/pWN7.126B; (b) *E. coli* KIT7/pWN7.126B; ■ D-1,2,4-butanetriol; □ 3,4-dihydroxy-D-butanoic acid; ▤ 3-deoxy-D-glycero-pentanoic acid; ▨ 4,5-dihydroxy-threo-L-norvaline; □ 3-deoxy-D-glycero-pentulosonic acid; ▩ D-xylonic acid.

It is known in the literature that the insertion of an antibiotic gene on the chromosome might exert a polarity effect on downstream genes, ultimately leading to growth issues.⁶⁷⁻⁷⁰ Excising the chloramphenicol resistance gene in *E. coli* D-1,2,4-butanetriol synthesizing constructs was of importance as observed in the cultivation experiments shown in Figure 86 and 65. *E. coli* microbes KIT3/pWN7.126B and KIT6/pWN7.126B synthesized D-1,2,4-butanetriol at a concentration of 10.0 g/L (47% yield, mol/mol) and 9.8 g/L (46% yield, mol/mol), respectively. A 15% increase of product concentration was observed in the cultivations using the antibiotic marker-free derivatives of these constructs. *E. coli* KIT4/pWN7.126B and *E. coli* KIT7/pWN7.126B synthesized D-1,2,4-butanetriol at a concentration of 11.5 g/L (55% yield, mol/mol) and 11.0 g/L (52% yield, mol/mol), respectively. *E. coli* KIT4/pWN7.126B was able to synthesize D-1,2,4-butanetriol at a concentration and yield 10% higher than the benchmark *E. coli* WN13/pWN7.126B. The success in using chromosomal *adhP* expression for the microbial D-1,2,4-butanetriol synthesis led to the construction of a second generation *E. coli* microbial catalyst.

Construction of the second generation D-1,2,4-butanetriol synthesizing E. coli microbe

The new generation of D-1,2,4-butanetriol synthesizing *E. coli* strain contains all the genetic modification developed in *E. coli* KIT4/pWN7.126B. In addition, chromosomal inactivations of *yiaE* and *yedW* genes were introduced. In an earlier study, *E. coli yiaE* and *yedW* genes encoding enzymes were identified to catalyze the reduction of 3-deoxy-D-glycero-pentulosonic acid to form 3-deoxy-D-glycero-pentanoic acid using NADPH as a cofactor.³⁶ Introducing these two inactivations into the second generation

microbe would likely reduce the production of 3-deoxy-D-*glycero*-pentanoic acid, ultimately increasing the yield for the desired product D-1,2,4-butanetriol.

Gene *yiaE* was inactivated by replacing it with a chloramphenicol resistant gene flanked with Flp recognition target sequence (FRT) amplified by PCR from plasmid pKD3. *E. coli* KIT16 containing this *yiaE* gene inactivation was prepared directly from *E. coli* KIT2 using the same homologous recombination method as described earlier. Gene *ycdW* was replaced by a kanamycin resistant gene that was obtained using a different template plasmid pKD4. The introduction of this kanamycin gene locus into *E. coli* KIT16 yielded a double knockout strain KIT17. Thermal induction of the Flp recombinase of KIT17/pCP20 excised the chloramphenicol and kanamycin marker in a single step and afforded strain KIT18. Both *E. coli* strains KIT17 and KIT18 were then transformed with *mdlC* containing plasmid pWN7.126B and evaluated (Table 24). Preliminary fermentation results were shown in Table 25. *E. coli* KIT17/pWN7.126B synthesized only 5.5 g/L (26% yield, mol/mol) D-1,2,4-butanetriol, presumably due to the polarity effect. KIT18/pWN7.126B yielded 11.2 g/L (53% yield, mol/mol) D-1,2,4-butanetriol, which is consistent with the earlier result using KIT4/pWN7.126B as catalyst (Table 24, line 3). Most importantly, a decrease in the concentration of 3,4-dihydroxy-D-butanoic acid and 3-deoxy-D-*glycero*-pentanoic acid was observed for this fermentation, indicating that the *adhP* genomic expression and the 2-keto-acid dehydrogenase inactivations were functional. Attempts to vary the amount of substrate D-xylose yielded an increase in D-1,2,4-butanetriol production (Table 25, line 4 – 6). Using 50 g/L D-xylose as substrate, *E. coli* KIT18/pWN7.126B was able to synthesize 18.0 g/L D-1,2,4-butanetriol in 54% yield.

Table 24. Creation of the second generation D-1,2,4-butanetriol synthesizing *E. coli*.

strain	genotype
KIT16	W3110yjhH::FRTyagE::FRTserAxylAB::xdh-adhP- <i>P_{tac}</i> -FRTyiaE::FRT-Cm ^R -FRT
KIT17	W3110yjhH::FRTyagE::FRTserAxylAB::xdh-adhP- <i>P_{tac}</i> -FRTyiaE::FRT-Cm ^R -FRT ycdW::FRT-Km ^R -FRT
KIT18	W3110yjhH::FRTyagE::FRTserAxylAB::xdh-adhP- <i>P_{tac}</i> -FRTyiaE::FRTycdW::FRT

Table 25. Preliminary results on the cultivation of *E. coli* KIT18/pWN7.126B under fermentor-controlled conditions.

construct	xylose (g)	time (h)	titer (g/L)					BT yield (%)
			BT ^a	DHBA ^b	DPA ^c	DHN ^d	DGP ^e	
WN13/pWN7.126B	30	48	10.2	4.6	5.1	3.8	0	48
KIT17/pWN7.126B	30	48	5.5	3.8	2.1	4.4	12.6	26
KIT18/pWN7.126B	30	48	11.2	3.9	2.9	5.3	0	53
KIT18/pWN7.126B	50	48	16.5	4.9	5.4	6	3	47
KIT18/pWN7.126B	50	54	18.0	5.2	5.5	5.9	0	54
KIT18/pWN7.126B	70	60	18.3	4.7	4.7	8.5	15.3	37

^aD-1,2,4-butanetriol (BT)^b3,4-dihydroxy-D-butanoic acid (DHBA)^c3-deoxy-D-glycero-pentanoic acid (DPA)^d4,5-dihydroxy-*threo*-L-norvaline (DHN)^e3-deoxy-D-glycero-pentulosonic acid (DGP)

Future work

Directed Evolution of 3-deoxy-D-glycero-pentulosonate decarboxylase

A variety of strategies was pursued to identify a 3-deoxy-D-glycero-pentulosonate decarboxylase with improved activity. Using the bioinformatics approach, four novel 3-deoxy-glycero-pentulosonate decarboxylases were successfully identified. Unfortunately, these decarboxylases were less active than the *P. putida* MdlC currently employed in the

D-1,2,4-butanetriol synthesizing construct. Attempts to evolve an active decarboxylase using family gene shuffling were unsuccessful due to insufficient screening capability and sequence identity. Recently, a robotic screening workstation was purchased and is available for access. This robotic screening workstation features a Genetic QPix2XT colony picker, a Tecan Freedom Evo 200 workstation and a Molecular Devices SpectraMax Plus 384 microplate reader. Mutant libraries will be generated by error-prone PCR. Mutants with improved activity will then be subject to DNA shuffling for further improvement. Library screening of *mdlC* mutants will be carried out in 96-well format. Purified DNA obtained from error-prone PCR⁷¹ will be transformed into electrocompetant *E. coli* DH5 α and plated on a QTray (Genetix) containing LB/Ap solid medium. This step will be optimized such that independent colonies appearing on each tray will be recognized by the CCD camera of the Genetix QPix2^{XT} colony picker. Single colonies on the tray will then be screened using the following protocol recently revised.

A sterilized deep-well block (Genetix) with each well containing 1 mL of LB/Ap liquid medium will be inoculated with single colonies from the mutant library using the 96-pin head of the colony picker. Simultaneously, a master plate for these mutants will be prepared by inoculating the material left on the 96-pin head into the wells of a microplate containing 100 μ L LB-freeze medium in each well. Wells A1 and B1 of the deep-well block will be inoculated with DH5 α /pKK223-3 and DH5 α /pML7.118 (*P. putida mdlC* gene cloned into pKK223-3) as controls, while well H12 will be left as blank. Therefore, a total of 93 independent mutants will be assayed in a typical screening cycle. The deep-well block will be covered by the Qiagen Airpore Sheet and incubated in the shaker at 37

°C for 12 h with agitation at 250 rpm, while the master plate will be incubated overnight in an oven at 37 °C.

After the incubation, glycerol will be added to the wells and the master plate will be stored at -80 °C. The deep-well block will be removed from the shaker and the cells will be pelleted by centrifugation at 4,000 rpm for 5 min. With the exception of the centrifugation steps, all subsequent microplates handling and liquid pipetting steps will be automated by the Tecan Freedom Evo 200 Workstation coupled with the Evoware Plus v1.4 software. The supernatant will be discarded and the cell pellets at the bottom of the block will be resuspended in 100 µL BugBuster solution (Novagen), transferred to a 96-well microplate (Nunc), and incubated in a microplate hotel at room temperature for 0.5 h without agitation. The cell lysate will then be centrifuged at 4,000 rpm for 20 min to remove cell debris. Protein concentration will be estimated by the measurement of optical density at 280 nm using the Molecular Devices SpectraMax Plus 384 microplate reader that is integrated to the Tecan workstation. This microplate will then be stored in the cooling hotel at 4 °C prior to use.

The enzymatic assay reaction for decarboxylases will be setup as described earlier followed by a 6 h incubation period at room temperature. The reactions will be quenched by the addition of 30 µL of 10% trichloroacetic acid solution followed by protein removal using centrifugation. 50 µL of Schiff's reagent⁷² will be added to each well and incubated at room temperature for another 30 min. Schiff's reagent has been demonstrated to react selectively with 3,4-dihydroxybutanal to give a red dye. The formation of this red dye in microplate wells will be recorded by measuring the absorbance at 550 nm. The higher

absorbance readings at 550 nm will be interpreted as indicative of a mutation leading to a more reactive decarboxylase.

Enzymes corresponding to these putatively more reactive decarboxylases will be assayed a second time at higher dilution using Schiff's reagent. The more active decarboxylase mutants will be assayed for a third time using an *in vivo* assay described earlier in this chapter. *E. coli* W3110 transformed with plasmid-localized mutant genes will be cultivated using D-xylonic acid as sole carbon source for growth. The resulting culture medium will be derivatized with BSTFA and the D-1,2,4-butanetriol will be quantified with our highly sensitive gas chromatography interfaced with a flame-ionization detector.

Chemo-microbial synthesis of (S)-3-hydroxy- γ -butyrolactone

As discussed in chapter one, (S)-3-hydroxy- γ -butyrolactone is an important synthetic intermediate for a variety of chiral compounds. For instance, the use of this precursor in the manufactures of Crestor and Zetia attracts enormous commercial interest. There are various strategies in the literature toward the synthesis of this molecule. (S)-3-hydroxy- γ -butyrolactone can be synthesized from the selective reduction of L-malic acid ester using a dimethyl sulfide complex of borane and a catalytic amount of sodium borohydride.⁷³ Methods of enzymatic or catalytic reduction of β -ketoester are known, however, with low enantioselectivity.⁷⁴ Many other methods were derived using carbohydrates such as cellobiose,⁷⁵ maltose,⁷⁶ lactose,⁷⁷ starch⁷⁸ or cellulose⁷⁸ as starting materials. However, issues associated with these methods such as low yielding and low stereoselective reactions, and byproducts formation of glycolic acid, isosaccharinic acid, formic acid, ketones and diketones increase the product cost and complicate the

downstream purification process.⁷⁵⁻⁷⁸ The recently discovered microbial synthesis of 3-deoxy-D-*glycero*-pentulosonic acid using *E. coli* microbe KIT15/pML7.135 from D-xylose provides a commercially viable alternative toward the synthesis of (*S*)-3-hydroxy- γ -butyrolactone. A simple oxidative decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid into 3,4-dihydroxy-D-butanoic acid followed by an acid-catalyzed lactonization to afford this important chiral synthon was envisioned. This stereospecific synthesis guarantees the enantio-purity of product (*S*)-3-hydroxy- γ -butyrolactone (Figure 88).

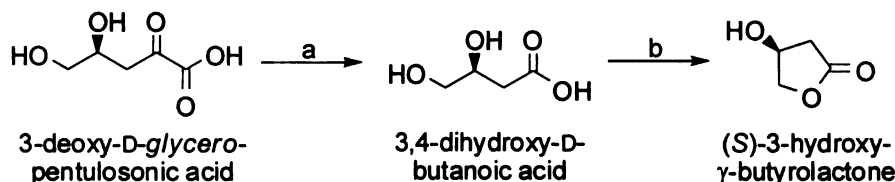


Figure 88. Proposed chemo-microbial synthesis of (*S*)-3-hydroxy- γ -butyrolactone.
(a) oxidative decarboxylation; (b) lactonization.

To facilitate this study, a simple purification method of 3-deoxy-D-*glycero*-pentulosonic acid was developed from the fermentation broth of KIT15/pML7.135. The fermentation broth was first passed through an anion-exchange column packed with Dowex 1 \times 8 (Cl⁻ form) and eluted with 100 mM HCl. The HCl eluent was then evaporated to dryness and the residue washed with methanol. Evaporation of the methanol layer yielded a colorless oily product 3-deoxy-D-*glycero*-pentulosonic acid with > 99% purity as indicated by ¹H and ¹³C NMR.

Oxidative decarboxylation of purified 3-deoxy-D-*glycero*-pentulosonic acid in water under reflux yielded trace amount of 3,4-dihydroxy-D-butanoic acid (Table 26, entry 1). Attempts to decarboxylate 3-deoxy-D-*glycero*-pentulosonic acid under nitrogen

atmosphere was unsuccessful (Table 26, entry 2). Addition of 3% H₂O₂ to the reaction mixture yielded 3,4-dihydroxy-D-butanoic acid in 90% under room temperature conditions (Table 26, entry 3). Quantitative conversion was obtained with the addition of 5 mol% FeSO₄ as a catalyst. This synthesis was further improved using 1.5% H₂O₂ and 1% FeSO₄ while the same conversion was maintained (Table 26, entry 4 – 6). To take a step further, quantitative oxidative decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid in fermentation broth KIT15/pML7.135 without prior purification was achieved even without extra FeSO₄ addition as catalyst (Table 26, entry 7 – 8). The fact that this reaction took only 15 min to complete certainly made it further appealing and commercially viable. Research is now focused on developing a reactive extraction method for (*S*)-3-hydroxy- γ -butyrolactone. In an ideal case, both acid-catalyzed lactonization and purification will be achieved in a single extraction step in Karr reciprocating column using ethyl acetate.

Table 26. Oxidative decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid.

entry	starting ketoacid	temp (°C)	time (min)	H ₂ O ₂ (%)	FeSO ₄ (mol%)	conversion (%)	note
1	1g purified	100	30	0	0	trace	
2	1g purified	100	30	0	0	trace	purged with N ₂
3	1g purified	rt	15	3	0	90	
4	1g purified	rt	15	3	5	quant.	
5	1g purified	rt	15	1.5	5	quant.	
6	1g purified	rt	15	1.5	1	quant.	
7	4g in broth ^a	rt	15	1.5	1	quant.	
8	4g in broth ^a	rt	15	1.5	0	quant.	

^aFermentation broth contained 3-deoxy-D-*glycero*-pentulosonic acid at a concentration of 21 g/L. Cells and protein were separated by ultra-filtration prior to use.

Enzyme activity studies for E. coli adhP expressing microbes

The studies on alcohol dehydrogenases is important to fulfill our understanding to otherwise unknown native *E. coli* enzyme conversions of 3,4-dihydroxy-D-butanal into D-1,2,4-butanetriol. Among the five *E. coli* alcohol dehydrogenases identified using bioinformatics, only three showed *in vitro* enzyme activity using racemic 3,4-dihydroxy-D-butanal as substrate. The benzoylformate decarboxylase activity time-course study indicated that both *E. coli* WN13/pWN7.126B and its derivative KIT10/pWN7.126B with chromosomal *adhP* inactivation appeared to be similar during fermentation, and achieved the highest value of 65 and 57 U/mg, respectively, at 48 h (Figure 89a). On the other hand, no measurable D-1,2,4-butanetriol dehydrogenase activity was observed for KIT10/pWN7.126B during the fermentation (Figure 89b). Since these two constructs expressed about the same level of benzoylformate decarboxylase activity, activity of AdhP became the only possible factor that could contribute to the difference in D-1,2,4-butanetriol production.

Construction of *E. coli* microbes by co-expressing 1,2,4-butanetriol dehydrogenase candidates *E. coli adhP*, *Z. mobilis adhA* and *K. pneumoniae dhaT* genes together with *P. putida mdlC* gene in a single multi-copy plasmid had led to poor cell growth and reduced decarboxylase activity (Figure 89a). Research was then focused on *E. coli adhP* to avoid potential complications due to heterologous gene expression of *dhaA* and *dhaT* in *E. coli*. Our efforts toward studying *adhP* expression for D-1,2,4-butanetriol synthesis are summarized in Table 27. In the case of *E. coli* WN13/pML6.195 (*E. coli adhP* on a plasmid), only 32 U/mg benzoylformate decarboxylase activity was expressed at the end of the fermentation, which was half of the activity observed for other constructs with a

single *mdlC* gene localized plasmid (Figure 89a). On the other hand, modulated expression of *adhP* on the *E. coli* chromosome afforded *E. coli* KIT4/pWN7.126B and *E. coli* KIT7/pWN7.126B. These microbes successfully produced 11.5 g/L and 11 g/L D-1,2,4-butanetriol, which was an average of 12% improvement in concentration over the benchmark *E. coli* WN13/pWN7.126B. The benzoylformate decarboxylase activity of these two constructs was restored as indicated by the enzyme activity time-course study (Figure 89a). In addition, while maintaining the same level of decarboxylase activity, *adhP* gene encoding dehydrogenase expression was also achieved in these constructs. In the case of *E. coli* KIT4/pWN7.126B, AdhP activity attained a maximum of 0.05 U/mg at 48 h, which corresponded to a five-fold increase comparing to the wild-type *E. coli* (Figure 89b). These enzyme activity studies provide a second line of evidence that *E. coli adhP* is indeed the dehydrogenase that participates in the D-1,2,4-butanetriol biosynthesis. Due to the fact that inactivating the *adhP* gene did not shut down D-1,2,4-butanetriol biosynthetic pathway, it becomes tempting to speculate that one or more other native *E. coli* dehydrogenases might be acting as *in vivo* D-1,2,4-butanetriol dehydrogenase. Last but not least, expressing *adhP* chromosomally under a *tac* or phage T5 promoter only afforded slight improvement in D-1,2,4-butanetriol production, which provides an additional support that MdlC catalyzed 3-deoxy-D-*glycero*-pentulosonate decarboxylation is indeed the limiting step in the pathway. Although preliminary cultivation of the second generation microbe *E. coli* KIT18/pWN7.126B successfully achieved D-1,2,4-butanetriol synthesis at a concentration of 18 g/L, it is believed that an active 3-deoxy-D-*glycero*-pentulosonate decarboxylase will determine the ultimate success of this pentose-based pathway.

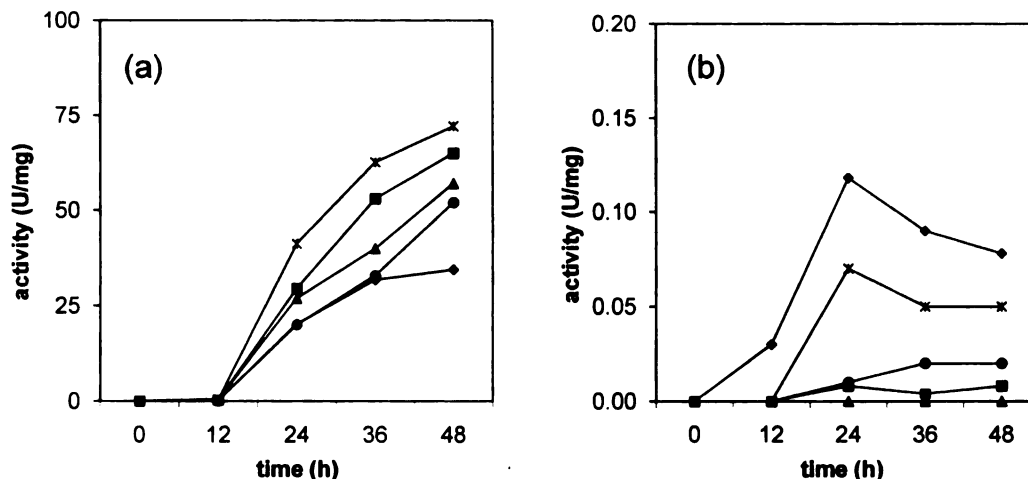


Figure 89. Enzyme activity time-course for microbial syntheses of D-1,2,4-butanetriol under fermentor-controlled conditions. (a) benzoylformate decarboxylase activity. (b) 1,2,4-butanetriol dehydrogenase activity. *E. coli* WN13/pWN7.126B (square); *E. coli* WN13/pML6.195 (diamond); *E. coli* KIT10/pWN7.126B (triangle); *E. coli* KIT4/pWN7.126B (asterisk); *E. coli* KIT7/pWN7.126B (circle). One unit of activity is defined as 1 μ mol of product formation per minute.

Table 27. *E. coli* strains with *adhP* gene expression.

construct	genomic <i>adhP</i>				titer, g/L (yield, %)		activity, U/mg	
	<i>P_{tac}</i>	<i>P_{T5}</i>	knockout	Cm ^R	BT	BA	MdIC	AdhP
WN13/pWN7.126B				✓	10.5(50)	4.5(19)	67	0.006
WN13/pML6.195				✓	4.6(22)	4.2(17)	38	0.08
KIT2/pWN7.126B					9.1(43)	4.5(19)	77	0.004
KIT3/pWN7.126B	✓			✓	10.0(47)	4.4(18)	35	0.02
KIT4/pWN7.126B	✓				11.5(55)	4.5(19)	64	0.02
KIT6/pWN7.126B		✓		✓	9.8(46)	4.5(19)		
KIT7/pWN7.126B		✓			11.0(52)	4.7(20)	51	0.02
KIT9/pWN7.126B			✓	✓	8.2(39)	5.0(21)		
KIT10/pWN7.126B			✓		6.5(31)	5.3(22)	57	0

Cm^R, chloramphenicol gene on the host chromosome; BT, D-1,2,4-butanetriol; BA, 3,4-dihydroxy-D-butanoic acid; MdIC, *P. putida* benzoylformate decarboxylase; AdhP, *E. coli* alcohol dehydrogenase.

REFERENCE

- ¹ Daniel, R.; Bobik, T. A.; Gottschalk, G. *FEMS Microbiol. Rev.* **1999**, *22*, 553.
- ² Lawson, M. E. In *Kirk-Othmer Encyclopedia of Chemical Technology*; Sugar Alcohols, 1991, Wiley.
- ³ http://www.cargill.com/today/releases/02_06_16iftconf.htm.
- ⁴ (a) Lin, S.-J.; Wen, C. J.; Liao, J.-C.; Chu, W. S. *Process Biochem.* **2001**, *36*, 1249. (b) Kasumi, T. *Jap. Agricul. Res. Quart.* **1995**, *29*, 49.
- ⁵ Veiga-da-Cunha, M.; Santos, H.; Van Schaftingen, E. *J. Bacteriol.* **1993**, *175*, 3941. (b) Veiga-da-Cunha, M.; Firme, P.; San Romao, M. V.; Santos, H. *Appl. Environ. Microbiol.* **1992**, *58*, 2271.
- ⁶ Seyfried, M.; Daniel, R.; Gottschalk, G. *J. Bacteriol.* **1996**, *178*, 5793.
- ⁷ Toraya, T.; Shirakashi, T.; Kosuga, T.; Fukui, S. *Biochem. Biophys. Res. Commun.* **1976**, *69*, 475.
- ⁸ Abeles, R. H.; Dolphin, D. *Acc. Chem. Res.* **1976**, *9*, 114.
- ⁹ Toraya, T. In *Vitamin B12 and B12-Proteins*; Krautler, B.; Arigoni, D.; Golding, B. T. Eds.; Wiley-VCH: Weinheim, 1998; p303.
- ¹⁰ Boenigk, R.; Bowien, S.; Gottschalk, G. *Appl. Microbiol. Biotechnol.* **1993**, *38*, 453.
- ¹¹ Daniel, R.; Stuerz, K.; Gottschalk, G. *J. Bacteriol.* **1995**, *177*, 4392.
- ¹² Roth, J. R.; Lawrence, J. G.; Bobik, T. A. *Annu. Rev. Microbiol.* **1996**, *50*, 137.
- ¹³ Toraya, T.; Honda, S.; Fukui, S. *J. Bacteriol.* **1979**, *139*, 39.
- ¹⁴ Badia, J.; Ros, J.; Aguilar, J. *J. Bacteriol.* **1985**, *161*, 435.
- ¹⁵ Scarlett, F. A.; Turner, J. M. *J. Gen. Microbiol.* **1976**, *95*, 173.
- ¹⁶ Tobimatsu, T.; Hara, T.; Sakaguchi, M.; Kishimoto, Y.; Wada, Y.; Isoda, M.; Sakai, T.; Toraya, T. *J. Biol. Chem.* **1995**, *270*, 7142.

- ¹⁷ Poznanskaja, A. A.; Tanizawa, K.; Soda, K.; Toraya, T.; Fukui, S. *Arch. Biochem. Biophys.* **1979**, *194*, 379.
- ¹⁸ Tobimatsu, T.; Azuma, M.; Matsubara, H.; Takatori, H.; Niida, T. Nishimoto, K.; Satoh, H.; Hayashi, R.; Toraya, T. *J. Biol. Chem.* **1996**, *271*, 22352.
- ¹⁹ (a) Bachovchin, W. W.; Moore, K. W.; Richards, J. H. *Biochemistry*, **1977**, *16*, 1082. (b) Schneider, Z.; Pawelkiewicz, J. *Acta Biochim. Pol.* **1966**, *13*, 311. (c) Poznanskaya, A. A.; Yakusheva, M. I.; Yakovlev, V. A. *Biochim. Biophys. Acta.* **1977**, *484*, 236.
- ²⁰ Toraya, T.; Fukui, S. In *B₁₂*; Dolphin, D. Ed.; John Wiley & Sons, Inc.: New York, 1982; p233.
- ²¹ (a) Honda, S.; Toraya, T.; Fukui, S. *J. Bacteriol.* **1980**, *143*, 1458. (b) Ushio, K.; Honda, S.; Toraya, T.; Fukui, S. *J. Nutr. Sci. Vitaminol.* **1982**, *28*, 225.
- ²² Homann, T.; Tag, C.; Biebl, H.; Deckwer, W.-D.; Schink, B. *Appl. Microbiol. Biotechnol.* **1990**, *33*, 121.
- ²³ Daniel, R.; Gottschalk, G. *FEMS Microbiol. Lett.* **1992**, *100*, 281.
- ²⁴ Tong, I.-T.; Liao, H. H.; Cameron, D. C. *Appl. Environ. Microbiol.* **1991**, *57*, 3541.
- ²⁵ Daniel, R.; Boenigk, R.; Gottschalk, G. *J. Bacteriol.* **1995**, *177*, 2151.
- ²⁶ Sprenger, G. A.; Hammer, B. A.; Johnson, E. A.; Lin, E. C. C. *J. Gen. Microbiol.* **1989**, *135*, 1255.
- ²⁷ Tobimatsu, T.; Azuma, M.; Hayashi, S.; Nishimoto, K.; Toraya, T. *Biosci. Biotechnol. Biochem.* **1998**, *62*, 1774.
- ²⁸ (a) Toraya, T.; Uesaka, M.; Kondo, M.; Fukui, S. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 350. (b) Toraya, T.; Uesaka, M.; Fukui, S. *Biochemistry* **1974**, *13*, 3895.
- ²⁹ Bobik, T. A.; Xu, Y.; Jeter, R. M.; Otto, K. E.; Roth, J. R. *J. Bacteriol.* **1997**, *179*, 6633.
- ³⁰ Dabrock, B.; Bahl, H.; Gottschalk, G. *Appl. Environ. Microbiol.* **1992**, *58*, 1233.
- ³¹ Heindrickx, M.; de Vos, P.; Vancanneyt, M.; de Ley, J. *Appl. Microbiol. Biotechnol.* **1991**, *34*, 637.

- ³² (a) Speranza, G.; Manitto, P.; Fontana, G.; Monti, D.; Galli, A. *Tetrahedron Lett.* **1996**, 37, 4247. (b) Speranza, G.; Corti, S.; Fontana, G.; Manitto, P. *J. Agric. Food Chem.* **1997**, 45, 3476. (c) Manitto, P.; Speranza, G.; Fontana, G.; Galli, A. *Helv. Chim. Acta* **1998**, 81, 2005. (d) Speranza, G.; Morelli, C. F.; Orlandi, M.; Scarpellina, M.; Manitto, P. *Helv. Chim. Acta* **2001**, 84, 335.
- ³³ Seifert, C.; Bowien, S.; Gottschalk, G.; Daniel, R. *Eur. J. Biochem.* **2001**, 268, 2369.
- ³⁴ Luers, F.; Seyfried, M.; Daniel, R.; Gottschalk, G. *FEMS Microbiol. Lett.* **1997**, 154, 337.
- ³⁵ Niu, W. *PhD Thesis*, Michigan State University, 2004.
- ³⁶ Niu W.; Frost, J. W. unpublished results.
- ³⁷ Toraya, T.; Kazutoshi, U.; Fukui, S.; Hogenkamp, H. P. C. *J. Biol. Chem.* **1977**, 252, 963.
- ³⁸ Johnson, E. A.; Lin, E. C. C. *J. Bacteriol.* **1987**, 169, 2050.
- ³⁹ (a) Niu, W.; Molefe, M. N.; Frost, J. W. *J. Am. Chem. Soc.* **2003**, 125, 12998. (b) Frost, J. W. WO 2005/068642, 2005.
- ⁴⁰ (a) Hegeman, G. D. *J. Bacteriol.* **1966**, 91, 1140. (b) Hegeman, G. D. *J. Bacteriol.* **1966**, 91, 1155. (c) Hegeman, G. D. *J. Bacteriol.* **1966**, 91, 1161. (d) Fewson, C. A. *FEMS Microbiol. Rev.* **1988**, 54, 85.
- ⁴¹ (a) Hasson, M. S.; Muscate, A.; Henehan, G. T.; Guidinger, P. F.; Petsko, G. A.; Ringe, D.; Kenyon, G. L. *Biochemistry* **1998** 37, 9918. (b) Klunger, R. *Pure Appl. Chem.* **1997**, 69, 1957. (c) Weiss, P. M.; Garcia, G. A.; Kenyon, G. L.; Cleland, W. W.; Cook, P. F. *Biochemistry* **1998**, 27, 2197. (d) Reynolds, L. J.; Garcia, G. A.; Kozarich, J. W.; Kenyon, G. L. *Biochemistry* **1998**, 27, 5530.
- ⁴² Polovnikova, E. S.; Mcleish, M. J.; Sergienko, E. A.; Burgner, J. T.; Anderson, N. L.; Bera, A. K.; Jordan, F.; Kenyon, G. L.; Hasson, M. S. *Biochemistry* **2003**, 42, 1820.
- ⁴³ (a) Bommarius, A. S.; Riebel, B. R. *Biocatalysis*; Wiley-VCH: Weinheim, 2004. (b) Krawetz, S. A.; Womble, D. D. *Introduction to Bioinformatics: A Theoretical and Practical Approach*; Humana Press: Totowa, NJ, 2003.
- ⁴⁴ <http://www.ncbi.nlm.nih.gov/>

- ⁴⁵ <http://www.ebi.ac.uk/>
- ⁴⁶ <http://wit.integratedgenomics.com/GOLD/>
- ⁴⁷ <http://www.ncbi.nlm.nih.gov/BLAST/>
- ⁴⁸ <http://www.ebi.ac.uk/clustalw/>
- ⁴⁹ <http://pymol.sourceforge.net/>
- ⁵⁰ <http://www.rcsb.org/pdb/home/home.do>
- ⁵¹ (a) Siegert, P.; Pohl, M.; Kneen, M. M.; Pogozeva, I. D.; Kenyon, G. L.; McLeish, M. J. In *Thiamine: Catalytic mechanisms in normal and disease states*; Jordan, F.; Patel, M. S. Eds.; Marcel Dekker, Inc.: NY, 2004; p275. (b) Pohl, M.; Siegert, P.; Mesch, K.; Bruhn, H.; Grotzinger, J. *Eur. J. Biochem.* **1998**, *257*, 538.
- ⁵² <http://www.stratagene.com/manuals/200518.pdf>
- ⁵³ Conway, T.; Osman, Y. A.; Konnan, J. I.; Hoffmann, E. M.; Ingram, L. O. *J. Bacteriol.* **1987**, *169*, 949.
- ⁵⁴ Arnold, F. H.; Georgiou, G. *Directed Evolution Library Creation: Methods and Protocols*; Humana Press: Totowa, NJ, 2003. (b) Arnold, F. H.; Georgiou, G. *Directed Enzyme Evolution: Screening and Selection Methods*; Humans Press: Totowa, NJ, 2003.
- ⁵⁵ (a) Chen, K.; Arnold, F. H. *Proc. Natl. Acad. Sci. USA* **1993**, *93*, 5618. (b) Powell, K. A.; Ramer, S. W.; Cardayre, S. B.; Stemmer, W. P. C.; Tobin, M. B.; Longchamp, P. F.; Huisman, G. W. *Angew. Chem. Int. Ed.* **2002**, *40*, 3948. (c) Schmidt-Dannert, C. *Biochemistry* **2001**, *40*, 13125. (d) Tao, H.; Cornish, V. W. *Curr. Opin. Chem. Biol.* **2002**, *6*, 858.
- ⁵⁶ Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A. *Trends Biochem. Sci.* **2001**, *26*, 100.
- ⁵⁷ Joern, J. M.; Meinhold, P.; Arnold, F. H. *J. Mol. Biol.* **2002**, *316*, 643.
- ⁵⁸ Burchert, J.; Viikari, L.; Linko, M.; Markkanen, P. *Biotechnol. Lett.* **1986**, *8*, 541.
- ⁵⁹ Frost, J. W.; Niu, W.; Lau, M. K.; Dodds, D.; Saffron, C. unpublished results.

- ⁶⁰ (a) Studier, F. W.; Moffatt, B. A. *J. Mol. Biol.* **1986**, *189*, 113. (b) Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorff, J. W. *Methods Enzymol.* **1990**, *185*, 60.
- ⁶¹ Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1074.
- ⁶² <http://www.emdbiosciences.com/docs/docs/PROT/TB031.pdf>
- ⁶³ (a) Thomas, J. G.; Ayling, A.; Baneyx, F. *Appl. Biochem. Biotechnol.* **1997**, *66*, 197. (b) Wall, J. G.; Pluchthun, A. *Curr. Opin. Biotechnol.* **1995**, *6*, 507.
- ⁶⁴ (a) Nichihara, K.; Kanemori, M.; Kitagawa, M.; Yanagi, H.; Yura, T. *Appl. Environ. Microbiol.* **1998**, *64*, 1694. (b) Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T. *Appl. Environ. Microbiol.* **2000**, *66*, 884.
- ⁶⁵ (a) Horovitz, A. *Curr. Opin. Struct. Biol.* **1998**, *8*, 93. (b) Sigler, P. B.; Xu, Z. H.; Rye, H. S.; Burston, S. G.; Fenton, W. A.; Horwich, A. L. *Annu. Rev. Biochem.* **1998**, *67*, 581. (c) Thirumalai, D.; Lorimer, G. H. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 245.
- ⁶⁶ Henning, H.; Leggewie, C.; Pohl, M.; Muller, M.; Eggert, T.; Jaeger, K.-E. *Appl. Environ. Microbiol.* **2006**, *72*, 7510.
- ⁶⁷ Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640.
- ⁶⁸ Cherepanov, P. P.; Wackernagel, W. *Gene* **1995**, *158*, 9.
- ⁶⁹ (a) McLaggan, D.; Naprstek, J.; Buurman, E.; Epstein, W. *J. Biol. Chem.* **1994**, *269*, 1991. (b) Yan, D.; Ikeda, T. P.; Shauger, A. E.; Kustu, S. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6527.
- ⁷⁰ (a) Ohta, K.; Beall, D. S.; Mejia, J. P.; Shanmugam, K. T.; Ingram, L. O. *Appl. Environ. Microbiol.* **1991**, *57*, 893. (b) Mantinez, A.; York, S. W.; Yomano, L. P.; Pineda, V. L.; Davis, F. C.; Shelton, J. C.; Ingram, L. O. *Biotechnol. Prog.* **1999**, *15*, 891.
- ⁷¹ Cadwell, R. C.; Joyce, G. F. *PCR Meth. Appl.* **1992**, *2*, 28.
- ⁷² Kiernan, J. A. *Histological and histochemical methods: Theory and practice* 3rd ed.; Butterworth Heinemann: Oxford, UK, 1999.
- ⁷³ Saito, S.; Hasegawa, T.; Inaba, M.; Nishida, R.; Fujii, T.; Nomizu, S.; Moriwake, T. *Chem. Lett.* **1984**, 1389.

⁷⁴ (a) Taber, D. F.; Raman, K. *J. Am. Chem. Soc.* **1983**, *105*, 5935. (b) Nakamura, K.; Kawai, Y.; Ohno, A. *Tetrahedron Lett.* **1990**, *31*, 267.

⁷⁵ Rowell, R. M.; Somers, P. J.; Barker, S. A.; Stacey, M. *Carbohydrate Res.* **1969**, *11*, 17.

⁷⁶ Green, J. W. *J. Am. Chem. Soc.* **1956**, *78*, 1894.

⁷⁷ Jacks, T. E. WO 9804543, 1998.

⁷⁸ (a) Whistler, R. L.; Schweiger, R. *J. Am. Chem. Soc.* **1959**, *81*, 3136. (b) Arts, S. J. H. F. F.; Mombarg, E. J. M.; van Bakkum, H.; Sheldon, R. A. *Synthesis* **1997**, 597.

CHAPTER FOUR

Downstream Purification of 1,2,4-Butanetriol

Background

The microbial synthesis of 1,2,4-butanetriol from renewable resources is an attractive process. However, the recovery of this high boiling hydrophilic compound from a fermentation broth is a challenging task. This is also considered critical for the development of a commercially viable process. Conventionally, vacuum distillation is often used in processes for the purification of other polyhydroxy compounds such as glycerol¹ and 1,3-propanediol.² To distill a dilute aqueous solution requires a substantial energy input that adds to the cost of the final product. Earlier attempts in the Frost group to distill the 1,2,4-butanetriol containing reaction mixture from catalytic malic acid hydrogenation³ and chemo-enzymatic synthesis using 2-deoxyribose-5-phosphate aldolase⁴ under vacuum resulted in poor product recovery. Undesired polymer and decomposition products were recovered at the end of the process. Thus, an alternative method for the purification of 1,2,4-butanetriol is needed. Replacement of distillation by energy efficient solvent extraction is believed to be an alternative to reduce the cost of the product separation significantly. It is also suitable for large scale operation. More importantly, solvent extraction does not require excessive heat input to vaporize the compound, therefore issues associated with polymerization and decomposition are thus eliminated. To this end, research involved a systematic and comprehensive examination of the liquid-liquid extraction of 1,2,4-butanetriol from fermentor broth. Potential solvent candidates were screened using a glass continuous liquid-liquid extractor apparatus. A

scalable downstream process was then designed using a bench scale Karr reciprocating column (Koch Modular Process Systems, KMPS).

Setting up an efficient liquid-liquid extraction process is often neglected because it only contributes to a minor part of the overall process scheme. However, significant reduction of process cost can be achieved by fine-tuning an extraction system. Liquid-liquid extraction is a mass transfer operation in which a liquid solution (the feed) contacts with an immiscible or nearly immiscible liquid (the solvent) that exhibits preferential affinity towards the target compound (the solute) in the feed.⁵ Feed and solvent streams resolve after the contact of these two phases. The extract is a solvent rich solution that contains the desired solute, while the raffinate refers to the solute lean feed solution after the extraction.

To design an efficient liquid-liquid extraction process, solvent selection, choice of extraction equipment and operating conditions represent different parameters that require careful consideration and optimization. Among these, a suitable solvent choice is of primary importance. This solvent must possess a relative low boiling point, therefore separates easily from the extracted components. Little or no miscibility with the feed solution is preferred. Corrosiveness, toxicity, flammability and stability are certainly other concerns. The solvent must also be cost effective, and readily available for a commercial operation. Additional consideration such as solvent recovery and solvent release are gaining more attention due to stricter regulations on solvent release and to lower the overall process cost. An ideal single solvent for an extraction system can usually be theoretically judged, but the performance of solvent blends is difficult to predict solely by theoretical means. Therefore, actual experimentation still occupies a

crucial role to make more accurate decisions. Once the solvent candidates are identified, measuring distribution coefficients (m) for the solute of interest and selectivity tests are the next tasks to be accomplished. The distribution coefficient is defined as the ratio of the concentration of solute in extract phase to the concentration in the raffinate.⁶ On the other hand, selectivity concerns the ability of the solvent to pick up that particular solute versus other components in the feed. The solvent of choice has a high distribution coefficient and a good selectivity towards the solute.

The most commonly used single-stage extracting device in the laboratory is a separatory funnel. The industrial equivalence of it is a mixer-settler that was developed in the past among pharmaceutical and agrochemical industries (Figure 90a).⁶ These extractor tanks are used today for washing and neutralization processes that require only a few stages. Extractions are usually operated in a cross-current mode and are particularly flexible and economical for batch operations. When this extraction tank is in operation, solvent is first added to the feed and mechanically mixed. This mixture will then be allowed to settle in a different chamber where it resolves into two layers. Single stage extraction can be easily achieved using this apparatus. If more than one theoretical stage is required, multiple tanks will be setup in series. In this case, a mixer-settler may not be a desirable apparatus. Using the wrong apparatus will lead to an unnecessary waste of solvent and low extraction yields. Another similar industrial scale extractor is a centrifugal extractor (Figure 90b). It features a high-speed centrifuge that offers the advantage of low residence time. Centrifugal extractors are ideal for systems in which the density difference between the feed and solvent is very small. A mechanical device is used to agitate the mixture by increasing the interfacial area and decreasing mass transfer

resistance. In most cases, the number of stages in a centrifugal extractor is limited to one. In contrast to the mixer-settler, centrifugal extractors are usually associated with high setup and maintenance cost.

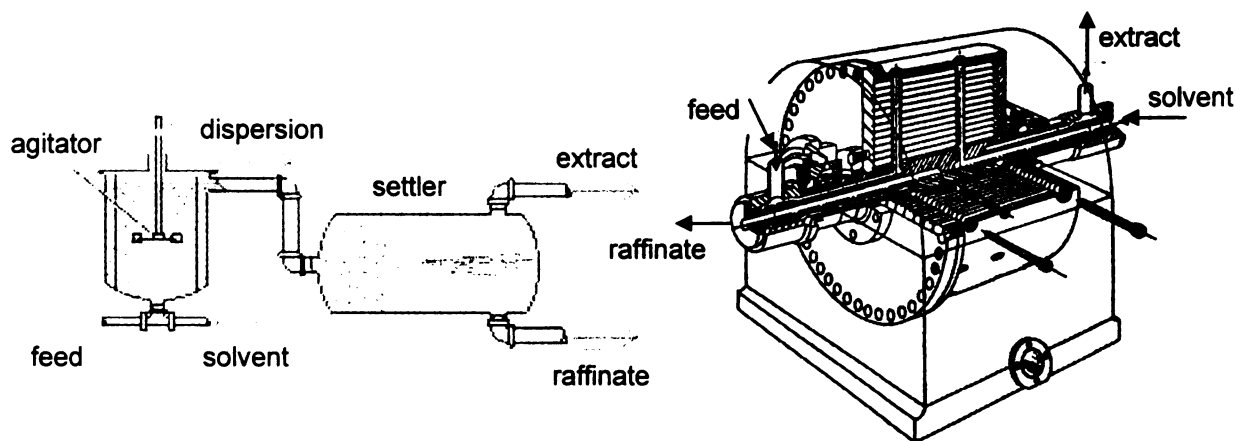


Figure 90. Single-stage liquid-liquid extractors. (a) a mixer-settler.⁷ (b) a centrifugal extractor (Baker-Perkins, Saginaw, MI).

Counter-current column extractors are probably the most popular in the chemical industry. They can be static or agitated. A static column provides a moderate number of stages with relatively low operating and maintenance cost. The packed column, tray column and spray column have different materials arranged in the column and these materials are used to increase the interfacial area in which the two phases contact. However, scaling up a laboratory-scale static column process into an industrial operation is not simple, and less efficient than the mixer-settler process. Agitated columns offer advantages such as reasonable capacity, low operating cost, and high efficiency. An agitated column typically requires less than a third of the number of theoretical stages as compared to a static column. These columns are therefore ideal for any system that a high

number of theoretical stages are required. An associated reduction of solvent usage is also beneficial to the process economics. Agitation in these columns are supplied either by a rotating disc, which usually runs at much higher speeds than a turbine type impeller (Scheibel column, Figure 91a), or with a reciprocating pump, in which a rapid reciprocating motion of baffled and perforated plates are used to create a dispersion (Karr column, Figure 91b). Among these two, Karr columns have an important advantage over Scheibel columns to handle difficult extraction systems that tend to emulsify or flood easily. In the Karr column, a rapid reciprocating motion with small amplitude is created by a pump equipped in the column to superimpose the normal counter-current flow of the liquid phases due to the solvent/feed density difference. The pulsation of the perforated plates causes the light liquid to be dispersed into the heavy phase on the upward stroke and heavy liquid phase to jet into the light phase on the downward stroke. It generates a certain number of theoretical stages required to transfer solute from one phase to the other. By increasing the surface area and turbulence, a more efficient extraction process will be achieved.

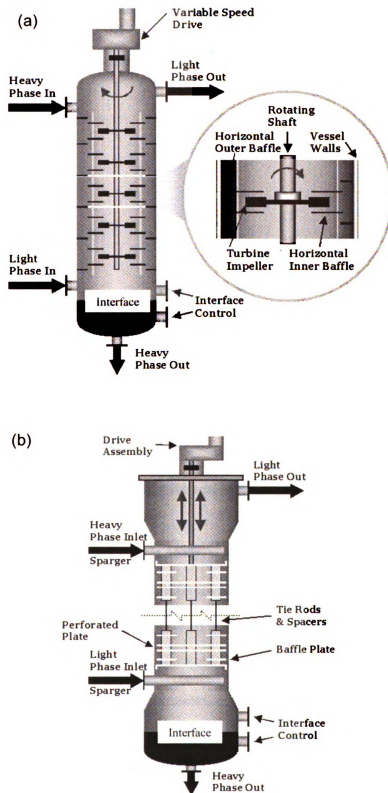


Figure 91. Agitated counter-current extraction columns.⁸ (a) Scheibel column. (b) Karr column.

Solvent candidates screening using continuous extraction method

Residual protein removal by membrane ultra-filtration

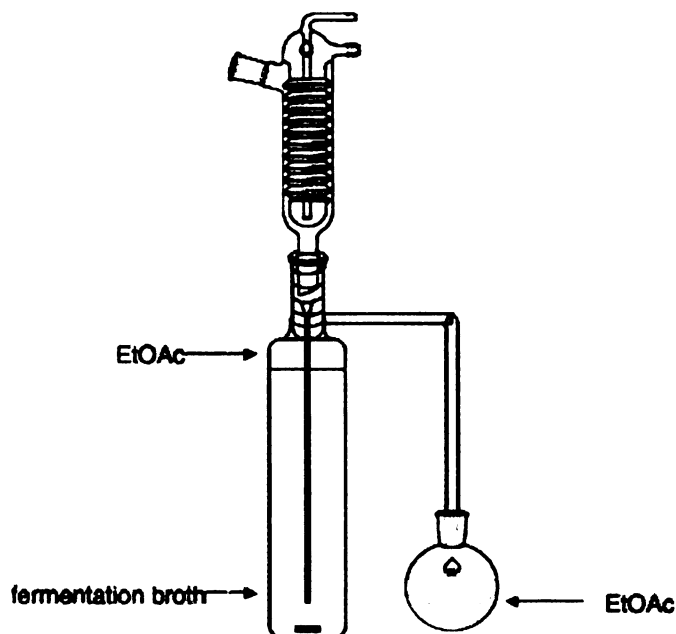


Figure 92. Glass continuous liquid-liquid extractor (Sigma-Aldrich).

Fermentation broth used in this study was obtained by the cultivation of *E. coli* microbe WN13/pWN7.126B. Cells were removed by centrifugation at 14,000 rpm for 10 min. Extraction experiments were carried out in a glass continuous liquid-liquid extraction apparatus obtained from Sigma-Aldrich. A typical extraction started by filling the extraction tower with 100 mL of fermentation broth. The solvent of interest was then poured into the same tower until it reached the side-arm. A 250 mL round-bottomed flask filled with 100 mL of the same solvent was attached to the other end of the side-arm. Finally, a condenser was fitted to the top of the extraction tower. The extraction began by refluxing the solvent inside the round-bottomed flask. The extractant evaporated,

condensed and directed back into the dispersion created inside the extraction tower by vigorous magnetic stirring. Due to the density difference, D-1,2,4-butanetriol-rich extractant was eventually separated from the feed and collected in the round-bottom flask. At the end of the extraction, D-1,2,4-butanetriol was accumulated with extractant inside the round-bottom flask. Subsequent removal of solvent was achieved by rotary evaporation and subjected to ^1H NMR and gas chromatographic (GC) analyses.

Direct D-1,2,4-butanetriol extraction from cell-free fermentation broth using ethyl acetate as solvent was unsuccessful due to protein precipitation (Table 28, entry 1). Methods were developed to remove residual soluble proteins in the fermentation broth. Earlier research in the Frost group clarified broth by adjusting the pH of the cell-free broth to 2.5 using concentrated H_2SO_4 . This method was used to treat D-1,2,4-butanetriol broth and the protein precipitate was pelleted using centrifugation. Subsequent extraction of this acidified broth led to decomposition of D-1,2,4-butanetriol (Table 28, entry 2). In another attempt, this acidified broth was neutralized using 10 N KOH and the extraction started once again. In this case, D-1,2,4-butanetriol was recovered in the solvent without decomposition (Table 28, entry 3). As an alternative approach, charcoal treatment was found to be an attractive way to remove protein by adsorption (Table 28, entry 4). It eliminated the extra salt stream created during acidification and neutralization of the fermentation broth. As an additional advantage, charcoal addition was also found useful in reducing the product color and unpleasant odor. The optimal way to remove protein was achieved by membrane ultra-filtration. Fitting an Amicon concentrator with a 10,000 MWCO Miomax membrane (Millipore) yielded clarified broth without extra process cost

due to acid/base or charcoal addition (Table 28, entry 5). Either charcoal or membrane-treated broth yielded D-1,2,4-butanetriol upon ethyl acetate extraction.

Table 28. Methods for residual protein removal from fermentation broth.

entry	conditions	recovery
1	direct extraction (pH 7)	failed
2	protein was precipitated at pH 2.5 (H ₂ SO ₄) extracted for 48 h, replaced w/fresh EtOAc every 12 h no organics left in the raffinate	BT decomposed
3	protein was precipitated at pH 2.5 (H ₂ SO ₄) pH adjusted to 6.5 w/ KOH extracted for 24 h, NMR samples were taken every 12 h	BT recovered
4	10 g/L charcoal was added to the cell-free broth filtered through Celite extracted for 24 h, NMR samples were taken every 12 h	BT recovered
5	cell-free broth was subjected to ultra-filtration extracted for 24 h, NMR samples were taken every 12 h	BT recovered

Liquid-liquid extraction using a single solvent system

A list of common industrial solvents was obtained from the Kirk-Othmer Encyclopedia of Chemical Technology.⁹ Potential solvent candidates for the D-1,2,4-butanetriol extraction were identified based on several considerations. Solvent boiling point lower than 120 °C was preferred to avoid potential D-1,2,4-butanetriol decomposition. Water solubility was certainly another important factor in liquid-liquid extraction. Organic solvents that are miscible with water could not be used. A difference in the density of saturated liquid phases should be as large as possible for physical separation of the two phases. As the solvent must be recovered for reuse, the formation of an azeotrope with the extracted solute is not desirable. The solvent should also be

inexpensive and non-toxic if possible. Lastly, the solvent should be chemically stable and inert to other components in the system.

Table 29. Single solvent candidates.

entry	solvent	boiling point (°C)	water solubility (wt%)
1	ethyl acetate	77	3
2	1-butanol	118	9
3	2-butanol	98	12.5
4	3-pentanone	101.5	5
5	2-butanone	80	25
6	nitromethane	101	10

Single solvent candidates were screened and are summarized in Table 30. The extraction experiments were performed using the same protocol as described previously. The 2-butanol extraction experiment was sampled every hour, while the rest of the solvent candidates were sampled every 12 h. Extractions were performed for a maximum of 24 h when needed. Both GC based and isolated recovery were calculated using these samples. The total recovery was obtained by adding up the recoveries at different time-points. The purity was based on the gas chromatogram of the isolated D-1,2,4-butanetriol. At this early stage of our study, final D-1,2,4-butanetriol quality was of primarily importance. The rate of recovery came second and was tuned at a later stage of this work. Experimental results revealed that ethyl acetate (Table 30, entry 1) and 2-butanone (Table 30, entry 5) possessed excellent selectivity toward D-1,2,4-butanetriol, however, the extractions were performed at a relatively low rate. On the other hand, extractions with 1-butanol (Table 30, entry 2) or 2-butanol (Table 30, entry 3) as solvent achieved at least ten-fold improvement in the rate of recovery. As a drawback, impurities were also extracted into these two solvents. Extraction using 3-pentanone (Table 30, entry 4) was

unsuccessful due to poor recovery (40% GC based). Attempts to use nitromethane as a solvent, which has a higher density than water also failed. A dispersion between the feed and the solvent was not formed during extraction with nitromethane. During the study, inconsistency between the GC-based and isolated total recovery was observed, indicating the accumulation of NMR and GC inactive impurities (ie. poly-glycols¹⁰) in the product after prolonged heating. In addition, attempts to scale up the process using a 1 L glass liquid-liquid extractor experienced a significant decrease in the rate of recovery. A dark yellowish oil was obtained even using ethyl acetate as solvent. These observations indicate that continuous extraction using a single solvent system is not a good method. To address this issue, other extraction alternatives using a solvent-blend system and counter-current extraction column were explored.

Table 30. Summary of single solvent screening experiments.

entry	solvent	time (h)	BT extracted ^a (g)		BT recovery (%)		total recovery (%)		purity ^b (%)
			GC	isolated	GC	isolated	GC	isolated	
1	ethyl acetate	12	0.24	0.22	44	40	64	51	>99
		24	0.11	0.06	20	11			
2	1-butanol	12	0.53	0.92	96	167	96	167	>90
3	2-butanol	1	0.29	0.36	52	65	99	113	>90
		2	0.26	0.26	47	48			
4	3-pentanone	12	0.14	0.22	25	40	40	67	>99
		24	0.08	0.15	15	27			
5	2-butanone	12	0.41	0.58	75	105	75	105	>99

^aThe charcoal clarified fermentation broth (100 mL) contained 0.55 g D-1,2,4-butanetriol.

^bPurity was calculated based on the gas chromatogram of isolated D-1,2,4-butanetriol.

Liquid-liquid extraction using a solvent-blend system

Given that ethyl acetate provided the best selectivity of D-1,2,4-butanetriol as discovered in the previous section, ethyl acetate and 2-butanone mixture were studied as the first solvent-blend system. Three experiments were setup to evaluate different concentration of 2-butanone (0%, 25% and 50%) in ethyl acetate. The extractions were carried out for 24 hours using the same protocol. A sample was taken every 12 h and analyzed by ^1H , ^{13}C NMR and GC to quantify D-1,2,4-butanetriol and to determine the product purity. All solvent system were successfully used to extract D-1,2,4-butanetriol in 99% purity (Figure 93). Increase in product recovery was observed upon increasing 2-butanone concentration as shown in Table 31.

Ethyl acetate extraction with ethanol as a co-extractant in the charcoal clarified fermentation broth was also studied. Different amount of ethanol (2%, 5%, 10%) was added into the clarified broth before extraction started. The extraction experiments were carried out for 24 h and was sampled every 12 h. Solvent was evaporated by rotary evaporation under reduced pressure and the product was analyzed by ^1H , ^{13}C NMR and GC. Similar to the 2-butanone/ethyl acetate case, EtOAc/EtOH extracted D-1,2,4-butanetriol with over 99% purity (Figure 94). In addition, 87% isolated recovery was successfully obtained by extracting clarified broth with 10% ethanol (Table 32). Although ethyl acetate/ethanol extraction was proven to be a better approach relative to other tested solvent systems, recycling two solvents was not economically viable for an industrial process. Furthermore, scaling up a continuous liquid-liquid extraction process into production scale is problematic. Nevertheless, solvent candidates screening using the glass continuous extractor provided fundamental knowledge towards extracting D-1,2,4-

butanetriol from the fermentation broth, leading to the development of a scalable extraction process using a Karr column.

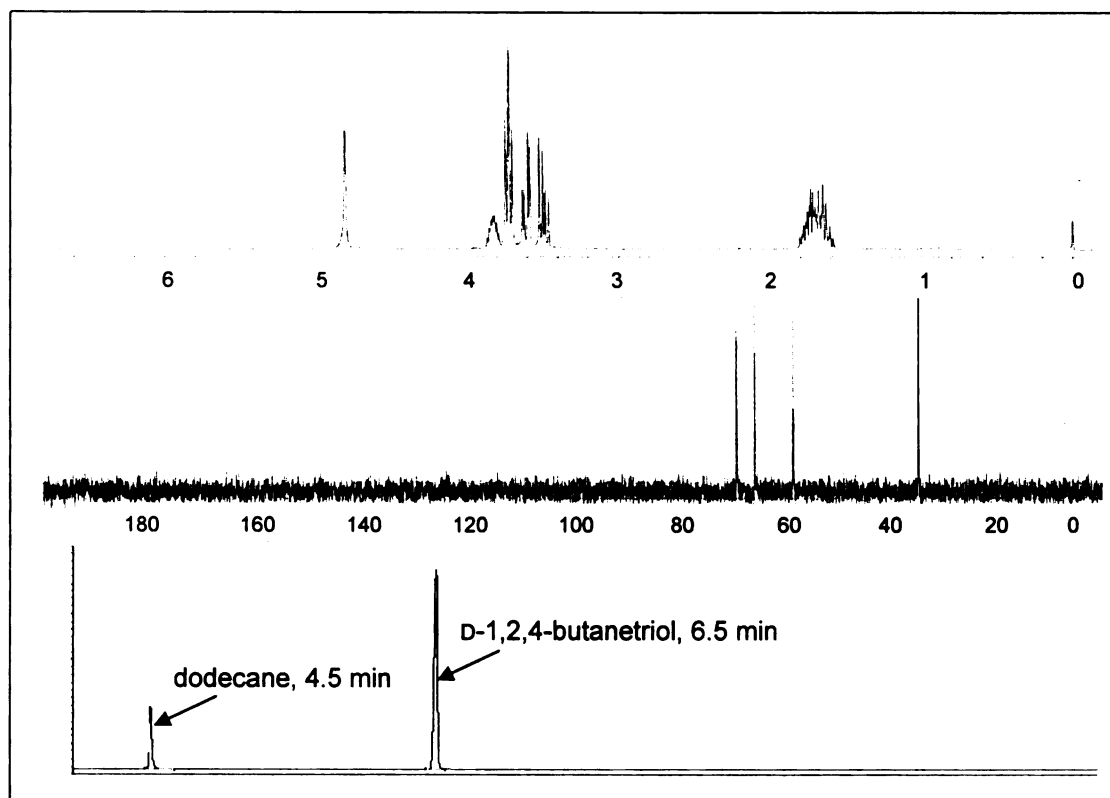


Figure 93. D-1,2,4-Butanetriol purified by 50% 2-butanone in ethyl acetate extraction. (upper) ^1H NMR spectrum, in ppm; (middle) ^{13}C NMR spectrum, in ppm; (lower) GC chromatogram.

Table 31. Liquid-liquid extraction using 2-butanone/ethyl acetate solvent-blend.

entry	2-butanone In EtOAc (%)	time (h)	BT extracted ^a (%)		BT recovery (%)		total recovery (%)		purity ^b (%)
			GC	isolated	GC	isolated	GC	isolated	
6	0	12	0.083	0.07	32	27	50	45	>99
		24	0.047	0.047	18	18			
7	25	12	0.091	0.1	36	39	58	66	>99
		24	0.057	0.07	22	27			
8	50	12	0.12	0.13	46	50	67	69	>99
		24	0.055	0.05	21	19			

^aThe charcoal clarified fermentation broth (100 mL) contained 0.26 g D-1,2,4-butanetriol.

^bPurity was calculated based on the gas chromatogram of isolated D-1,2,4-butanetriol.

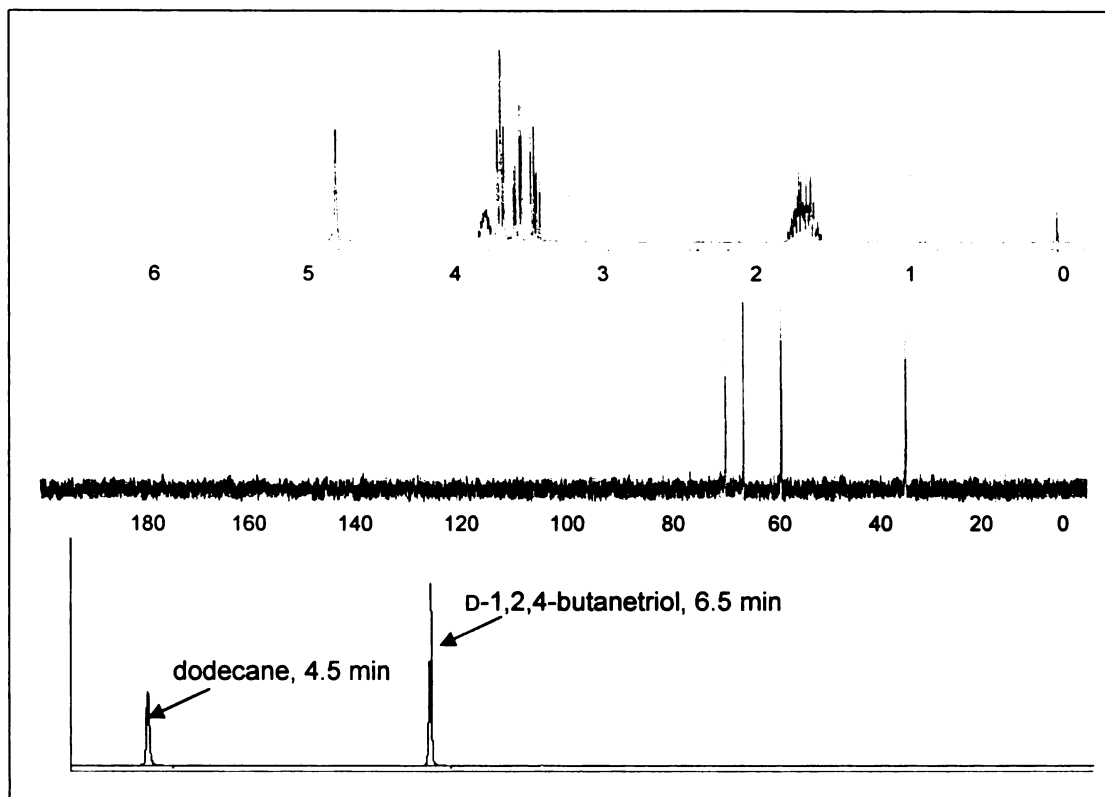


Figure 94. D-1,2,4-Butanetriol purified by 10% ethanol in ethyl acetate extraction. (upper) ^1H NMR spectrum, in ppm; (middle) ^{13}C NMR spectrum, in ppm; (lower) GC chromatogram.

Table 32. Liquid-liquid extraction using ethanol/ethyl acetate solvent blend.

entry	Ethanol in broth (%)	time (h)	BT extracted ^a (g)		BT recovery (%)		total recovery (%)		purity ^b (%)
			GC	isolated	GC	isolated	GC	isolated	
9	0	12	0.09	0.09	41	41	55	59	>99
		24	0.031	0.04	14	18			
10	2	12	0.106	0.1	48	46	63	55	>99
		24	0.033	0.02	15	9			
11	5	12	0.104	0.14	47	64	70	82	>99
		24	0.054	0.04	24	18			
12	10	12	0.11	0.14	50	64	80	87	>99
		24	0.064	0.05	30	23			

^aThe charcoal clarified fermentation broth (100 mL) contained 0.22 g D-1,2,4-butanetriol.

^bPurity was calculated based on the gas chromatogram of isolated D-1,2,4-butanetriol.

Development of a scalable extraction process using a bench-top Karr column

Bench-top Karr column unit setup

A Karr reciprocating column (Koch Modular Process System, KMPS) is a counter-current column extractor that is designed to handle difficult extraction systems having a low interfacial tension and tends to emulsify. The bench-top Karr column unit is widely used for feasibility testing during early stages of process development. By using this apparatus, overall process feasibility, economics and different solvents will be evaluated. In most applications, data generated in this bench-top unit will form the basis for a larger scale pilot test program or even the design of a production scale column.

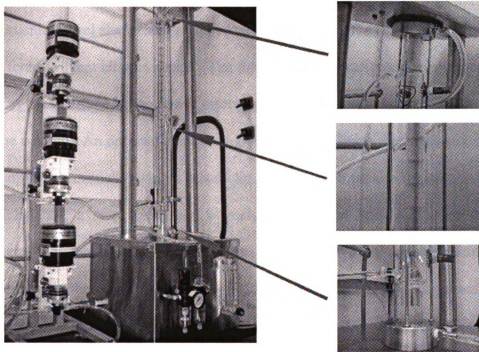


Figure 95. Benchtop Karr column in action.

The bench-top Karr column was set up in a laboratory fumehood. Three Fluid Metering QD-Pump drive modules fitted with RH pump heads were used to control liquid flows within this extraction system. This pump was setup to deliver a liquid flow from 8 mL/min to 86.25 mL/min to the system. Precise control over the flow could be adjusted by a control ring graduated in 450 divisions. Two among these variable speed pumps were used for charging feed and solvent, and a special Teflon coated Tygon tubing was employed to deliver solvent to the column. The agitator assembly was powered by an air driven motor to drive the agitation. A needle valve flow meter was provided to regulate the gas flow, therefore providing an avenue to vary the agitation speed of the plate stack assembly. In addition, the agitation speed (stroke per minute, spm) was monitored by a handheld touchless digital tachometer. A Teflon and a stainless steel plate stack assembly built on stainless steel shafts were supplied with the system. They were interchangeable to meet different solvent characteristics. The third variable speed pump was responsible for controlling an interface in the lower disengaging chamber and discharging the raffinate out of the column. An optional control over very small raffinate discharge rate could also be accomplished manually using a metal clip. To maintain precise input/output rate and material balance, a stop watch and a graduated cylinder were also used during the experiment. The total volume of this column is 220 mL, and the operating throughput is expected to be in the range between 15 – 120 mL/min. Manufacturer suggested agitation speed is between 100 – 400 spm. While systems with a large density difference between the phases or high interfacial tension will be expected to operate in the high end of this range, systems with opposite properties will be operated in the low end of this range.

Distribution coefficient for solvent candidates

Transition of the earlier knowledge obtained using a continuous liquid-liquid extractor to a scalable process using a counter-current Karr extraction column can be facilitated by understanding distribution coefficients between solvent candidates and D-1,2,4-butanetriol. The distribution coefficient (m) is defined using the following formula: $m = y_a / x_a$, where a is the solute, y_a is the concentration of component a in the solvent and x_a is the concentration of component a in the raffinate. The distribution coefficient for our solvent candidates was calculated from the liquid-liquid equilibrium data set, which had been generated experimentally using known concentration of authentic 1,2,4-butanetriol in water and the solvent of interest. A typical liquid-liquid equilibrium data set was obtained by vigorously mixing 1,2,4-butanetriol solution (50 mL) and solvent (50 mL) in a separatory funnel and resolved back into two phases. The solvent was collected and the concentration of 1,2,4-butanetriol in solvent and raffinate were determined using GC. The raffinate was extracted again twice more and the distribution coefficient calculated from each extraction was averaged at the end of the experiment. This averaged number represents the distribution coefficient of a particular solvent. This method was applied on eight solvent candidates and their distribution coefficients are summarized in Table 33.

Table 33. Distribution coefficients for various solvent candidates.

entry	solvent	m	entry	solvent	m
1	2-butanol	0.24	5	1-pentanol	0.06
2	1-butanol	0.14	6	2-pentanol	0.06
3	iso-butanol	0.10	7	2-butanone	0.03
4	isoamyl alcohol	0.07	8	ethyl acetate	0.004

Abbreviation: distribution coefficient (m)

According to earlier experiments, ethyl acetate provided the best selectivity over the others toward extracting D-1,2,4-butanetriol from the fermentation broth. However, it is not suitable to be used in an industrial column extraction process due to its small distribution coefficient towards D-1,2,4-butanetriol (Table 33, entry 8). Although 2-butanol and 1-butanol were less selective toward D-1,2,4-butanetriol that yielded products with only 90% purity, they possess distribution coefficients that are two orders of magnitudes greater than that using ethyl acetate as solvent (Table 33, entries 1 and 2). Attempts to explore solvent alternatives yielded an extra five solvent candidates as iso-butanol, isoamyl alcohol, 1-pentanol, 2-pentanol and 2-butanone. From a distribution coefficient standpoint, however, none of them shows a better potential over 2-butanol (Table 33, entries 3 – 7). To this end, 2-butanol was the only solvent employed in the following Karr column extraction experiments.

Start-up conditions for Karr column extractions

There are several conditions associated with the use of the bench-top Karr column. Careful optimization of all parameters will not only maximize the performance of this laboratory scale column but also provide a solid foundation to perform tests in a medium scale pilot size column that is available on a rental basis at the KMPS facility located in Houston, Texas. The pilot test is important to determine other system limitations before significantly increasing the capacity by building an operating column.

The initial conditions for the Karr column extraction are different for each specific case and were determined for the D-1,2,4-butanetriol extraction based on the operating manual and consultation with Mr. Donald J. Glatz at KMPS. Prior to start-up, it was

important to determine the continuous phase of the system. Generally speaking, the continuous phase is the higher flowing phase and will be filled in the column before the extraction begins. In contrast, the lower flowing phase will be dispersed. Conventionally, nearly all old cases involving fermentation broth column extraction, the broth (the feed) is used as the continuous phase, while the solvent is the disperse phase.¹¹ The material used in making the plate stack poses another option based on the ‘wetting’ characteristics of the system. If coalescing of the dispersed phase occurs on the plates and shaft spacers, then either the phase should be reversed or the plate stack should be replaced with the alternate plate stack made from a different material. In general, stainless steel works best for aqueous phase continuous extractions, while teflon will be used for organic-based phases. In our initial plan, since fermentor broth was the continuous phase, a stainless steel shaft was used. The initial agitation speed is usually defined to void flooding, which is a phenomenon where the upward or downward flow of the disperse phase ceases to form a second interface in the column. This value was set to 200 rpm for the D-1,2,4-butanetriol extraction. By knowing that 2-butanol has a low distribution coefficient at 0.24 for D-1,2,4-butanetriol extraction, it was therefore reasonable to adjust the starting total throughput to the low end, which was 15 mL/min. To select a reasonable solvent to feed ratio is of almost importance and this requires knowledge of the solvents distribution coefficients. It is not difficult to understand that a high recovery can be obtained with a high solvent to feed ratio. However, using a large amount of solvent will certainly increase the cost of the process. These two factors compete against each other and a formula for the extraction factor (E) exists to make a reasonable judgment on the initial amount of solvent: $E = m \times (\text{vol. of solvent/vol. of feed}) \geq 1$.¹¹ For instance, in the case

of 2-butanol extraction, the minimum solvent to feed ratio would be 4:1. After operating the Karr column at these initial conditions, evaluations were made for the extraction performance. Related parameters were changed accordingly to maximize the efficiency of the Karr column extraction.

Preliminary extraction attempts using Karr column

A typical column extraction started by filling up the column with the continuous phase, either with clarified fermentation broth or fresh solvent. Once the column was full of liquid, the agitator was turned on at a slow rate (50 rpm). The pumps for the feed and solvent were then turned on with the flow rate being set as required to establish an interface. A stable interface was achieved by controlling the bottom raffinate take-off rate. The agitation was then brought up slowly to the initial set point. During the first run after a shutdown period, the column was allowed to operate for five column turnovers before sampling. A turnover is defined as the column volume divided by the combined flow rate. Solvent and raffinate (5 mL) were obtained for each set of parameters and analyzed by ^1H , ^{13}C NMR and GC. All the NMR spectra were obtained on a 500 MHz spectrometer for better resolution. The results were used to calculate the recovery and purity. For each additional run on the same day, samples were collected after three column turnovers.

In the first five entries, the solvent to feed ratio was set to 3:1 to allow precise broth flow rate and maintained a low throughput at 15 mL/min (feed flow rate below 4 mL was not accurately controlled by the pump). Among these experiments, attempts to extract cell-free broth as a continuous phase failed shortly after five minutes due to

severe protein precipitation that resulted in column blockage (Table 34, entry 1). Further attempts using different clarified broths were not successful due to entrainment (a portion of feed was carried by the solvent and collected at the wrong end), leading to the change of the continuous phase to solvent (Table 34, entry 2 and 3). In entry 4, it was discovered that using solvent as the continuous phase provided a more defined interface and resulted in stable extraction. Unfortunately, acidified broth yielded D-1,2,4-butanetriol with serious contamination, as indicated earlier during the continuous extraction studies. Karr column extraction of neutralized broth under the same conditions led to the isolation of 1,2,4-butanetriol with 86% purity (Table 34, entry 5). Unfortunately, the recovery was only 15%. Increasing the agitation to 260 spm was unsuccessful and resulted in entrainment (Table 34, entry 6). Doubling the total throughput to 30 mL/min did not make an impact on the recovery (Table 34, entry 7), however, reducing the feed flow rate by half increased the recovery to 22% (Table 34, entry 8). At this point, it was interesting to investigate the recovery with even smaller feed flow rates. To this end, a Q pump (Fluid Metering, Inc.) was setup to provide a finely controlled feed delivery to the column. By lowering the total throughput back to 15 mL/min with a broth flow rate at 2.5 mL/min, a 43% recovery was successfully achieved using the Karr column (Table 34, entry 9). More importantly, entries 6 – 8 yielded D-1,2,4-butanetriol with purity around 90%.

Table 34. Preliminary Karr column extraction studies.

entry	broth ^a	continuous phase	S/F ^b	agitation (spm ^c)	flow rate (mL/min)			recovery (%)	purity (%)
					total	broth	solvent		
1	cell-free ^d	broth	3:1	200	15	4	11	column blocked	
2	acid ^e	broth	3:1	200	15	4	11	entrainment	
3	acid/base ^f	broth	3:1	200	15	4	11	entrainment	
4	acid ^e	solvent	3:1	200	15	4	11	poor BT quality	
5	acid/base ^f	solvent	3:1	200	15	4	11	15	86
6	acid/base ^f	solvent	3:1	260	15	4	11	entrainment	
7	acid/base ^f	solvent	3:1	200	30	8	22	14	87
8	acid/base ^f	solvent	7:1	200	32	4	28	22	92
9	acid/base ^f	solvent	4:1	200	12.5	2.5	10	43	92

^aFor entries 1 – 3, the titer of D-1,2,4-butanetriol is 11.0 g/L; for 4 – 8, 11.7 g/L; for 9, 10.6 g/L.

^bSolvent to feed ratio (S/F)

^cStroke per minute (spm), stainless steel plate stack was used in all experiments.

^dFermentation broth was cell-free, without underwent protein precipitation steps.

^eBroth residual protein was precipitated at pH 2.5.

^fAcidified broth was adjusted back to pH 7 after protein precipitation.

Further optimization of Karr column extraction

A new set of experiments was designed to optimize the Karr column extraction in an attempt to increase the recovery. Given the earlier experience, several conditions were adjusted. First, protein-free clarified fermentation broth was obtained from membrane ultrafiltration as discussed earlier. This method avoided the introduction of salt streams and reduced the cost by eliminating unnecessary use of acid and base. Second, coalescing of the dispersed phase (feed) was observed in the previous experiments. Interfacial area was therefore decreased as expected. To address this issue, the stainless steel plate stack was replaced by the one made of Teflon. Third, broth delivery was tightly controlled by the pump that was made to handle low flow rate (< 10 mL). More accurate and reproducible extraction data were thus expected.

Attempts to extract D-1,2,4-butanetriol at a solvent to feed ratio at 4:1, where the flow rate for broth and solvent were at 2 mL/min and 8 mL/min, respectively, recovered the target product in 70% (Table 35, entry 1). Given this success, the total throughput were increased and two extractions were run at 22 and 30 mL/min without significant decrease of recovery (Table 35, entry 2 and 3). Increasing the throughput beyond this point was not encouraging. A 54% recovery of D-1,2,4-butanetriol was obtained with a total throughput at 40 mL/min (Table 35, entry 4), and 50 mL/min throughput ended up in entrainment (Table 35, entry 5). Further attempts to increase the recovery by increasing solvent to feed ratio to 8:1 and 12:1 yielded little success. D-1,2,4-Butanetriol was recovered in 72% and 78%, respectively (Table 35, entry 6 and 7). The exhaustive use of solvent in these cases would increase the cost of the process and therefore was not preferred. From these experiments, it is clear that obvious advantages exists for filtering off protein by ultrafiltration. Using the Teflon plate stack for agitation certainly created a better dispersion to facilitate improved extraction performance.

Table 35. Further optimization for Karr column extraction.

entry	broth ^a	continuous phase	S/F ^b	agitation (spm ^c)	flow rate (mL/min)			recovery (%)	TS ^d
					total	broth	solvent		
1	ultrafil.	solvent	4:1	200	10	2	8	70	2.5
2	ultrafil.	solvent	4:1	200	22	4.5	18	74	2.5
3	ultrafil.	solvent	4:1	200	30	6	24	64	1.5
4	ultrafil.	solvent	4:1	200	40	8	32	54	1.3
5	ultrafil.	solvent	4:1	200	50	10	40	entrainment	
6	ultrafil.	solvent	8:1	260	56	6	50	72	1.2
7	ultrafil.	solvent	12:1	200	78	6	72	78	<1

^aThe titer of D-1,2,4-butanetriol is 11.9 g/L. Residual protein was membrane filtered.

^bSolvent to feed ratio (S/F)

^cStroke per minute (spm), Teflon plate stack was used in all experiments.

^dTheoretical stage (TS), estimated by using the Klemser type plot.

The operating efficiency of Karr column extractions can also be evaluated by using the Kremser equation.^{7,11} The Kremser equation can be used to calculate the number of theoretical stages (N_s) and the equation is defined as follows (Figure 96), where X_F is the mass concentration of solute in the feed, X_N is the mass concentration of solute in the raffinate, Y_S is the mass concentration of solute in the solvent, m is the distribution coefficient, S/F is the mass ratio of solvent rate to feed rate. Finally, E is the extraction factor that was discussed earlier. As an alternative, the number of theoretical stages can be easily determined using the Kremser type plot (Figure 97).¹¹ Theoretical stages for the Karr column extraction experiments were determined and shown in Table 35. According to all the evidence obtained, entry 2 represents the most efficient Karr column extraction process by having a 22 mL/min throughput, a theoretical stage of 2.5 and a recovery at 74%. These conditions will be used for future process scale-up.

Final step toward purified D-1,2,4-butanetriol

The Karr column was successfully employed to yield D-1,2,4-butanetriol in 90% purity using 2-butanol as solvent. Under optimized conditions, recovery of this product achieved up to 74% at a solvent to feed ratio of 4:1. The GC chromatographic traces for the extraction process are shown in Figure 98 using dodecane as an internal standard. Negative adsorption using a small plug of Dowex 1×8 (OH^-) resin was employed in an attempt to obtain D-1,2,4-butanetriol with higher purity.

Figure 96. Kremser equation.

$$N_S = \frac{\ln \left[\left(\frac{X_F - \frac{Y_S}{m}}{X_N - \frac{Y_S}{m}} \right) \left(1 - \frac{1}{E} \right) + \frac{1}{E} \right]}{\ln E}$$

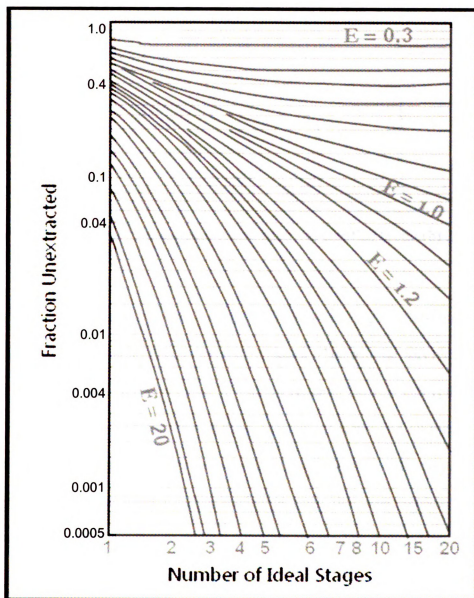


Figure 97. Kremser type plot.¹¹

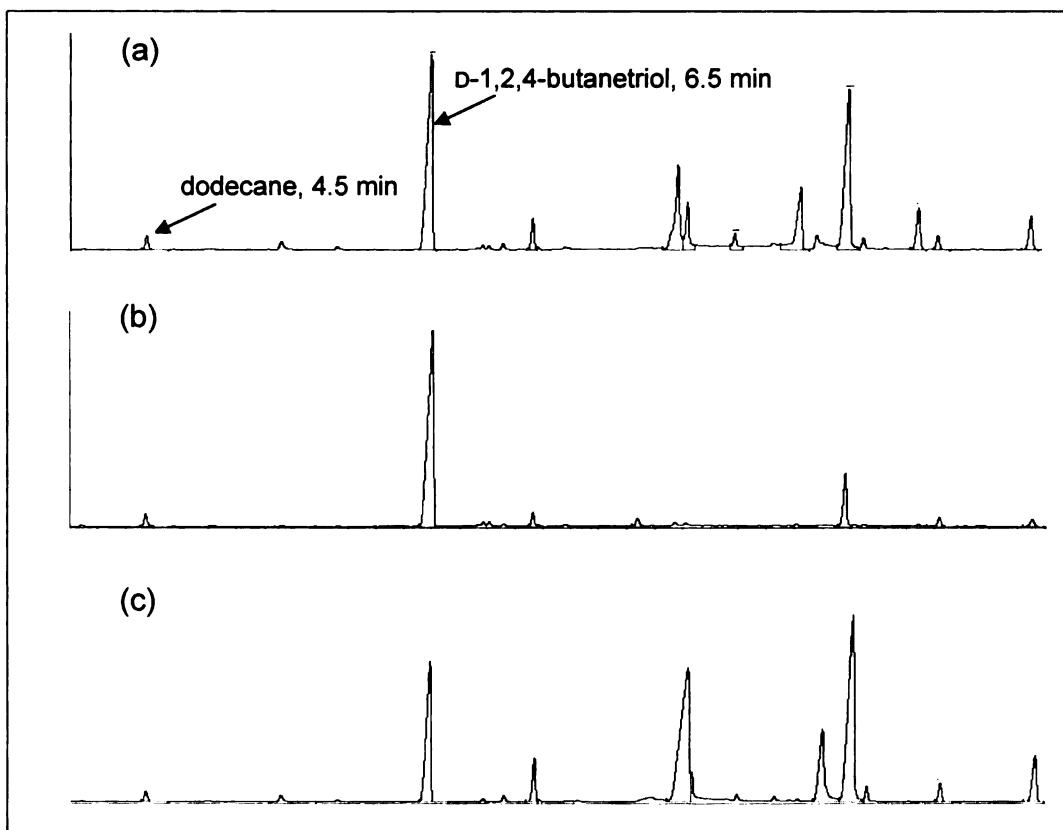


Figure 98. GC traces for Karr column extraction. (a) clarified broth; (b) 2-butanol extract; (c) raffinate.

2-Butanol extract eluted from the Karr column containing D-1,2,4-butanetriol (500 mL) was evaporated to dryness under reduced pressure. The residue was redissolved into 10 mL H₂O and filtered through a plug of anion-exchange Dowex 1×8 (OH⁻) resin (30 mL). The resin was washed with two column volumes of water and all the flowthrough was allowed to filter through another cation-exchange Dowex 50 (H⁺) column (30 mL). The second column contained a cation-exchange resin that was expected to remove residual metal ion contamination in the D-1,2,4-butanetriol. The filtrate was evaporated to dryness under reduced pressure to yield >99% pure salt-free D-1,2,4-butanetriol. Combining the Karr column extraction and the resin filtration methodologies, 7 g D-1,2,4-

butanetriol was successfully purified from a 1 L fermentation broth of *E. coli* WN13/pWN7.126B with >99% purity (Figure 99).

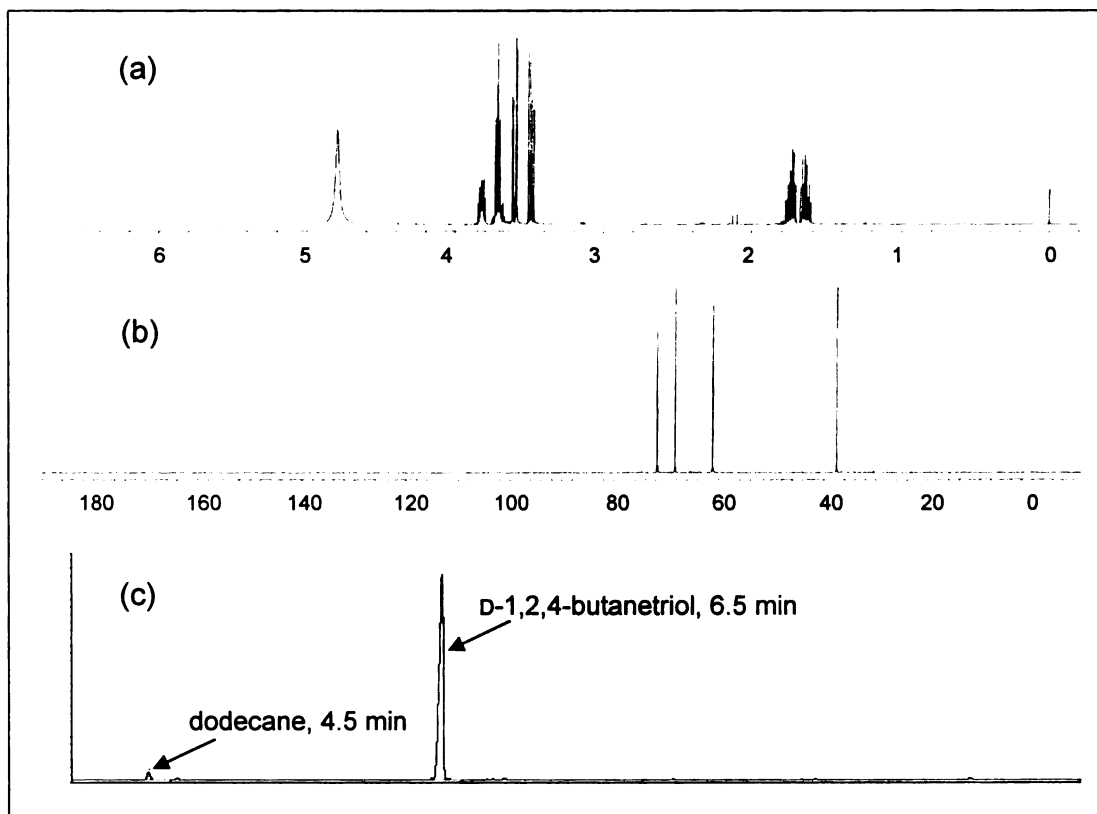


Figure 99. Large scale purification of *E. coli* WN13/pWN7.126B fermentation broth. (a) ^1H NMR spectrum, in ppm; (b) ^{13}C NMR spectrum, in ppm; (c) GC chromatogram.

Alternative method for D-1,2,4-butanetriol purification

Given success in using Dowex 1 \times 8 (OH^-) anion-exchange resin as a second purification step to obtain 99% pure D-1,2,4-butanetriol, it became interesting to investigate whether Karr column extraction was indeed necessary for our purpose. To this end, D-1,2,4-butanetriol was purified directly from membrane clarified fermentation broth using Dowex 1 \times 8 (OH^-) resin followed by a Dowex 50 (H^+) column. Although the amount of cation-exchange resin was proven to be similar to the earlier process, it turned out that the amount of the anion-exchange resin played an important role in the final

quality of product in this extraction-free purification. Therefore experiments were setup to optimize this parameter and are shown in Table 36. D-1,2,4-Butanetriol with >99% purity was obtained from 1 L fermentation broth using 2 L of Dowex 1×8 (OH⁻) resin. It was an eight-fold increase in the amount of resin involved comparing to the earlier process. In addition, since more resin was involved in the purification, more water was used to wash off D-1,2,4-butanetriol from the resin, leading to a requirement of 4 L of water to obtain a recovery of 96%. Evaporation of this volume of water was deemed problematic (Table 37). Finally, regenerating the Dowex 1×8 (OH⁻) resin after each purification was also found to be problematic. Ten column volumes of 5 N NaOH was used to wash off the bound material, including organic acids, salts and most importantly, some unidentified colored impurities in the fermentor broth. Exhaustive resin regeneration is known to shorten the resin lifetime, ultimately leading to an increase of product cost. MBI International purified 500 g D-1,2,4-butanetriol with >99% purity using this extraction-free method and a cost study was also prepared.¹² It costs \$164 to purify 1 kg D-1,2,4-butanetriol from the fermentation broth, while \$48 was the cost of sodium hydroxide involved in the resin regeneration. Apparently Karr column extraction using 2-butanol did not only selectively extract D-1,2,4-butanetriol over other impurities, most of the salt and coloring material were also separated from the product and discarded in the raffinate. Although extra equipment cost associated with the construction of a production scale Karr column will be involved when the process is commercialized, the solvent involved can be easily recycled by distillation. As a long term consideration, the process involving the Karr column extraction will lower the production cost by avoiding exhaustive use of expensive Dowex 1 resin and associated NaOH.

Table 36. Dowex 1×8 (OH⁻) for extraction-free purification.

resin (L resin per L broth)	purity (%)
0.5	95
1	96
1.5	97
2	99

Table 37. D-1,2,4-Butanetriol recovery from Dowex 1×8 (OH⁻).

Water (CV)	recovery (%)
0.5	2
1	76
1.5	93
2	96

Abbreviation: column volume (CV)

Future plan

Compared to the conventional purification method for polyols using distillation, Karr column extraction was proven to be useful in D-1,2,4-butanetriol purification to obtain product in a 74% recovery and 99% purity . Using liquid-liquid extraction would significantly reduce the cost associated in the product separation. In addition, residual salt and impurities would separate from the product during extraction. However, the distribution coefficient of D-1,2,4-butanetriol into extraction solvent appears to be too small to make the process efficient enough. A similar extraction approach using the Karr column might serve as a solution to tackle this issue. D-1,2,4-Butanetriol is not easily extracted into organic solvent due to the extensive hydrogen bonds formed between the molecule and water. A possible avenue toward easier recovery of 1,2,4-butanetriol would therefore be to react them with a suitable aldehyde to form the corresponding acetal

(Figure 100). Simultaneously, this aldehyde will also serve as solvent for the liquid-liquid extraction process. By lowering the polarity of the compound, this D-1,2,4-butanetriol derivative will therefore possess a larger distribution coefficient and readily extract into the aldehyde solvent. This acetal cyclization only selectively occurs with compounds having hydroxyl groups attached, therefore additional purification will be expected. In the end, the aldehyde will be recycled by distillation. This reactive extraction methodology had been successfully applied toward the recovery of 1,3-propanediol from a dilute stream.¹³ With the similarity of D-1,2,4-butanetriol and 1,3-propanediol, it is believed that reactive extraction will apply to D-1,2,4-butanetriol. Solvent screening using different aldehydes involving acetaldehyde, isobutyraldehyde and benzaldehyde is currently underway.

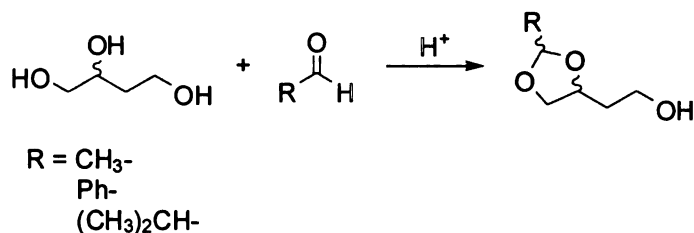


Figure 100. Cyclic acetal formation of 1,2,4-butanetriol.

REFERENCE

- ¹ Brockmann, R.; Jeromin, L.; Johannisbauer, W.; Meyer, H.; Michel, O.; Plachenka, J. US 4,655,879, 1987.
- ² (a) Ames, T. T. US 6,361,983, 2002. (b) Haas, T.; Wiegand, N.; Arntz, D. US 5,334,778, 1994.
- ³ Molefe, M. N. *PhD Thesis*, Michigan State University, 2005.
- ⁴ See Chapter 2.
- ⁵ (a) Robbins, L. A. *Chem. Eng. Prog.* **1980**, 76, 58. (b) <http://www.cheresources.com/extraction.html>
- ⁶ Belter, P. A.; Cussler, E. L.; Hu, W.-S. *Bioseparations: Downstream Processing for Biotechnology*; Wiley-VCH: Weinheim, 1988.
- ⁷ Gu, T. In *Handbook of Bioseparation*; Ahuja, A. Ed.; Academic Press: NY, 2000; p329.
- ⁸ <http://www.liquid-extraction.com/default.htm>
- ⁹ Sullivan, D. A.; Shell Chemical Company In *Kirk-Othmer Encyclopedia of Chemical Technology*; Wiley: 1997, DOI: 10.1002/0471238961.1915122219211212.a01.
- ¹⁰ Dhale, A. D.; Myrant, L. K.; Chopade, S. P.; Jackson, J. E.; Miller, D. J. *Chem. Eng. Sci.* **2004**, 59, 2881.
- ¹¹ Glatz, D. J. personal communications.
- ¹² Saffron, C.; Dodds, D.; Lau, M. K.; Frost, J. W. unpublished results.
- ¹³ Broekhuis, R. R.; Lynn, S.; King, C. J. *Ind. Eng. Chem. Res.* **1994**, 33, 3230. (b) Malinowski, J. J. *Biotechnol. Prog.* **2000**, 16, 76. (c) Hao, J.; Liu, H.; Liu, D. *Ind. Eng. Chem. Res.* **2005**, 44, 4380.

CHAPTER FIVE

Experimental

General chemistry

All reactions sensitive to air and moisture were carried out in oven and/or flame dried glassware under positive argon pressure. Air or moisture sensitive reagents and solvents were transferred to reaction flasks fitted with rubber septa via syringes or cannula. Solvents were removed using either a Büchi rotary evaporator at water aspirator pressure or under high vacuum.

Reagents and solvents

CH_2Cl_2 and benzene were distilled from calcium hydride under argon. CH_3OH was distilled from sodium metal under argon and stored over Linde 4 Å molecular sieves under argon. THF and diethyl ether were distilled under nitrogen from sodium benzophenone ketyl. DMF, DMSO, hexanes and acetone were dried over activated Linde 4 Å molecular sieves under nitrogen. Water was glass distilled and deionized. All reagents and solvents were used as available from commercial sources or purified according to published procedures. Organic solutions of products were dried over anhydrous MgSO_4 .

Chromatography

Gas chromatography was performed on an Agilent 6890N equipped with an HP-5 capillary column (30 m \times 0.25 mm \times 0.25 micron). Temperature programming began with an initial temperature of 120 °C for 3 min. The temperature was increased to 210 °C at a rate of 15 °C/min, and held at the final temperature for 1 min. The split injector was maintained at a temperature of 300 °C and the FID detector was kept at 350 °C. Samples analyzed by gas chromatography were derivatized using bis(trimethylsilyl)trifluoroacetamide and quantified relative to an internal standard of dodecane against a calibration curve.

HPLC analysis was performed on an Agilent 1100 HPLC installed with ChemStation acquisition software (Rev. A.08.03). Columns used in the enantiomeric purity analysis of 1,2,4-butanetriol include Chiralpak AD column (Daicel Chemical, 4.6 mm \times 250 mm) and Chiralpak ADH column (Daicel Chemical, 4.6 mm \times 250 mm). Protein purification utilized the same HPLC system equipped with a Pharmacia Resource Q column (6.4 mm \times 30 mm, 1 mL). Solvents were routinely filtered through 0.45- μ m membranes (Gelman Science) prior to use.

Dowex 50W \times 8-200 (H⁺) and Dowex 1 \times 8-400 (Cl⁻) were purchased from Sigma-Aldrich. Previously used Dowex 50 (H⁺) was cleaned by treatment with bromine. An aqueous suspension of resin was adjusted to pH 14 by addition of solid KOH. Bromine was added to the solution until the suspension turned a golden yellow color. Additional bromine was added (1-2 mL) to obtain a saturated solution. The mixture stood at room temperature overnight, and the Dowex 50 resin was collected by filtration and washed exhaustively with water followed by 6 N HCl. Dowex 50 (H⁺) was stored at 4 °C. AG-

1X8 (acetate form and chloride form) and hydroxyapatite Bio-Gel HTP gel were purchased from Bio-Rad. Phenylsepharose was purchased from Pharmacia. Diethylaminoethyl cellulose (DEAE) was purchased from Whatman. Ni-NTA resin was purchased from Qiagen.

Spectroscopic measurements

^1H NMR and ^{13}C NMR spectra were recorded on either a Varian VX-300 or a Varian VXR-500 FT-NMR spectrometer. Chemical shifts for ^1H NMR and ^{13}C NMR spectra were reported in parts per million (ppm) relative to sodium 3-(trimethylsilyl)propionate-2,2,3,3- d_4 (TSP, $\delta = 0.0$ ppm) with D_2O as the solvent. UV and visible measurements were recorded on a Perkin-Elmer Lambda 3b UV-vis spectrophotometer or on a Hewlett Packard 8452A Diode Array Spectrophotometer equipped with HP 89532A UV-Visible Operating Software. Measurements of multiple samples in microplates were carried out using a Benchmark microplate reader (Bio-Rad Laboratories) equipped with Microplate Manager III Macintosh Data Analysis and Kinetics Software.

Microbial strains and plasmids

Plasmid constructions were carried out in *E. coli* DH5 α and XL1-Blue.

Table 38. Microbial strains and plasmids.

strain	relavant characteristics	source
DH5 α	<i>F'</i> $\phi 80lacZ\Delta M15$ $\Delta(lacZYA-argF)U169$ <i>deoR</i> <i>recA1</i> <i>endA1</i> <i>hsdR17</i> (r_k^- , m_k^+) <i>phoA</i> λ^- <i>supE44</i> <i>thi-1</i> <i>gyrA96</i> <i>relA1</i>	Invitrogen
BL21(DE3)	<i>E. coli</i> <i>B F</i> <i>dcm</i> <i>ompT</i> <i>hsdS</i> (r_B^- m_B^-) <i>gal</i> λ (DE3)	Novagen

Table 38. (continue)

strain	relavnt characteristics	source
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacIqZΔM15 Tn10</i> (Tet)].	Stratagene
W3110	wild-type K12	CGSC
<i>Acidithiobacillus ferrooxidans</i>	wild-type	ATCC
<i>Burkholderia fungorum</i> LB400	wild-type	ARS
<i>Caulobacter crescentus</i> CB15	wild-type	ATCC
<i>Klebsiella oxytoca</i>	wild-type	ATCC
<i>Klebsiella pneumoniae</i>	wild-type	ATCC
<i>Lactobacillus brevis</i> LB18	wild-type	CNRZ
<i>Lactobacillus brevis</i> LB19	wild-type	CNRZ
<i>Mycobacterium smegmatis</i>	wild-type	ATCC
<i>Novosphingobium aromaticivorans</i>	wild-type	ATCC
<i>Pseudomonas fragi</i>	wild-type	ATCC
<i>Pseudomonas fluorescens</i> Pf5	wild-type	ATCC
<i>Pseudomonas putida</i>	wild-type	ATCC
<i>Rhodopseudomonas palustris</i>	wild-type	ATCC
WN7	W3110 <i>serA</i> yj <i>hH::FRT</i> yag <i>E::FRT</i>	ref 1
WN13	WN7 <i>xylAB::xdh-FRT-Cm^R-FRT</i>	ref 1
KIT1	W3110 <i>xylAB::xdh-adhP-P_{tac}-FRT-Cm^R-FRT</i>	Chapter 3
KIT2	WN7 <i>xylAB::xdh-FRT</i>	Chapter 3
KIT3	WN7 <i>xylAB::xdh-adhP-P_{tac}-FRT-Cm^R-FRT</i>	Chapter 3
KIT4	WN7 <i>xylAB::xdh-adhP-P_{tac}-FRT</i>	Chapter 3
KIT5	W3110 <i>P_{T5}-FRT-Cm^R-FRT</i>	Chapter 3
KIT6	KIT2 <i>P_{T5}-FRT-Cm^R-FRT</i>	Chapter 3
KIT7	KIT2 <i>P_{T5}-FRT</i>	Chapter 3
KIT8	W3110 <i>adhP::FRT-Cm^R-FRT</i>	Chapter 3
KIT9	KIT2 <i>adhP::FRT-Cm^R-FRT</i>	Chapter 3
KIT10	KIT2 <i>adhP::FRT</i>	Chapter 3
KIT16	KIT4 <i>yiaE::FRT-Cm^R-FRT</i>	Chapter 3
KIT17	KIT16 <i>ycdW::FRT-Km^R-FRT</i>	Chapter 3
KIT18	KIT4 <i>yiaE::FRTycdW::FRT</i>	Chapter 3

Table 38. (continue)

plasmid	relavant characteristics	source
pVH17	Ap ^R , <i>deoC</i> in pBR322	ATCC, ref 2
pCS120	Ap ^R , <i>P_{lac} dhaBCE</i> from <i>C. freundii</i>	ref 3
pFL1	Ap ^R , <i>dhaBCE</i> from <i>C. pasteurianum</i>	ref 4
pXY39	Ap ^R , <i>lacI^Q</i> , <i>P_{lac} pduCDE</i> from <i>S. typhimurium</i>	ref 5
pUC18	Ap ^R , <i>P_{lac}</i>	ref 6
pLOI135	Ap ^R , <i>Z. mobilis adhA</i> in pUC18	ref 7
pLOI276	Ap ^R , <i>Z. mobilis pdc</i> in pUC18	ref 29
pLOI295	Ap ^R , <i>Z. mobilis adhB</i> in pUC18	ref 8
pKD3	Ap ^R , FRT-flanked Cm ^R	ref 35
pKD4	Ap ^R , FRT-flanked Km ^R	ref 35
pKD46	Ap ^R , <i>araC</i> , <i>P_{araB}γ</i> , <i>β</i> , <i>exo</i> , <i>ts-pA101</i> replicon	CGSC ³⁵
pCP20	Ap ^R , Cm ^R , Flp ⁺ , λ cI857 ⁺	ref 34
pG-Tf2	Cm ^R , Pzt1 <i>groESL-tig</i>	Takara Bio
pRC1.55B	<i>serA</i> in pSU18	lab
pT7-7	Ap ^R , <i>P_{T7}</i>	lab ⁹
pKK223-3	Ap ^R , <i>P_{lac}</i>	ref 10
pJF118EH	Ap ^R , <i>lacI^Q</i> in pKK223-3	ref 11
pET28c(+)	Km ^R , <i>lacI^Q</i> , <i>P_{T7}</i>	Novagen
pQE30	Ap ^R , <i>lacO</i> , <i>P_{T5}</i>	Qiagen
pWN5.220A	Ap ^R , <i>dhaT</i> in pT7-7	lab ¹
pWN5.260A	Ap ^R , <i>pddABC</i> in pWN5.220A	lab ¹
pWN5.258A	Ap ^R , <i>pddABC-ddrAB</i> in pWN5.220A	lab ¹
pWN5.238A	Ap ^R , <i>lacI^Q</i> , <i>mdlC</i> in pJF118EH	lab ¹
pWN6.186A	Km ^R , <i>lacI^Q</i> , <i>P_{lac} mdlC</i>	lab ¹
pWN6.222A	Km ^R , <i>P_{lac} aatp</i> , <i>P_{T7} aadh</i> in pWN6.186	lab ¹
pWN7.096B	Ap ^R , <i>lacI^Q</i> , <i>P_{lac} mdlC</i> , <i>P_{lac} adhA</i>	lab ¹
pWN7.126B	<i>serA</i> in pWN5.238A	lab ¹
pML2.118	<i>P. fluorescens</i> decarboxylase in pJF118EH	Chapter 3
pML2.123	<i>P. aeruginosa</i> decarboxylase in pJF118EH	Chapter 3
pML2.162	<i>B. fungorum</i> decarboxylase in pJF118EH	Chapter 3
pML2.208	<i>S. coelicolor</i> decarboxylase in pJF118EH	Chapter 3
pML2.214	<i>N. aromaticivorans</i> decarboxylase in pJF118EH	Chapter 3
pML2.256	<i>R. palustris</i> decarboxylase in pJF118EH	Chapter 3
pML2.286	<i>M. smegmatis</i> decarboxylase in pJF118EH	Chapter 3
pML3.040	<i>A. ferrooxidans</i> decarboxylase in pJF118EH	Chapter 3
pML3.054	<i>P. putida mdlC</i> mutant A460I in pJF118EH	Chapter 3
pML3.062	<i>Z. mobilis pdc</i> mutant I472A in pUC18	Chapter 3
pML3.084	<i>Z. mobilis pdc</i> mutant I472S in pUC18	Chapter 3
pML3.086	<i>Z. mobilis pdc</i> mutant I472G in pUC18	Chapter 3
pML3.104	<i>Z. mobilis pdc</i> mutant I472T in pUC18	Chapter 3
pML3.171	<i>Bf</i> gene with <i>Eco</i> RI site deletion in pJF118EH	Chapter 3
pML3.200	<i>Bf</i> gene with <i>Bam</i> HI site deletion in pML3.200	Chapter 3
pML3.224	<i>P. fluorescens</i> decarboxylase in pJF118HE	Chapter 3

Table 38. (continue)

plasmid	relavant characteristics	source
pML3.225	<i>P. aeruginosa</i> decarboxylase in pJF118HE	Chapter 3
pML3.226	<i>B. fungorum</i> decarboxylase mutant in pJF118HE	Chapter 3
pML5.176	Ap ^R , <i>gldBC</i> in pJF118EH	Chapter 3
pML5.179	Ap ^R , <i>dhaT</i> in pJF118EH	Chapter 3
pML5.200	Ap ^R , <i>gldA</i> in pML5.176	Chapter 3
pML5.226	Ap ^R , <i>dhaT</i> in pML5.200	Chapter 3
pML6.090	Ap ^R , <i>dhaT</i> in pKK223-3	Chapter 3
pML6.128	Ap ^R , <i>P_{tac}-dhaT</i> in pWN5.238	Chapter 3
pML6.133	<i>serA</i> in pWN7.096B	Chapter 3
pML6.135	<i>serA</i> in pML6.128	Chapter 3
pML6.166	Ap ^R , <i>adhP</i> in pKK223-3	Chapter 3
pML6.168	Ap ^R , <i>adhE</i> in pKK223-3	Chapter 3
pML6.185	Ap ^R , <i>P_{tac}-adhP</i> in pWN5.238A	Chapter 3
pML6.195	<i>serA</i> in pML6.185	Chapter 3
pML6.255	Ap ^R , Cm ^R , <i>P_{tac}-adhP</i> in pKD3	Chapter 3
pML6.259	Ap ^R , <i>ydH</i> in pJF118EH	Chapter 3
pML6.261	Ap ^R , <i>yiaY</i> in pJF118EH	Chapter 3
pML6.263	Ap ^R , <i>ydjO</i> in pJF118EH	Chapter 3
pML6.272	Ap ^R , Cm ^R , <i>xdh</i> in pML6.255	Chapter 3
pML7.042	Ap ^R , Cm ^R , <i>P_{T5}</i> in pKD3	Chapter 3
pML7.128	Km ^R , <i>mdlC</i> in pET28c(+)	Chapter 3
pML7.135	<i>serA</i> in pML7.128	Chapter 3
pML7.166	Ap ^R , <i>groESL-tig</i> in pJF118EH	Chapter 3
pML7.175	Ap ^R , <i>groESL</i> in pWN5.238A	Chapter 3
pML7.180	<i>serA</i> in pML7.175	Chapter 3
pML7.202	<i>groESL</i> in pML7.135	Chapter 3

Storage of microbial strains and plasmids

All bacterial strains were stored at -78 °C in glycerol. Plasmids were transformed into DH5α for long-term storage. Glycerol samples were prepared by adding 0.75 mL of an overnight culture to a sterile vial containing 0.25 mL of 80% (v/v) glycerol. The solution was mixed, left at room temperature for 2 h, and then stored at -78 °C.

Culture medium

Bacto tryptone, Bacto yeast extract, nutrient broth, casamino acids, agar, and MacConkey agar base were purchased from Difco. Casein hydrolysate was obtained from Sigma. Nutrient agar was purchased from Oxoid.

All solutions were prepared in distilled, deionized water. LB medium¹² (1 L) contained Bacto tryptone (10 g), Bacto yeast extract (5 g), and NaCl (10 g). LB-glucose medium contained glucose (10 g), MgSO₄ (0.12 g), and thiamine hydrochloride (0.001 g) in 1 L of LB medium. LB-freeze buffer contained K₂HPO₄ (6.3 g), KH₂PO₄ (1.8 g), MgSO₄ (1.0 g), (NH₄)₂SO₄ (0.9 g), sodium citrate dihydrate (0.5 g) and glycerol (44 mL) in 1 L of LB medium. M9 salts¹² (1 L) contained Na₂HPO₄ (6 g), KH₂PO₄ (3 g), NH₄Cl (1 g), and NaCl (0.5 g). M9 minimal medium¹² contained D-glucose (10 g), MgSO₄ (0.12 g), and thiamine hydrochloride (0.001 g) in 1 L of M9 salts. Other M9 media contained carbon sources (10 g) including D-lactose, L-arabinose, glycerol, D-xylonate, or L-arabinonate in place of D-glucose in M9 minimal medium. For example, M9 glycerol medium contained glycerol (10 g), MgSO₄ (0.12 g), and thiamine hydrochloride (0.001 g) in 1 L of M9 salts. M9 L-arabinose medium also contained 0.4% (w/v) casamino acids, which were added into the M9 salts solution before sterilization in an autoclave. M9 medium (1 L) was supplemented where appropriate with L-phenylalanine (0.040 g), L-tyrosine (0.040 g), L-tryptophan (0.040 g), *p*-hydroxybenzoic acid (0.010 g), potassium *p*-aminobenzoate (0.010 g), and 2,3-dihydroxybenzoic acid (0.010 g). L-Serine was added to a final concentration of 40 mg/L where indicated. Antibiotics were added where appropriate to the following final concentrations: ampicillin (Ap), 50 µg/mL; chloramphenicol (Cm), 20 µg/mL; kanamycin (Kan), 50 µg/mL; tetracycline (Tc),

12.5 μ g/mL. Stock solutions of antibiotics were prepared in water with the exceptions of chloramphenicol which was prepared in 95% ethanol and tetracycline which was prepared in 50% aqueous ethanol. Aqueous stock solutions of isopropyl- β -D-thiogalactopyranoside (IPTG) were prepared at various concentrations. Solutions of LB medium, M9 inorganic salts, MgSO₄, D-glucose, D-xylose, glycerol and erythritol were autoclaved individually and then mixed. Solutions of potassium D-xyloate, thiamine hydrochloride, antibiotics, and IPTG were sterilized through 0.22- μ m membranes. Other solid media were prepared by addition of Difco agar to a final concentration of 1.5% (w/v) to the liquid medium.

Pseudomonas fragi and *Pseudomonas putida* were grown either in nutrient broth or on solid nutrient agar plates. Nutrient broth and nutrient agar were prepared according to procedures recommended by the manufacturers. *Klebsiella pneumoniae* was cultured in liquid or on solid ATCC medium 561. ATCC medium 561 (800 mL) contained yeast extract (1 g), casein hydrolysate (1.4 g), K₂HPO₄ (0.6 g), MgSO₄ (0.5 g), K₂SO₄ (1 g), and glycerol (20 mL). Solid ATCC medium 561 was prepared by addition of Difco agar to a final concentration of 1.5% (w/v) to the liquid medium.

The standard fermentation medium (1 L) utilized in Chapter 3 contained K₂HPO₄ (7.5 g), ammonium iron (III) citrate (0.3 g), citric acid monohydrate (2.1 g), and concentrated H₂SO₄ (1.2 mL). Fermentation medium was adjusted to pH 7.0 by addition of concentrated NH₄OH before autoclaving. The following supplements were added immediately prior to initiation of the fermentation: D-glucose, MgSO₄ (0.24 g), potassium and trace minerals including (NH₄)₆(Mo₇O₂₄)·4H₂O (0.0037 g), ZnSO₄·7H₂O (0.0029 g), H₃BO₃ (0.0247 g), CuSO₄·5H₂O (0.0025 g), and MnCl₂·4H₂O (0.0158 g). IPTG stock solution was added as necessary to the indicated final concentration. Glucose feed

solution and MgSO_4 (1 M) solution were autoclaved separately. Glucose feed solution (650 g/L) was prepared by combining 300 g of glucose and 280 mL of H_2O . Solutions of trace minerals and IPTG were sterilized through 0.22- μm membranes. Antifoam (Sigma 204) was added to the fermentation broth as needed.

General fed-batch fermentation conditions

Fermentations employed a 2.0 L working capacity B. Braun M2 culture vessel. Utilities were supplied by a B. Braun Biostat MD controlled by a DCU-3. Data acquisition utilized a Dell Optiplex Gs⁺ 5166M personal computer (PC) equipped with B. Braun MFCS/Win software (v1.1) or a Dell Optiplex GX200 personal computer (PC) equipped with B. Braun MFCS/Win software (v2.0). Temperature, pH, and carbon source feeding were controlled with PID control loops. pH was maintained at 7.0 by addition of concentrated NH_4OH or 2 N H_2SO_4 . Dissolved oxygen (D.O.) was measured using a Mettler-Toledo 12 mm sterilizable O_2 sensor fitted with an Ingold A-type O_2 permeable membrane. Inoculants were started by introduction of a single colony picked from an agar plate into 5 mL of M9 medium. Cultures were grown at 37 °C with agitation at 250 rpm until they were turbid and subsequently transferred to 100 mL of M9 medium. Cultures were grown at 37 °C and 250 rpm for an additional 10 h. The inoculant ($\text{OD}_{600} = 1.0\text{-}3.0$) was then transferred into the fermentation vessel and the batch fermentation was initiated ($t = 0$ h).

Three staged methods were used to maintain D.O. concentrations at desired air saturation during the fermentations. With the airflow at an initial setting of 0.06 L/L/min,

the D.O. concentration was maintained by increasing the impeller speed from its initial set point of 50 rpm to its preset maximum rate. With the impeller speed constant, the mass flow controller then maintained the D.O. concentration by increasing the airflow rate from 0.06 L/L/min to a preset maximum of 1.0 L/L/min. At constant impeller speed and constant airflow rate, the D.O. concentration was finally maintained at the desired air saturation for the remainder of the fermentation by oxygen sensor-controlled carbon source feeding. At the beginning of this stage, the D.O. concentration fell below the desired air saturation due to residual initial carbon source in the medium. This lasted for approximately 10 min to 30 min before carbon source feeding commenced. The carbon source feed PID control parameters were set to 0.0 s (off) for the derivative control (τ_D) and 999.9 s (minimum control action) for the integral control (τ_I). X_P was set to 950% to achieve a K_c of 0.1.

Analysis of fermentation broth

Samples (5-10 mL) of fermentation broth were removed at the indicated timed intervals. Cell densities were determined by dilution of fermentation broth with water (1:100) followed by measurement of absorption at 600 nm (OD_{600}). Dry cell weight of *E. coli* cells (g/L) was calculated using a conversion coefficient of 0.43 g/L/ OD_{600} . The remaining fermentation broth was centrifuged to obtain cell-free broth.

Glucose concentrations in cell-free broth were measured using the Glucose Diagnostic Kit purchased from Sigma. For the biosynthesis of D-1,2,4-butanetriol, organic acid byproducts such as D-xylonic acid, 3-deoxy-D-*glycero*-pentulosonic acid, 3-

ddeoxy-D-*glycero*-pentanoic acid, and 4,5-dihydroxy-*threo*-L-norvaline concentrations in the cell-free broth were quantified by ^1H NMR. A portion (0.5-2.0 mL) of the cell-free broth was concentrated to dryness under reduced pressure, concentrated to dryness one additional time from D_2O , and then redissolved in D_2O containing a known concentration of the sodium salt of 3-(trimethylsilyl)propionic-2,2,3,3- d_4 acid (TSP, Lancaster Synthesis Inc.). Concentrations were determined by comparison of integrals corresponding to each compound with the integral corresponding to TSP ($\delta = 0.00$ ppm) and were converted by response factors determined using authentic materials. Compounds were quantified using the following resonances: D-xylonic acid (δ 4.08, d, 1 H); 3-deoxy-D-*glycero*-pentulosonic acid (δ 4.58, m, 1 H); 3-deoxy-D-*glycero*-pentonic acid (δ 1.94, ddd, 1 H); 4,5-dihydroxy-*threo*-L-norvaline (δ 4.01, dd, 1 H). A standard concentration curve was determined for metabolites using solutions of authentic samples. The concentration of D-1,2,4-butanetriol and 3,4-dihydroxy-D-butanoic acid in cell-free broth was quantified by GC analysis. A portion of the fermentation broth (0.5-1.0 mL) was concentrated to dryness under reduced pressure, and the residue was redissolved in pyridine (0.9 mL). To this pyridine solution, dodecane (0.1 mL) and bis(trimethylsilyl)trifluoroacetamide (BSTFA, 2 mL, 7.53 mmol) were sequentially added. Silylation of D-1,2,4-butanetriol and 3,4-dihydroxy-D-butanoic acid were carried out at room temperature with stirring for 10 h. Samples were then analyzed using gas chromatography.

Genetic manipulations

Recombinant DNA manipulations generally followed methods described by Sambrook et al.¹³ Restriction enzymes were purchased from Invitrogen or New England

Biolabs. T4 DNA ligase was obtained from Invitrogen. Fast-Link™ DNA Ligation Kit was obtained from Epicentre. Zymoclean Gel DNA Recovery Kit and DNA Clean & Concentrator Kit was obtained from Zymo Research Company. Maxi and Midi Plasmid Purification Kits were obtained from Qiagen. Calf intestinal alkaline phosphatase was obtained from Boehringer Mannheim. Agarose (electrophoresis grade) was obtained from Invitrogen. Phenol was prepared by addition of 0.1 % (w/v) 8-hydroxyquinoline to distilled, liquefied phenol. Extraction with an equal volume of 1 M Tris-HCl (pH 8.0) two times was followed by extraction with 0.1 M Tris-HCl (pH 8.0) until the pH of the aqueous layer was greater than 7.6. Phenol was stored at 4 °C under an equal volume of 0.1 M Tris-HCl (pH 8.0). SEVAG was a mixture of chloroform and isoamyl alcohol (24:1, v/v). TE buffer contained 10 mM Tris-HCl (pH 8.0) and 1 mM Na₂EDTA (pH 8.0). TAE buffer contained 40 mM Tris-acetate (pH 8.0) and 2 mM Na₂EDTA. Endostop solution (10× concentration) contained 50% glycerol (v/v), 0.1 M Na₂EDTA, pH 7.5, 1% sodium dodecyl sulfate (SDS) (w/v), 0.1% bromophenol blue (w/v), and 0.1% xylene cyanole FF (w/v) and was stored at 4 °C. Prior to use, 0.12 mL of DNase-free RNase was added to 1 mL of 10X Endostop solution. DNase-free RNase (10 mg mL⁻¹) was prepared by dissolving RNase in 10 mM Tris-HCl (pH 7.5) and 15 mM NaCl. DNase activity was inactivated by heating the solution at 100 °C for 15 min. Aliquots were stored at -20 °C. PCR amplifications were carried out as described by Sambrook et al.¹³ A standard reaction (0.1 mL) contained 10 mM KCl, 20 mM Tris-HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.1% Triton X-100, dATP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), dTTP (0.2 mM), template DNA, 0.5 μM of each primer, and 2 units of *Taq* polymerase. Template concentration varied from 0.02 μg to 1.0 μg.

Large scale purification of plasmid DNA

Plasmid DNA was purified on a large scale using a modified alkaline lysis method described by Sambrook et al.¹³ In a 2 L Erlenmeyer flask, 500 mL of LB containing the appropriate antibiotics was inoculated from a single colony, and the culture was incubated in a gyratory shaker at 37 °C for 14 h with agitation at 250 rpm. Cells were harvested by centrifugation (4,000g, 5 min, 4 °C) and then resuspended in 10 mL of cold GETL solution (50 mM glucose, 20 mM Tris-HCl (pH 8.0), 10 mM Na₂EDTA, pH 8.0) into which lysozyme (5 mg/mL) had been added immediately before use. The suspension was stored at room temperature for 5 min. Addition of 20 mL of 1% sodium dodecyl sulfate (w/v) in 0.2 N NaOH was followed by gentle mixing and storage on ice for 15 min. Fifteen milliliters of an ice-cold solution containing 3 M KOAc (prepared by combining 60 mL of 5 M potassium acetate, 11.5 mL of glacial acetic acid, and 28.5 mL of H₂O) was added. Vigorous shaking resulted in formation of a white precipitate. After the suspension was stored on ice for 10 min, the cellular debris was removed by centrifugation (48,000g, 20 min, 4 °C). The supernatant was transferred to two clean centrifuge bottles and isopropanol (0.6 volumes) was added to precipitate the DNA. After the samples were left at room temperature for 15 min, the DNA was recovered by centrifugation (20,000g, 20 min, 4 °C). The DNA pellet was then rinsed with 70% ethanol and dried.

Further purification of the DNA sample involved precipitation with polyethylene glycol (PEG). The isolated DNA was dissolved in TE (3 mL) and transferred to a Corex tube. Cold 5 M LiCl (3 mL) was added and the solution was gently mixed. The sample was then centrifuged (12,000g, 10 min, 4 °C) to remove high molecular weight RNA. The

clear supernatant was transferred to a clean Corex tube and isopropanol (6 mL) was added followed by gentle mixing. The precipitated DNA was collected by centrifugation (12,000g, 10 min, 4 °C). The DNA was then rinsed with 70% ethanol and dried. After re-dissolving the DNA in 0.5 mL of TE containing 20 µg/mL of RNase, the solution was transferred to a 1.5 mL microcentrifuge tube and stored at room temperature for 30 min. DNA was precipitated from solution upon addition of 500 µL of 1.6 M NaCl containing 13% PEG-8000 (w/v) (Sigma). The solution was mixed and centrifuged (microcentrifuge, 10 min, 4 °C) to recover the precipitated DNA. The supernatant was removed, and the DNA was then re-dissolved in 400 µL of TE. The sample was extracted sequentially with phenol (400 µL), phenol and SEVAG (400 µL each), and finally SEVAG (400 µL). Ammonium acetate (10 M, 100 µL) was added to the aqueous DNA solution. After thorough mixing, 95% ethanol (1 mL) was added to precipitate the DNA. The sample was left at room temperature for 5 min and then centrifuged (microcentrifuge, 5 min, 4 °C). The DNA was rinsed with 70% ethanol, dried, and then redissolved in 200-500 µL of TE.

Alternatively, DNA was purified using a Qiagen Maxi Kit or Midi Kit as described by the manufacturer. The purity of DNA isolated by these kits was adequate for DNA sequencing.

Small scale purification of plasmid DNA

An overnight culture (5 mL) of the plasmid-containing strain was grown in LB containing the appropriate antibiotics. Cells from 3 mL of the culture were collected in a 1.5 mL microcentrifuge tube by centrifugation. The resulting cell pellet was liquefied by vortexing (30 sec) and then resuspended in 0.1 mL of cold GETL solution into which

lysozyme (5 mg/mL) had been added immediately before use. The solution was stored on ice for 10 min. Addition of 0.2 mL of 1% sodium dodecyl sulfate (w/v) in 0.2 N NaOH was followed by gentle mixing and storage on ice for 5-10 min. To the sample was added 0.15 mL of cold KOAc solution. The solution was shaken vigorously and stored on ice for 5 min before centrifugation (15 min, 4 °C). The supernatant was transferred to another microcentrifuge tube and extracted with equal volumes of phenol and SEVAG (0.2 mL). The aqueous phase (approximately 0.5 mL) was transferred to a fresh microfuge tube, and DNA was precipitated by the addition of 95% ethanol (1 mL). The sample was left at room temperature for 5 min before centrifugation (15 min, room temperature) to collect the DNA. The DNA pellet was rinsed with 70% ethanol, dried, and redissolved in 50 to 100 μ L TE. DNA isolated using this method was used for restriction enzyme analysis, although the concentration of DNA could not be accurately determined by spectroscopic methods.

Restriction enzyme digestion of DNA

Restriction enzyme digests were performed in buffers provided by Invitrogen or New England Biolabs. A typical restriction enzyme digest contained 0.8 μ g of DNA in 8 μ L of TE, 2 μ L of restriction enzyme buffer (10 \times concentration), 1 μ L of bovine serum albumin (0.1 mg/mL), 1 μ L of restriction enzyme and 8 μ L TE. Reactions were incubated at 37 °C for 1 h, terminated by addition of 2.2 μ L of 10X Endostop solution and analyzed by agarose gel electrophoresis. When DNA was required for cloning experiments, the digest was terminated by addition of 1 μ L of 0.5 M Na₂EDTA (pH 8.0) or by heating at

70 °C for 15 min followed by extraction of the DNA using Zymoclean gel DNA recovery kit.

Determination of DNA concentration

The concentration of DNA in the sample was determined as follows. An aliquot (10 μ L) of DNA was diluted to 1 mL in TE and the absorbance at 260 nm was measured relative to the absorbance of TE. The DNA concentration was calculated based on the fact that the absorbance at 260 nm of 50 μ g/mL of double stranded DNA is 1.0.

Agarose gel electrophoresis

Agarose gel typically contained 0.7% agarose (w/v) in TAE buffer. Ethidium bromide (0.5 μ g/ml) was added to the agarose to allow visualization of DNA fragments under a UV lamp. Agarose gel was run in TAE buffer. The size of the DNA fragments were determined using two sets of DNA molecular weight standards: λ DNA digested with *Hind*III (23.1-kb, 9.4-kb, 6.6-kb, 4.4-kb, 2.3-kb, 2.0-kb and 0.6-kb) and λ DNA digested with *Eco*RI and *Hind*III (21.2-kb, 5.1-kb, 5.0-kb, 4.3-kb, 3.5-kb, 2.0-kb, 1.9-kb, 1.6-kb, 1.4-kb, 0.9-kb, 0.8-kb and 0.6-kb).

Isolation of DNA from agarose

The band of agarose containing DNA of interest was excised from the gel while visualized with long wavelength UV light. Two methods were used for isolating DNA from agarose gels. The first method used Zymoclean gel DNA recovery kit to isolate

DNA from the agarose gel according to the procedure provided by Zymo Research. Alternatively, the agarose gel containing DNA was chopped thoroughly with a razor and then transferred to a 0.5 mL microfuge tube packed tightly with glass wool and having an 18 gauge hole at the bottom. The tube was centrifuged for 5 min using a Beckman microfuge to extrude the DNA solution from the agarose into a second 1.5 mL microfuge tube. The DNA was precipitated using 3 M NaOAc and 95% ethanol as described previously and subsequently redissolved in TE.

Treatment of DNA with Klenow fragment

DNA fragments with recessed 3' termini were modified to DNA fragments with blunt ends by treatment with the Klenow fragment of *E. coli* DNA polymerase I. After restriction digestion (20 μ L) of the DNA (0.8-2 μ g) was complete, a solution (1 μ L) containing each of the four dNTPs was added to a final concentration of 1 mM for each dNTP. Addition of 1-2 units of Klenow fragment was followed by incubation at room temperature for 20-30 min. Since the Klenow fragment works well in the common restriction enzyme buffers, there was generally no need to purify the DNA after restriction digestion and prior to filling recessed 3' termini. Klenow reactions were quenched by extraction with equal volumes of phenol and SEVAG. DNA was recovered using the Zymoclean gel DNA recovery kit or by precipitation as described previously and subsequently dissolved in TE.

Treatment of vector DNA with calf intestinal alkaline phosphatase

Following restriction enzyme digestion, plasmid vectors were dephosphorylated to prevent self-ligation. Digested vector DNA was dissolved in TE (88 μL). To this sample was added 10 μL of dephosphorylation buffer (10 \times concentration) and 2 μL of calf intestinal alkaline phosphatase (2 units). The reaction was incubated at 37 °C for 1-2 h. The phosphatase was inactivated by the addition of 1 μL of 0.5 M Na_2EDTA (pH 8.0) followed by heat treatment (70 °C, 15 min). The sample was extracted with phenol and SEVAG (100 μL each) to remove the protein, and the DNA was purified as previously described and subsequently dissolved in TE.

Ligation of DNA

Molar ratios of insert to vector were typically maintained at 3 to 1 for DNA ligations. A typical reaction contained 0.1 μg of vector DNA and 0.05 to 2.0 μg of insert DNA in a total volume of 7 μL . To this was added 2 μL of T4 ligation buffer (5 \times concentrations) and 1 μL of T4 DNA ligase (2 units). The reaction was incubated at 16 °C for at least 4 h and then used to transform competent cells.

Alternatively, A Fast-Link™ DNA Ligation Kit (Epicentre, Madison, WI) was used for ligation of insert DNA with cohesive or blunt ends into predigested vectors with compatible ends according to the protocol provided by Epicentre.

Preparation and transformation of competent *E. coli* cells

Competent cells were prepared according to a procedure modified from Sambrook et al.¹³ LB medium (5 mL) containing antibiotics where appropriate, was inoculated with a single colony from a LB plate containing antibiotics where appropriate. The culture was grown at 37 °C with shaking at 250 rpm for 10-12 h. An aliquot (1 mL) from the culture (5 mL) was used to inoculate LB (100 mL) containing the appropriate antibiotics. The culture was grown at 37 °C with shaking at 250 rpm in a NBS series 25 incubator shaker until the optical density at 600 nm was between 0.4 and 0.6. The culture was transferred to a centrifuge bottle that had been sterilized with a 25 % (v/v) bleach solution and rinsed four times with sterile, deionized water. The cells were harvested by centrifugation (4000 g, 5 min, 4 °C) and the culture medium was decanted. All subsequent manipulations were carried out on ice. The harvested cells were resuspended in ice-cold 0.9 % NaCl (100 mL), and the cells were collected by centrifugation (4,000 g, 5 min, 4 °C). The 0.9 % NaCl solution was decanted, the cells were resuspended in ice-cold 100 mM CaCl₂ (50 mL) and stored on ice for 30 min. After centrifugation (4,000 g, 5 min, 4 °C), the cells were resuspended in 4 mL of ice-cold 100 mM CaCl₂ containing 15% glycerol (v/v). Aliquots (0.25 mL) of competent cells were added to 1.5 mL microfuge tubes, immediately frozen in liquid nitrogen, and stored at -78 °C.

Frozen competent cells were thawed on ice for 5 min before transformation. A small aliquot (1 to 10 μ L) of plasmid DNA or a ligation reaction was added to the thawed competent cells (0.1 mL). The solution was gently mixed by tapping and stored on ice for 30 min. The cells were then heat shocked at 42 °C for 30 seconds and returned to ice briefly (1 min). LB (0.5 mL, no antibiotics) was added to the cells, and the sample was

incubated at 37 °C (no agitation) for 1 h. Cells were collected by centrifugation (30 s) in a microcentrifuge. If the transformation was to be plated onto LB plates, 0.5 mL of the culture supernatant was removed, and the cells were resuspended in the remaining 0.1 mL of LB and subsequently spread onto plates containing the appropriate antibiotics. If the transformation was to be plated onto minimal medium plates, the cells were washed twice with a solution of M9 salts (0.5 mL). After resuspension in a fresh aliquot of M9 salts (0.1 mL), the cells were spread onto a plate. An aliquot of competent cells with no DNA added was also carried through the transformation protocol as a control. These cells were used to check the viability of the competent cells and to verify the absence of growth on selective medium.

Transformations were also performed by electroporation using electrocompetent cells. An aliquot (1 mL) from an overnight culture (5 mL) was used to inoculate 500 mL of 2×YT containing the appropriate antibiotics. The cells were cultured at 37 °C with shaking at 250 rpm. Once an absorbance of 0.6-0.8 at 600 nm was observed, the cells were kept on ice for 10 min and harvested (3,000g, 5 min, 4 °C). The cells were gently washed three times with sterile, cold water (450 mL once and 250 mL twice) and then resuspended in 100 mL sterile, ice-cold aqueous 10% glycerol (v/v). After centrifugation (3,000g, 5 min, 4 °C), the cells were resuspended in 1.5 mL sterile ice-cold aqueous 10% glycerol (v/v). Aliquots (0.1 mL) of electrocompetent cells were dispensed into 1.5 mL microfuge tubes, and immediately frozen in liquid nitrogen and stored at -78 °C.

The electroporation was performed in Bio-Rad Gene Pulser cuvettes with an electrode gap of 0.2 cm. The cuvettes were chilled on ice for 5 min prior to use. Electrocompetent cells were thawed in ice for 5 min, and 40 µL of thawed cells was added

to the chilled cuvette. To this was added 1-10 μL of plasmid DNA ($1\text{ }\mu\text{g mL}^{-1}$), and the mixture was gently shaken. The Bio-Rad Gene Pulser was set at 2.5 kV, 25 μF and 200 Ω . The outside surface of the cuvette was wiped clean and it was placed in the sample chamber. A single pulse was applied, the cuvette was removed, and 1 mL of freshly prepared SOC was added into it. The contents of the cuvette were transferred to a 15 mL sterile centrifugation tube. The cells were incubated at 37 °C for 1 h with shaking at 250 rpm. The transformed cells were plated in the same manner as in the transformation with chemically competent cells.

Purification of *E. coli* genomic DNA

Genomic DNA was purified using a method modified from Silhavy et al.¹⁴ A single colony of an *E. coli* strain was inoculated into 100 mL of TB medium (500 mL Erlenmeyer flask). The cells were cultured in a gyratory shaker (37 °C, 250 rpm) for 12 h. Centrifugation (4,000g, 5 min, 4 °C) of the culture was followed by resuspension of the cell pellet in 5 mL of buffer (50 mM Tris-HCl, 50 mM Na₂EDTA, pH 8.0) and storage at -20 °C for 20 min to freeze the suspension. To the frozen cells was added 0.5 mL of 0.25 M Tris-HCl (pH 8.0) that contained 5 mg of lysozyme. The suspension was thawed at room temperature in a water bath with gentle mixing and then stored on ice for 45 min. The sample was then transferred to a Corex tube. After addition of 1 mL of STEP solution (25 mM Tris-HCl (pH 7.4), 200 mM Na₂EDTA (pH 8.0), 0.5% SDS (w/v), and proteinase K (1 mg mL^{-1}), prepared just before use), the mixture was incubated at 50 °C for at least 1 h with gentle, periodic mixing. The solution was then divided into two Corex tubes, and the contents of each tube were extracted with phenol (4 mL). The organic and

aqueous layers were separated by centrifugation (1,000g, 15 min, room temperature), and the aqueous layer was transferred to a fresh Corex tube. All transfers of the aqueous layer were carried out using wide bore pipette tips to minimize shearing of the genomic DNA. The contents of each tube were extracted again with a mixture of phenol (3 mL) and SEVAG (3 mL). Extractions with phenol/SEVAG were repeated (approximately 6 times) until the aqueous layer was clear.

Genomic DNA was precipitated by addition of 0.1 volume of 3 M NaOAc (pH 5.2) followed by gentle mixing and addition of 2 volumes of 95% ethanol. Threads of DNA were spooled onto a sealed Pasteur pipette and transferred to a Corex tube that contained 5 mL of 50 mM Tris-HCl (pH 7.5), 1 mM Na₂EDTA (pH 8.0), and 1 mg of RNase. The mixture was stored at 4 °C overnight to allow the DNA to dissolve completely. The solution was then extracted with SEVAG (5 mL) and centrifuged (1,000 g, 15 min, room temperature). The aqueous layer was transferred to a fresh Corex tube and the genomic DNA was precipitated as described above. The threads of DNA were spooled onto a Pasteur pipette and redissolved in 2 mL of 50 mM Tris-HCl (pH 7.5) and 1 mM Na₂EDTA (pH 8.0). Genomic DNA was stored at 4 °C.

Alternatively, genomic DNA was purified using a method described by Pitcher et al.¹⁵ A single colony of an *E. coli* strain was inoculated into 20 mL of LB medium. The cells were culture in a gyratory shaker (37 °C, 250 rpm) overnight, and were subsequently harvested by centrifugation (1,000g, 15 min, room temperature). The cell pellet obtained was resuspended into 100 μ L of TE and incubated at 37°C for 30 min. Cells were lysed with 0.5 mL 5 M guanidine thiocyanate, 100 mM EDTA and 0.5% (v/v) sarkosyl (GES reagent), which was prepared as follows. Guanidine thiocyanate (60 g), 0.5 M Na₂EDTA,

pH 8.0 (20 mL), and deionized water (20 mL) were heated at 65°C with mixing until dissolved. After cooling, 5 mL of 10% v/v sarkosyl were added, the solution was made up to 100 mL with deionized water, filtered through a 0.22- μ m membrane and stored at room temperature.

The cell suspension was vortexed briefly and checked for lysis (clear solution) after 5-10 min. The lysate was cooled on ice, and a cold solution of ammonium acetate (7.5 M, 0.25 mL) was added with mixing on ice for 10 min. To this sample, 0.5 mL SEVAG was added and mixed thoroughly. After centrifugation in a 1.5 mL Eppendorf tube (25,000g, 10 min, room temperature), the supernatant was transferred to an Eppendorf tube and 0.54 volumes of cold 2-propanol was added. The tubes were inverted for 1 min to mix the solutions and the fibrous DNA precipitate was collected by centrifugation (6,500g, 20 sec, room temperature). DNA pellet was washed five times with 70% ethanol and dried in air at room temperature for 20 min. Finally, the genomic DNA was redissolved in 100 μ L TE.

P1 phage-mediated transduction

Transduction with P1 phage was carried out using a method modified from Miller.¹⁶ P1 phage lysate was prepared by propagation of phage in the donor strain using the following procedure. Serial dilutions of P1 phage stock (0.1 mL, 10^{-1} to 10^{-5}) in LB were prepared in sterile test tubes (13 \times 100 mm). An aliquot (0.1 mL, approximately 5×10^8 cells) of an overnight culture of the donor strain was added to each tube. Sterile, molten soft agar (45 °C) was added to each tube. The contents of each tube were mixed and poured immediately onto a pre-warmed (37 °C) L plate and swirled gently to achieve

uniform coverage of the plate. After the agar had solidified, the plates were incubated at 37 °C until confluent lysis had occurred (approximately 8 h). Because the multiplicity of infection is critical to phage generation, confluent lysis occurred on only one or two of the plates. To the plates displaying confluent lysis, L-broth (4 mL) was added, and the plates were then stored overnight at 4 °C to allow the phage particles to diffuse into the broth. The L-broth was collected from the plate and vortexed with several milliliters of CHCl_3 . The solution was centrifuged (2,000g, 5 min, room temperature) to separate the layers. Aqueous phage lysate was stored in 1.5 mL microfuge tubes over several drops of CHCl_3 at 4 °C.

Infection of the recipient strain with phage lysate proceeded as follows. An overnight culture (2 mL) of the recipient strain was centrifuged (microfuge, 30 sec, 4 °C) and the growth medium was discarded. The cells were resuspended in 1 mL sterile solution containing of 5 mM CaCl_2 and 100 mM MgSO_4 , and shaken (250 rpm) at 37 °C for 15 min to promote aeration of the cells. In the meantime, 0.1 mL serial dilutions (10^0 to 10^{-3}) of phage lysate in LB were prepared in sterile microfuge tubes. An aliquot (0.1 mL) of aerated recipient cells was added to each of the phage dilutions, the samples were gently mixed, and then incubated at 37 °C for 20 min without shaking. Sodium citrate (1 M, 0.2 mL) was added to each sample, and the cells were harvested (microfuge, 30 s, room temperature) and resuspended in 0.2 mL of LB containing 100 mM sodium citrate. After incubation at 30 °C for 30 min, cells were again harvested (microfuge, 30 sec, room temperature), resuspended in 0.1 mL of growth medium, and plated out onto appropriate agar plates.

Enzyme assays

After collected and resuspended in the proper resuspension buffer, the cells were disrupted by two passages through a French pressure cell (SLM Aminco) at 16,000 psi. Cellular debris was removed from the lysate by centrifugation (48,000g, 20 min, 4 °C). Protein was quantified using the Bradford dye-binding procedure.¹⁷ A standard curve was prepared using bovine serum albumin. Protein assay solution was purchased from Bio-Rad and used as described by the manufacture.

2-Deoxyribose-5-phosphate aldolase assay

2-Deoxyribose 5-phosphate aldolase activity was assayed as described by Wong, C.-H. in literature.¹⁸ 2-Deoxyribose-5-phosphate (1 mM), 0.28 mM NADH, and a mixture of glycerophosphate dehydrogenase and triose phosphate isomerase was incubated in 50 mM Tris, 0.1 mM EDTA, pH 7.6 at 25 °C. The assay was initiated by the addition of 2-deoxyribose 5-phosphate aldolase, and the decrease in the absorbance at 340 nm was monitored. The extinction coefficient for NADH was taken as $6.22 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$.

Glycerol and diol dehydratase assay

Glycerol and diol dehydratase was assayed according to the procedure described by Abeles.¹⁹ A coupled enzyme reaction in which the aldehyde formed by the dehydratase reaction was reduced to the corresponding alcohol. The assay mixture was composed of an appropriate amount of dehydratase, 0.2 M substrate, 12 U yeast alcohol dehydrogenase (Sigma-Aldrich), 0.2 mM NADH, 0.04 M potassium phosphate buffer (pH 8.0), and 10

μ M coenzyme B₁₂ in a total volume of 1.0 mL. The reaction was carried out at 37 °C in a cuvette with a 1.0 cm light path and started by adding coenzyme to the reaction mixture which was preincubated at 37 °C for 5 min. One unit is defined as the amount of enzyme activity catalyzing the formation of 1 μ mol of aldehyde per min under the standard conditions.

2-Keto acid decarboxylase assay

The specific activities of 2-keto acid decarboxylases including pyruvate decarboxylase and benzoylformate decarboxylase were determined by coupling the decarboxylation reaction with aldehyde-dependent oxidation of NADH by equine liver alcohol dehydrogenase. Resuspension buffer contained sodium phosphate (50 mM, pH 6.5) and MgCl₂ (10 mM). The enzyme assay solution (1 mL) contained sodium phosphate (50 mM, pH 6.5), MgCl₂ (10 mM), thiamine pyrophosphate (0.15 mM), NADH (0.2 mM), and an appropriate amount of cell lysate.^{20,21} When pyruvate decarboxylase was assayed for specific activity on pyruvate, equine liver alcohol dehydrogenase (0.05 U) and pyruvate (5 mM) were also included in the assay.²⁰ When benzoylformate decarboxylase was assayed for specific activity on benzoylformate, equine liver alcohol dehydrogenase (0.05 U) and benzoylformate (2 mM) were also included in the assay.²¹ When 2-keto decarboxylases were assayed for specific activity on 3-deoxy-D,L-*glycero*-pentulosonic acid, equine liver alcohol dehydrogenase (1 U) and 3-deoxy-D,L-*glycero*-pentulosonate (50 mM) were also included in the assays. Enzyme activity was measured spectrophotometrically by monitoring the oxidation of NADH at 340 nm. One unit of 2-

keto acid decarboxylase was defined as the conversion of 1 μmol of NADH ($\epsilon = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$) per min at room temperature.

Alcohol dehydrogenase assay

The enzyme activities of alcohol dehydrogenases including *adhA*- and *adhB*-encoded alcohol dehydrogenases²² of *Zymomonas mobilis* and 1,3-propanediol oxidoreductase²³ were measured in the oxidative direction by monitoring the formation of NADH. Resuspension buffer for *adhA*- and *adhB*-encoded alcohol dehydrogenase contained potassium phosphate (30 mM, pH 6.5), sodium ascorbate (10 mM), and $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ (0.5 mM).²² Enzyme assays (1 mL) of AdhA and AdhB on the native substrate contained Tris-HCl (30 mM, pH 8.5), NAD (1 mM), appropriate amount of cell lysate, and ethanol (1 mM).²² *E. coli* alcohol dehydrogenase candidates were assayed under the same conditions as AdhA and AdhB. Cell resuspension buffer for 1,3-propanediol oxidoreductase contained potassium carbonate (100 mM, pH 9.0) and ammonium sulfate (35 mM).²³ When the specific activity of 1,3-propanediol oxidoreductase on native substrate was measured, the enzyme assay solution (1 mL) contained potassium carbonate (100 mM, pH 9.0), NAD (1 mM), appropriate amount of cell lysate, and 1,3-propanediol (0.1 mM).²³ When the above alcohol dehydrogenases were assayed for specific activity on nonnative substrate, 1,2,4-butanetriol (50 mM) was included in the assay in place of the native substrate. Enzyme activity was measured spectrophotometrically by monitoring the formation of NADH at 340 nm. One unit of

alcohol dehydrogenase was defined as the formation of 1 μmol of NADH ($\epsilon = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$) per min at room temperature.

Protein SDS-PAGE analysis

Protein SDS-PAGE analysis followed the procedure described by Harris.²⁴ Preparation of a 10% separating gel started from mixing 3.33 mL of 30% (w/v) aqueous acrylamide stock solution containing N,N'-methylene-bisacrylamide (0.8% (w/v)), 2.5 mL of 1.5 M Tris-HCl (pH 8.8), and 4 mL of distilled deionized water. After degassing the solution using a water aspirator for 30 min, 0.1 mL of 10% (w/v) aqueous ammonium persulfate solution, 0.1 mL 10% (w/v) aqueous SDS solution, and 0.005 mL of N, N, N', N'-tetramethylethylenediamine (TEMED) were added. The solution was mixed thoroughly and poured into a 0.1 cm-width gel cassette to about 1.5 cm below the top of the gel cassette. *t*-Amyl alcohol was overlaid on top of the solution and the gel was allowed to polymerize for 1 h at rt. The stacking gel was prepared by mixing 1.7 mL 30% acrylamide stock solution containing N,N'-methylene-bisacrylamide (0.8% (w/v)), 2.5 mL Tris-HCl solution (0.5 M, pH 6.8), and 5.55 mL of distilled deionized water. After degassing for 30 min, 0.1 mL of 10% ammonium persulfate, 0.1 mL 10% SDS, and 0.01 mL of TEMED was added, and the solution was mixed thoroughly. *t*-Amyl alcohol was removed from the top of the gel cassette, which was subsequently rinsed with water and wiped dry. After insertion of the comb, the gel cassette was filled with stacking gel solution, and the stacking gel was allowed to polymerize for 1 h at rt. After removal of the comb, the gel cassette was installed into the electrophoresis apparatus. The electrode chamber was then filled with electrophoresis buffer containing glycine (192 mM), Tris

base (25 mM), and 0.1% SDS (w/v). Following dilution with Laemmli sample buffer (10 μ L, Sigma S-3401) consisting of 4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue, and Tris-HCl (125 mM, pH 6.8), each protein sample (10 μ L) was heated at 100 °C for 10 min. Samples and markers (MW-SDS-200, Sigma) were then loaded into the sample wells and the gel was run under constant current at 30 mA until the blue tracking dye (bromophenol blue) reached the interface of stacking gel and separating gel. The protein gel was then run at a higher current (50 mA). When the blue tracking dye reaches the bottom of the gel, electrophoresis was terminated. The protein gel was subsequently removed from the cassette and submerged in 10% (w/v) aqueous trichloroacetic acid solution with constant shaking for 30 min. The protein gel was then transferred into a solution containing 0.1% (w/v) Coomassie Brilliant Blue R, 45% (v/v) MeOH, 10% (v/v) HOAc in H₂O and stained with constant shaking for 4 h. Destaining of the protein gel was carried out in a solution containing 45% (v/v) MeOH, 10% (v/v) HOAc in H₂O for 2-3 h. For long-term storage, SDS-PAGE gels were sealed in plastic bags containing 10% glycerol.

Chapter 2

Purification of 2-deoxyribose-5-phosphate aldolase.

E. coli strain DH5 α /pVH17 was obtained from the American Type Culture Collection (ATCC 86963). Cells were grown in rich medium (1 L) containing 30 g tryptone, 20 g yeast extract, and 10 g 4-morpholine-propanesulfonic acid at pH 7. Cultures were initiated by inoculating a single colony into 100 mL of rich medium

containing 50 µg/mL ampicillin. Inoculants were grown at 37 °C with agitation at 250 rpm for 12 h, and a 10 mL aliquot was subsequently transferred to 1 L of rich growth medium. A total of 4-1 L cultures of inoculated rich media were then cultured for 12 h at 37 °C. Following centrifugation at 4,000g for 5 min, approximately 50 g of wet cells were obtained. All subsequent manipulations were performed at 4 °C. Cells were resuspended in 100 mL 100 mM Tris, 2 mM EDTA, pH 7.6 and lysed by two passages through a French press (1220 atm). Cellular debris was removed by centrifugation at 20,000g for 30 min to afford cell-free lysate containing 38,000 units of aldolase activity. The crude lysate was applied to Q-Sepharose Fast-Flow anion exchange resin (300 mL), which had been equilibrated with 100 mM Tris and 2 mM EDTA, pH 7.6. The column was eluted with a linear gradient (500 mL + 500 mL, 0-1 M) of NaCl in the same buffer. The fractions were assayed as described previously.^{18,25} Active fractions were pooled and concentrated to afford a total of 29,400 units of aldolase activity with a specific activity of 43.9 U/mg.

Enzymatic synthesis of 3,4-dihydroxy-D-butanal

The synthesis of 3,4-dihydroxy-D-butanal was modified from a literature procedure.²⁵ 2-Deoxyribose-5-phosphate aldolase (20,000 U) was added to a 400 mL solution of 50 mM triethanolamine buffer (pH 7.3) containing glycoaldehyde dimer (2.25 g, 18.7 mmol), acetaldehyde (4.94 g, 112.2 mmol), and EDTA (0.15 g, 0.4 mmol). The resulting solution was stirred in the dark for 72 h under Ar. The reaction was quenched by addition of 1 L of acetone and then cooled to 0 °C for 1 h. Precipitated protein was removed by centrifugation. After removal of the solvent under reduced pressure, the residue was passed through Dowex 50 (H⁺) to remove triethanolamine and then

concentrated to a solution of 150 mL. This crude 3,4-dihydroxy-D-butanal was characterized by ^1H and ^{13}C NMR spectroscopy²⁶ as the major product and subsequent hydrogenation was carried out without further purification.

Catalytic hydrogenation of 3,4-dihydroxy-D-butanal

The solution obtained from the previous step was placed in a glass reaction vessel along with 5 wt % Ru on C (0.76 g, 1.0 mol %). The glass reaction vessel was inserted into the 500 mL Parr 4575 stainless steel high temperature-high pressure reactor and the vessel sealed. The temperature and stirring rate were controlled by a Parr 4842 temperature controller. Hydrogen was bubbled through the reaction mixture for 10-15 min to remove air while stirring at 100 rpm. The vessel was then charged with 13.6 atm H_2 . After heating the reaction to 30 °C, it was allowed to stir for 5 h at 30 °C. After removal of the catalyst by filtration through Celite, the reaction solution was concentrated to a yellow syrup, and a portion of the residue was derivatized by bis(trimethylsilyl)trifluoroacetamide followed by GC analysis. D-1,2,4-Butanetriol was obtained as a yellow oil (0.34 g, 17%) following Kugelrohr distillation (165 °C/ 0.2 mm Hg) of the syrup. The ^1H NMR and ^{13}C NMR were identical to authentic sample.

Chapter 3

Artificial biosynthesis of 1,2,4-butanetriol from erythritol

Plasmid pML5.176

The *gldBC* gene was amplified from *K. pneumoniae* ATCC25955 genomic DNA using 5'-GTAGTGGTCGACGTGCAACAGACAACCCAAAT as the forward primer, 5'-CTATACAAAGCTTGCTGACCTCCGCTTAGCTTC as the reverse primer (*Sa*I and *Hind*III restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1 kb was cloned into pJF118EH to afford pML5.176. Transcription of *gldBC* is in the same orientation as the *P_{lac}* promoter.

Plasmid pML5.179

The *dhaT* gene was amplified by PCR from *K. pneumoniae* ATCC25955 genomic DNA using 5'-CGGCAGAAAGCTTGAGAAGGTATATTATGAGCT-3' as the forward primer and 5'-GTTAGCAAAGCTTCCTCGTTAACTCAGAATG-3' as the reverse primer (*Hind*III restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.2 kb fragment was cloned into pJF118EH to afford pML5.179. Transcription of *dhaT* is in the same orientation as the *P_{lac}* promoter.

Plasmid pML5.200

The *gldA* gene (glycerol dehydratase α subunit) was amplified by PCR from *K. pneumoniae* ATCC25955 genomic DNA using 5'-CGTCGCGAATTCCCTTTGAGC-CGATGAACAA as the forward primer, 5'-GCTAACGTCGACACCGCCTTATTCA-ATGGTGT as the reverse primer (*Eco*RI and *Sa*I restriction sites are underlined), and

PfuTurbo as the DNA polymerase. The resulting 1.7 kb fragment was digested and ligated with pML5.176 to produce pML5.200. The *gldA* gene is transcribed in the same direction as *gldBC*.

Plasmid pML5.226

The *dhaT* gene was amplified by PCR from pWN5.022A using 5'-CGGCAG-AAGCTTGAGAAGGTATATTATGAGCT-3' as the forward primer and 5'-GTTAGCAAAGCTTCCTCGTTAACTCAGAAATG-3' as the reverse primer (*Hind*III restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.2 kb fragment was cloned into pML5.200 to afford pML5.226. The *dhaT* gene cluster is transcribed in the same direction as *gldABC*.

Plasmid pWN5.220A

The *dhaT* gene was amplified by PCR from *K. pneumoniae* ATCC25955 genomic DNA using 5'-ACGGATCCGCGAGAAGGTATATTATGAGC as the forward primer, 5'-TCGGATCCCCTCGTTAACTCAGAAATGC as the reverse primer (*Bam*HI restriction sites are underlined), and *PfuTurbo* as DNA polymerase. The resulting 1.2 kb fragment was digested and ligated with pT7-7 to produce pWN5.220A. Transcription of *dhaT* is in the same orientation as the phage T7 promoter.

Plasmid pWN5.260A

The *pddABC* gene cluster was amplified by PCR from *K. pneumoniae* ATCC25955 genomic DNA using 5'-CCGAATTCACATTTAATATCCGCGAGGC as the forward primer, 5'-CGGAATTCGCTTACCCCTGTTAATCGTC as the reverse primer (*Eco*RI restriction sites are underlined), and *PfuTurbo* as DNA polymerase. The

resulting 3.0 kb fragment was digested and ligated with pWN5.220A to produce pWN5.260A. The *pddABC* gene cluster is transcribed in the same direction as *dhaT*.

Plasmid pWN5.258A

The *pddABC-ddrAB* gene cluster was amplified by PCR from *K. oxytoca* ATCC8724 genomic DNA using 5'-CCGAATTCCGTCCTACATCTGATACCCA as the forward primer, 5'-CGGAATTCTGTTTCAACCAGTCCCAGTG as the reverse primer (*EcoRI* restriction sites are underlined), and *PfuTurbo* as DNA polymerase. The resulting 3.0 kb fragment was digested and ligated with pWN5.220A to produce pWN5.258A. The *pddABC-ddrAB* gene cluster is transcribed in the same direction as *dhaT*.

Anaerobic cultivation of Klebsiella and Lactobacillus strains

Anaerobic culturing of *Klebsiella pneumoniae*, *Klebsiella oxytoca* and *Lactobacillus brevis* strains followed the literature protocol.²⁷ All solutions were prepared in distilled, deionized water. Culture medium contained KH₂PO₄ (5.4 g), (NH₄)₂SO₄ (1.2 g), MgSO₄·7H₂O (0.4 g), tryptone (2.0 g), yeast extract (2.0 g) and carbon source (9 g) in 1 L H₂O. The buffering reagent, growth factor and carbon sources were prepared and autoclaved separately. Prior to autoclaving, solutions were saturated with N₂/CO₂ (95:5, v/v) gas by bubbling through a boiled solution. The process continued until the solution was cooled down to rt. Serum bottles containing treated solution were sealed with a butyl rubber stopper and aluminum cap. The cell culture was initiated by incubation of a single colony of bacterial strain into nutrient broth (5 mL) and cultured at 37 °C with agitation. A 5 mL aliquot of overnight culture was transferred into the serum bottle using a syringe. The inoculant was cultured at 37 °C without agitation. Four different carbon sources were

fed to the cultures. These include glycerol, erythritol, glycerol/erythritol (1:1, wt/wt), and glycerol/erythritol (1:8, wt/wt). Cell free culture medium of each strain was analyzed by GC.

In vitro dehydratase reaction of erythritol

In vitro enzymatic reactions contained potassium phosphate (33 mM, pH 8.6), KCl (50 mM), vitamin B₁₂ (15 µM), erythritol (0.2 M), and an appropriate amount of cell lysate. A 5 mL reaction mixture was setup and stirred for 24 h at rt. Formation of product 1,2,4-butanetriol was monitored by GC.

In search of 3-deoxy-glycero-pentulosonate decarboxylase genes for the microbial synthesis of 1,2,4-butanetriol from pentoses

Plasmid pML2.118

The potential decarboxylase gene was amplified from *Pseudomonas fluorescens* Pf-5 ATCC BAA477 genomic DNA using 5'-CGGAATTCGGCGGCTGCAAGAGT-AACTC as the forward primer, 5'-TCGGATCCGCGTCATCTATCACAGCAGT as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.118. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.123

The potential decarboxylase gene was amplified from *Pseudomonas aeruginosa* ATCC47085 genomic DNA using 5'-CGGAATTCGTAGCGCAAATCCAGGAAAC as the forward primer, 5'-ATGGATCCTCAGGGTTCGATGGTTTGCG as the reverse

primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.123. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.162

The potential decarboxylase gene was amplified from *Burkholderia fungorum* LB400 genomic DNA using 5'-GCGCCCCCGGGTCATTTGATAGATATTAAGG as the forward primer, 5'-TTCCCCCGGGTAAGTGGCTTGCTCATGGCT as the reverse primer (*Sma*I restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.162. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.208

The potential decarboxylase gene was amplified from *Streptomyces coelicolor* ATCC BAA471 genomic DNA using 5'-CGGAATTCTCACTTCGAAAGGAACGTCC as the forward primer, 5'-TAGGATCCTGATCTTCGGTGTGGTGTCTCG as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.208. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.214

The potential decarboxylase gene was amplified from *Novosphingobium aromaticivorans* ATCC700278 genomic DNA using 5'-CGGAATTCTGCGGCAGTTC-ATCCTCGGT as the forward primer, 5'-ATGGATCCAAGGGCGGTCGTCGATAAGG as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as

the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.214. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.256

The potential decarboxylase gene was amplified from *Rhodopseudomonas palustris* ATCC BAA98 genomic DNA using 5'-ATGAATTCAGGAAGCTTCGGCGG-TGGG as the forward primer, 5'-TCGGATCCGCTGGACAACCTGATGACGCA as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.256. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML2.286

The potential decarboxylase gene was amplified from *Mycobacterium smegmatis* ATCC700084 genomic DNA using 5'-ATGAATTCACTGGCAGGACGGCGTGAGC as the forward primer, 5'-GCGGATCCCATCGTTCGGTCTCCTGTTT as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML2.286. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML3.040

The potential decarboxylase gene was amplified from *Acidithiobacillus ferrooxidans* ATCC BAA477 genomic DNA using 5'-CGGAATTCGTTGCTTTGGGG-GGTTGACA as the forward primer, 5'-TCGGATCCTATGCCCAACCATTCATTC as the reverse primer (*Eco*RI and *Bam*HI restriction sites are underlined), and *PfuTurbo* as

the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118EH to afford pML3.040. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML3.054

All point mutations were introduced into the plasmid with a QuikChange site directed mutagenesis kit (Stratagene). The general protocol was described in the manufacturer's manual. Oligonucleotide primers carrying the mutation BFD A460I and flanked by complementary sequence of the plasmid template pWN5.238A were designed as follows: 5'-ATGAACAACGGCC_gTACGGTATCTTGCGATGG and 5'-CCATCGCAAGATA-CCGTAcGTGCCGTTGTTCAT (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type).²⁸ The PCR reaction mixture containing the primers (125 ng each), 1 µL supplied dNTP mix, 10 ng plasmid pWN5.238A, 5 µL reaction buffer and 2.5 U *PfuTurbo* DNA polymerase was made up to 50 µL using membrane filtered deionized H₂O. Following temperature cycling, the reaction mixture containing the 1.7 kb nicked plasmid was treated with *DpnI*, subsequently the mixture was transformed into XL1-Blue supercompetent cells. The transformants were plated on LB solid medium containing ampicillin. Plasmid was purified from a single colony and the mutation was confirmed by sequencing using primers 5'-ATAACGGTTCTGGC-AAATATT, 5'-GCCGGCTTTCATTAACCTGCAT, 5'-GATCCTCAGTCCCACCAC-CTTT, 5'-CGAAGGTCACGATGTGGTTTTG, and 5'-GAGAATGCGATTTACCTG-AACG. The resulting plasmid was designated as pML3.054.

Plasmid pML3.062

Oligonucleotide primers carrying the mutation PDC I472A and flanked by complementary sequence of the plasmid template pLOI276²⁹ were designed as follows: 5'-GATCAATAACTATGGTTACACCgcCGAAGTTATGATCCATGATG and 5'-CATC-ATGGATCATAACTTCGgcGGTGTAACCATAGTTATTGATC (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type). The point mutation was generated as described earlier and confirmed by sequencing analysis using primers 5'-CAAGGAAACAGCTATGACC, 5'-ACAACCTCGTCCT-TCTTGAC, 5'-GAAGAAGCCGGTTTATCTC, 5'-ACCGCGTT-CTGTCGTCGTTA, and 5'-CAGGAAGTCGCTCAATGGT. The resulting plasmid was designated as pML3.062.

Plasmid pML3.084

Oligonucleotide primers carrying the mutation PDC I472S and flanked by complementary sequence of the plasmid template pLOI276²⁹ were designed as follows: 5'-TCAATAACTATGGTTACACCAgCGAAGTTATGATCCATGATG and 5'-CATCATGGATCATAACTTCGcTGGTGTAACCATAGTTATTGA (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type). The point mutation was generated as described earlier and confirmed by sequencing analysis using the same primer set as for generating plasmid pML3.062. The resulting plasmid was designated as pML3.084.

Plasmid pML3.086

Oligonucleotide primers carrying the mutation PDC I472G and flanked by complementary sequence of the plasmid template pLOI276²⁹ were designed as follows: 5'-GATCAATAACTATGGTTACACCggCGAAGTTATGATCCATGATG and 5'-CATCATGGATCATAACTTCGccGGTGTAACCATAGTTATTGATC (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type). The point mutation was generated as described earlier and confirmed by sequencing analysis using the same primer set as for generating plasmid pML3.062. The resulting plasmid was designated as pML3.086.

Plasmid pML3.104

Oligonucleotide primers carrying the mutation PDC I472T and flanked by complementary sequence of the plasmid template pLOI276²⁹ were designed as follows: 5'-TCAATAACTATGGTTACACCAAcCGAAGTTATGATCCATGATG and 5'-CATCATGGATCATAACTTCGgTGGGTGTAACCATAGTTATTGA (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type). The point mutation was generated as described earlier and confirmed by sequencing analysis using the same primer set as for generating plasmid pML3.062. The resulting plasmid was designated as pML3.104.

Directed evolution of 3-deoxy-*glycero*-pentulosonate decarboxylase for microbial 1,2,4-butanetriol synthesis

Plasmid pML3.200

Oligonucleotide primers carrying the deletion of *Eco*RI site and flanked by complementary sequence of the plasmid template pML2.162 were designed as follows: 5'-CCAACGCGTGGAAAcTCGCATACGCCG and 5'-CGGCGTATGCGAgTTCCACG-CGTTGG (The mutated codons are underlined, and the lowercase letters indicate a base change from wild-type). The resulting plasmid was designated as pML3.171. To delete the *Bam*HI site, another set of primers were designed as follow: 5'-GCGGCGGCGGGCATCCAGCTCGCG and 5'-CGCGAGCTGGATGCCCCGCCGCC-
GC. The resulting plasmid was designated pML3.200. These mutations were verified by PCR analysis using primers 5'-ATAACGGTTCTGGCAAATATT and 5'-GTGGGACCACCGCGCTACTGC.

Plasmid pML3.224

The *P. fluorescens* parent gene candidate was amplified from plasmid pML2.118 using 5'-TACGCTCAAAGCTTGCCTCCATGAAAACCGTCCA as the forward primer, 5'-GCGTCTGAATTCTCTCTCATCCGCCAAAACAG as the reverse primer (*Hind*III and *Eco*RI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118HE to afford pML3.224. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML3.225

The *P. aeruginosa* parent gene candidate was amplified from plasmid pML2.123 using 5'-TACGCTCAAAGCTTGCCTCCATGAAAACCGTCCA as the forward primer, 5'-GCGTCTGAATTCTCTCTCATCCGCCAAAACAG as the reverse primer (*Hind*III and *Eco*RI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118HE to afford pML3.225. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Plasmid pML3.226

The *B. fungorum* parent gene candidate was amplified from plasmid pML3.200 using 5'-TACGCTCAAAGCTTGCCTCCATGAAAACCGTCCA as the forward primer, 5'-GCGTCTGAATTCTCTCTCATCCGCCAAAACAG as the reverse primer (*Hind*III and *Eco*RI restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.7 kb gene was cloned into pJF118HE to afford pML3.226. Transcription of this ORF is in the same orientation as the *P_{lac}* promoter.

Chimera library generation for ssDNA family shuffling

The original DNA shuffling was invented by Stemmer.³⁰ ssDNA shuffling is a similar DNA chimera library generation technique developed by Zhao and Arnold.³¹ The decarboxylase genes of *P. fluorescens*, *P. aeruginosa* and *B. fungorum* were amplified from plasmids pML3.224, pML3.225 and pML3.226, respectively, using the following primers: 5'-TACGCTCAAAGCTTGCCTCCATGAAAACCGTCCA and 5'-GCGTCT-GAATTCTCTCTCATCCGCCAAAACAG. The *Hind*III and *Eco*RI sites are underlined. Phosphorylated forward or reverse primer was used to yield the desired single DNA

strand. A 100 μ L reaction mixture contained: 10 μ L of 10 \times *Pfu* buffer, 10 μ L of 2 mM dNTP mix, 50 pmol of each primer, 8 μ L of dimethylsulfoxide (DMSO), 50 ng of template plasmid and 2.5 U of *PfuTurbo* DNA polymerase. Amplification of DNA was carried out for 35 cycles under the following conditions: 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 2.5 min.

5 μ g of each purified decarboxylase gene candidates were diluted to 45 μ L with sterile water and 5 μ L of 10 \times lambda exonuclease buffer was added. Each mixture was digested with 10 U lambda exonuclease at 37 °C for 90 min. The resulting ssDNA was then purified by agarose gel electrophoresis. 0.5 μ g of each purified ssDNA was mixed and diluted to 42.5 μ L, treated with 5 μ L of 0.5 M Tris-HCl (pH 7.5) and 2.5 μ L of 0.2 M MnCl₂, and digested with 0.02 U DNase I at 15 °C for 30 sec. The reaction was terminated by the addition of endostop and was purified by low melting agarose gel electrophoresis (1%). 500 – 1,000 bps DNA fragments were recovered using a Zymoclean gel extraction kit.

Fragments were reassembled in a 20 μ L reaction mixture containing 5 μ L of purified fragment DNA, 2 μ L of 10 \times *Pfu* buffer, 1 μ L DMSO, 2 μ L dNTP mix (2 mM each) and 1 U *PfuTurbo* DNA polymerase. Cycling protocol: 96 °C for 90 seconds; 50 cycles of (94 °C for 1 min; 55 °C for 1 min; 72 °C for 2.5 min); 72 °C for 7 min and 4 °C thereafter. This step was monitored by running a small aliquot on an agarose gel until a smear extending through 1.6 kb was seen.

To amplify full-length 1.6 kb genes, the reassembly reaction served as the DNA template and was diluted 2-fold. The same 50 μ L PCR mixture used to acquire DNA for lambda exonuclease digestion was then setup. The PCR program is as follow: 4 min at

96 °C followed by 35 cycles of 1 min at 94 °C, 1 min at 58 °C, 2.5 min at 72 °C and finally 7 min at 72 °C. The final full-length PCR product was purified by agarose gel electrophoresis. The purified genes were digested by *Hind*III and *Eco*RI and subsequently cloned into pJF118HE for *in vivo* screening.

Chimera library generation for dsDNA family shuffling

Chimera library for dsDNA family shuffling was generated according to a similar procedure described by Joern and Arnold.³² *P. fluorescens*, *P. aeruginosa* and *B. fungorum* genes were amplified using non-phosphorylated primers: 5'-TACGCTCAAG-CTTGCCTCCATGAAAACCGTCCA and 5'-GCGTCTGAATTCTCTCTCATCCGCCAAAACAG from pML3.224, pML3.225, and pML3.226, respectively. A 100 µL reaction mixture contained: 10 µL of 10× *Pfu* buffer, 10 µL of 2 mM dNTP mix, 50 pmol of each primer, 8 µL of dimethylsulfoxide (DMSO), 50 ng of template plasmid and 2.5 U of *PfuTurbo* polymerase was setup. Amplification of DNA was carried out for 35 cycles under the following conditions: 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 2.5 min. The *P. putida* gene was obtained by restriction enzyme digestion from pWN5.238A using enzymes *Eco*RI and *Bam*HI. All four gene candidates were purified by agarose gel electrophoresis prior to use.

After purification and quantification, equal amounts of parent DNA (as determined spectroscopically by UV-vis absorbance, $\lambda = 260$ nm) were mixed and subjected to DNase I digestion. A 50 µL digestion contained 4 µg of parent DNA mix, 5 µL of 0.5 M Tris-HCl (pH 7.4), 2.5 µL of 0.2 M MnCl₂, and 0.02 U of DNase I. After 4 min digestion at 15

°C, 5 µL of endostop was added. Using Zymoclean gel extraction kit, fragments from 100 – 300 kb were extracted.

Fragments were reassembled in a 50 µL reaction containing all the purified DNA fragments, 10 µL of 2 mM dNTP mix, 5 µL of *Pfu* buffer, 2 µL DMSO and 2.5 U of *PfuTurbo* DNA polymerase. Cycling was according to the following protocol: 96 °C, 90 sec; 35 cycles of (94 °C, 30 s; 65 °C, 90 s; 62 °C, 90 s; 59 °C, 90 s; 56 °C, 90 s; 53 °C, 90 s; 50 °C, 90 s; 47 °C, 90 s; 44 °C, 90 s; 41 °C, 90 s; 72 °C, 4 min); 72 °C, 7 min; 4 °C thereafter. 5 µL of the reaction mixture was analysed on the gel and a smear of reassembled DNA that extended above the parent gene size (1.6 kb) was seen.

PCR was carried out for the reassembly mixture using primers 5'-TACGC-TCAAGCTTGCCTCCATGAAAACCGTCCA, and 5'-GCGTCTGAATTCTCTCTC-ATCCGCCAAAACAG under the same conditions used to acquire DNA for fragmentation. The DNA template used in the reaction, however, was diluted 80-fold. Cycling was done using the following program: 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 4 min. The final full-length PCR product was purified by agarose gel electrophoresis. The purified genes were digested by *HindIII* and *EcoRI* and subsequently cloned into pJF118HE for *in vivo* screening.

Library screening for 3-deoxy-glycero-pentulosonate decarboxylase

Library screening was carried out in a 96-well format. Purified genes obtained from the family shuffling were transformed into electrocompetant *E. coli* W3110 and plated on M9/xylonate/Ap plates. This step was optimized such that the amount of colonies appeared on each plate was around 200. Single colonies were then used for

screening using the protocol as follows. The daily throughput of this assay is 200 colonies.

A fresh growth block (Qiagen) with each well containing 1 mL of M9/xylionate/Ap medium was inoculated with single colonies from the chimera library. The first column of each block was inoculated with the controls. Wells A1 and H1 were blanks. Wells B1 – G1 were inoculated with *E. coli* W3110/pWN5.238A, W3110/pWN5.238A* (this plasmid harbors an *mdlC* mutant isolated earlier with 2-fold activity improvement³³), W3110/pML3.224, W3110/pML3.225, W3110/pML3.226, and W3110/pWN5.238A in order. The growth block was then covered by Qiagen Airpore Sheet and incubated in the shaker at 37 °C for 36 hours.

The growth block was removed from the shaker and 150 µL was withdrawn from each well and poured into an autoclaved 96-well plate (Nunc Inc.) containing 50 µL of 80% sterilized glycerol. The glycerol freeze plate was then covered with aluminum tape and stored at –80 °C. From the same growth block, another 10 µL was drawn from each well and used to inoculate another fresh block. The new block was then covered with Qiagen Airpore Sheet and incubated at 37 °C for another 24 hours.

The growth block was removed from the shaker and 100 µL aliquot was poured into a 96-well optical plate and the OD₆₀₀ was measured. At this stage, the OD₆₀₀ should be around 0.6 and IPTG was added to a final concentration of 0.5 mM. The induced culture block was then covered with a fresh piece of Airpore Sheet and was incubated at 37 °C for 48 hours.

The growth block was removed from the shaker and the final OD₆₀₀ was recorded. The cell was then centrifuged down at 400 rpm for 10 min. 100 µL supernatant from each

well was pipetted to a new block (Fisherbrand, polypropylene) and this block was placed in the CentriVap (Labconco). The supernatant was evaporated to dryness within 2 hours by setting the core temperature to 70 °C. The residue in each well was then resuspended into 10 μ L H₂O and submitted to GC derivatization using bis(trimethylsilyl)-trifluoroacetamide (BSTFA). The BSTFA solution mix was prepared by mixing 50 mL pyridine, 50 mL BSTFA and 500 μ L dodecane *in situ*. The reaction began by pouring 1 mL of this solution mix into each well of the block. The block was sealed tight with a polypropylene lid (Fisherbrand) and allowed to react overnight with agitation at 250 rpm in a shaker.

Finally, the solution from the block was pipetted into GC sample vials for analysis. The GC method used was developed to be isothermal with a 7 min method runtime. Therefore, the time required for analyzing 200 samples is 24 hours. In a typical GC chromatogram, the BT comes out at 2.86 min while the internal standard, dodecane comes out at 2.34 min. Enzymes corresponding to these putatively more reactive decarboxylases were assayed for a second time using the same assay. Confirmed decarboxylase mutants were assayed for a third time by culturing in 100 mL M9/xylonate/Ap medium.

Bacteriophage T7 promoter and *E. coli* chaperones for *mdlC* expression

E. coli KIT2

The genomic antibiotic marker on *E. coli* WN13 was excised by the method described by Cherepanov et al.³⁴ *E. coli* WN13 was transformed with plasmid pCP20, and

ampicillin resistant transformants were selected at 30 °C, after which a few were colony-purified once non selectively at 43 °C . They were then tested for the loss of all antibiotic resistances. Disruption of the genomic chloramphenicol marker was confirmed by PCR analysis using the following primers: 5'-TACGACATCATCCATCACCCGCGGCA-TTAC and 5'-CAGAAGTGGC-TGATAGAGGCGACGGAACGT. This W3110 Δ yjhH- Δ yagEserAxylAB::xdh-FRT strain was designated as *E. coli* KIT10.

E. coli KIT15

E. coli KIT2 was grown in LB supplemented with 0.2% maltose, 10 mM MgSO₄ at 37 °C in the shaker to an OD₆₀₀ = 0.5. The DE3 lysogen of *E. coli* KIT2 was then prepared using the DE3 lysogenation kit of Novagen according to the manufacturer's protocol. *E. coli* KIT15 was the result and this new strain was used as host for phage T7 expression vectors.

Plasmid pML7.128

A 1.6 kb fragment encoding the *P. putida* benzolformate decarboxylase gene *mdlC* was excised from plasmid pWN5.238A by digestion with *Eco*RI and ligated to the 5.4 kb plasmid pET28c(+), which had been previously treated with *Eco*RI, followed by CIAP, to afford a 7.0 kb plasmid pML7.128.

Plasmid pML7.135

A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the 6.5 kb plasmid pML7.128, which had been previously treated with *Sma*I and CIAP. The ligation mixture was transformed into *E. coli* WN13. Transformants

carrying the *serA* insert were selected on M9 medium plates. In the resulting 8.7 kb plasmid pML7.135, the *serA* and *mdlC* genes are convergently transcribed.

Plasmid pML7.166

To construct a plasmid harboring the *P_{tac}-groEL-groES* gene fragment, a 3.4 kb fragment carrying the *groEL-groES* chaperones and the trigger factor (TF) genes without the tetracycline induced promoter Pzt1 sequence was amplified by PCR using 5'-CGG-AAGCTTATGAATATTCGTCCATTGCA as the forward primer and 5'-ATAAAGCTTA-CGGGCCTTTGTGCGAATTTA as the reverse primer (*Hind*III restriction sites are underlined) from the plasmid pG-Tf2 (Takara Bio Inc.). The PCR product was digested with *Hind*III and ligated with *Hind*III digested pJF118HE to afford the 8.8 kb plasmid pML7.166.

Plasmid pML7.175

The 1.5 kb *P_{tac}-groEL-groES* locus was amplified by PCR using 5'-ACCCCGGGCGTAAATCACTGCATAATTCG as the forward primer and 5'-ACCCCGGGTACCTCAAAAAATCACAGTGG as the reverse primer (*Sma*I restriction sites are underlined) from plasmid pML7.166. The PCR product was digested with *Sma*I and ligated to the 7.0 kb plasmid pWN5.238A, which had been sequentially treated with *Sma*I and CIAP, to afford the 8.5kb plasmid pML7.175.

Plasmid pML7.180

A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the 8.5 kb plasmid pML7.128, which had been previously treated with *Sca*I and CIAP. The ligation mixture was transformed into *E. coli* WN13. Transformants

carrying the *serA* insert were selected on M9 medium plates. In the resulting 10.1 kb plasmid pML7.135, the *serA* and *P_{tac}-mdlC-P_{tac}-groEL-groES* are convergently transcribed.

Plasmid pML7.202

A 1.5 kb fragment carrying the *P_{tac}-groEL-groES* region was amplified by PCR using the same method as described earlier. The *Sma*I digested PCR product was ligated into pML7.135, which had been sequentially treated with *Bgl*II, Klenow and CIAP, to afford the 10.2 kb plasmid pML7.202.

Discovery of alcohol dehydrogenases for D-1,2,4-butanetriol biosynthesis

E. coli KIT1

The chromosomal gene insertion in *E. coli* was generated by modifying a method described by Datsenko *et al.*³⁵ To replace native *E. coli xylAB* with the *xdh-adhP-P_{tac}* along with a chloramphenicol selectable marker in the chromosome, the PCR-mediated λ -Red recombination methodology was employed. A linear DNA cassette including these genes was generated by PCR from pML6.272 using Platinum *Taq* High Fidelity DNA polymerase (Invitrogen), with 5'-CACCCGCGGCATTACCTGATTATGGAGTTCA-ATATGTCCTCAGCCATCTATCCC as the forward primer and 5'-CAGAAGTTGCTGATAGAGGCGACGGAACGTTTCTCATATGAATATCCTCCTTA-GT as the reverse primer. The sequences underlined represent the homology arms, while the remainders are the priming sequences for hybridization to complementary nucleotide sequences on the template plasmid pML6.272. The 3 kb PCR product was agarose gel-

purified and transformed into electrocompetent *E. coli* W3110/pKD46. SOC medium was added after electroporation and incubated at 37 °C for 1 h. The cells were spread on LB plates containing 25 µg/mL chloramphenicol and incubated overnight at 37 °C to select for antibiotic transformants. The λ-Red recombinase expression plasmid pKD46 has a temperature sensitive replication origin and can be eliminated at 37 °C. Strain KIT1 (W3110*xylAB::xdh-adhP-P_{tac}*-FRT-Cm^R-FRT) was therefore constructed. The gene insertion was verified by PCR analysis using primers 5'-TACGACATCATCC-ATCACCCGCGGCATTAC and 5'-CAGAAGTGGCTGATAGAGGCGACGGAACGT.

E. coli KIT3

P1 phage mediated transduction was used to transfer the *xylAB::xdh-adhP-P_{tac}*-FRT-Cm^R-FRT mutation into *E. coli* WN7. *E. coli* strain KIT1 was used as the donor strain while strain WN7 served as the recipient strain. The transformants were selected on LB/Cm plate to produce another new strain KIT3 (W3110Δ*yjhH*Δ*yagE**serA**xylAB::xdh-adhP-P_{tac}*-FRT-Cm^R-FRT). The gene insertion was verified by PCR analysis using primers 5'-TACGACATCATCCATCACCCGCGGCATTAC and 5'-CAGAAGTGGC-TGATAGAGGCGACGGAACGT.

E. coli KIT4

The genomic antibiotic marker on *E. coli* KIT3 was excised by the method described by Cherepanov *et al.*³⁴ *E. coli* strain KIT3 was electroporated with plasmid pCP20 containing Flp recombinase gene, and grown on LB plate containing 100 µg/mL ampicillin at 30 °C overnight. The Flp recombinase excises the chloramphenicol gene flanked by FRT sites from the chromosome. After growth at 42 °C for one day in LB to

cure KIT3/pCP20 that has a temperature sensitive replication origin, ampicillin and chloramphenicol sensitivity were tested to verify loss of pCP20 and the antibiotic marker. The loss of the antibiotic gene marker was verified by PCR analysis using primers 5'-TACGACATCATCCATCACCCGCGGCATTAC and 5'-CAGAAGTGGCTGATAGAG-GCGACGGAACGT. The resulting strain was designated as *E. coli* KIT4 (W3110 Δ yjhH Δ yagEserAxylAB::xdh-adhP-*P_{lac}*-FRT).

E. coli KIT5

The chromosomal gene insertion in *E. coli* was generated by the same protocol used to generate strain KIT1. A 1.2 kb fragment carrying the *P₇₅*-FRT-Cm^R-FRT sequence was amplified by PCR using Platinum *Taq* High Fidelity DNA polymerase, with 5'-TCTCCCTTTTTCAGATCTCATCCATACTGGGTAGTGGCGAATAAACGTCTTG-AGCGATTGTGTAG as the forward primer and 5'-AACTGCAGCCTTCATAGTTC-CTCCTTTTCGGATGATGTTCTGCATAGTTAATTTCTCCTCTTTAATG as the reverse primer. The sequences underlined represent the homology arms, while the remainders are the priming sequences for hybridization to complementary sequences on the template plasmid pML7.042. The PCR product was agarose gel purified and was transformed into electrocompetent *E. coli* W3110/pKD46. The transformants were selected on LB/Cm plate and verified by PCR analysis using primers 5'-DAGGG-ATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAAT-CAATCAC. *E. coli* strain KIT5 (W3110*P₇₅*-FRT-Cm^R-FRT) was thus constructed.

E. coli KIT6

E. coli strain KIT6 (W3110 Δ yjhH Δ yagEserAxylAB::xdh-FRT P_{T5} -FRT-Cm^R-FRT) was made from *E. coli* strain WN13 by P1 phage mediated transduction using KIT5 as the donor strain. The newly constructed strain KIT6 was selected on an LB/Cm plate. The inserted P_{T5} -FRT-Cm^R-FRT gene fragment was verified by PCR analysis using primers 5'-DAGGGATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAATCAATCAC.

E. coli KIT7

The genomic chloramphenicol marker was excised by the same protocol for generating *E. coli* KIT4. Thermal induction of the Flp recombinase in KIT6/pCP20 excised the chloramphenicol marker and afforded strain KIT7 (W3110 Δ yjhH Δ yagEserAxylAB::xdh-FRT P_{T5} -FRT). The loss of the antibiotic marker gene was verified by PCR analysis using primers 5'-DAGGGATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAATCAATCAC.

E. coli KIT8

The genomic *adhP* inactivation was generated by the methods described by Datsenko et al.³⁵ Linear DNA with an FRT-flanked chloramphenicol gene was amplified by PCR from pKD3 using Platinum *Taq* DNA polymerase, with 5'-ACTGGGTAGTGGCGAATAAATCTCATTTCGCTCACCTGCTGTGTAGGCTGGA-GCTGCTTC as the forward primer, and 5'-GATGCTGCGACCCGAACATGGCAGTC-GCAGCAAAGGCCTCCATATGAATATCCTCCTTAG as the reverse primer. The sequence underlined represent the homology arms, while the remainders are the priming

sequences for hybridization to complementary sequences on the template plasmid pKD3. The 1 kb PCR product was agarose gel purified and subsequently transformed into electrocompetent *E. coli* W3110/pKD46. The transformants were selected on an LB/Cm plate. Disruption of *adhP* was confirmed by PCR analysis using the following primers: 5'-DAGGGATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAATCAATCAC. The W3110*adhP*::FRT-Cm^R-FRT strain was designated as *E. coli* KIT8.

E. coli KIT9

P1 phage mediated transduction was employed to transfer the *adhP*::FRT-Cm^R-FRT mutation into *E. coli* KIT2 to produce KIT9 (W3110Δ*yjhH*Δ*yagE**serAxylAB::xdh*-FRT*adhP*::FRT-Cm^R-FRT). Disruption of *adhP* in *E. coli* KIT9 was confirmed by PCR analysis using the following primers: 5'-DAGGGATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAATCAATCAC.

E. coli KIT10

The genomic chloramphenicol marker was excised by the same protocol for generating *E. coli* KIT4. Thermal induction of the Flp recombinase in KIT9/pCP20 excised the chloramphenicol marker and afforded strain KIT7. The loss of the antibiotic marker gene was verified by PCR analysis using primers 5'-DAGGGATCCATGCAGAACATCATCCGAAAA and 5'-AGGGATCCTTAGTGACGGAAATCAATCAC. This W3110Δ*yjhH*Δ*yagE**serAxylAB::xdh*-FRT*adhP*::FRT strain was designated as KIT10.

E. coli KIT16

The *E. coli* genomic *yiaE* inactivation was generated by the same protocol for generating *E. coli* KIT8. The FRT-flanked chloramphenicol gene was amplified from pKD3 using 5'-CGCCCGAAAAATTCTGCTGATGCACACGCCAACCGTATTATGTGTAGGCTGGAGCTGCTTCG as the forward primer, and 5'-AGATCCAATATGCGGTACTGCGACGACGTTGGCCATTGAGCATATGAATATCCTCCTTAGT as the reverse primer. The sequence underlined represent the homology arms, while the remainders are the priming sequences for hybridization to complementary sequences on the template plasmid pKD3. The 1 kb PCR product was agarose gel purified and subsequently transformed into electrocompetent *E. coli* W3110/pKD46. The transformants were selected on an LB/Cm plate. Disruption of *yiaE* was confirmed by PCR analysis using the following primers: 5'-CGGAATTCGAATGGAGAGAAGCATGAAG and 5'-CGGAATTCGATTATCGGGCTTTACTCTA. This *E. coli* W3110 $\Delta yjhH\Delta yagE-serAxylAB::xdh-adhP-P_{lac}$ -FRT*yiaE*::FRT-Cm^R-FRT was designated as KIT16.

E. coli KIT17

P1 phage mediated transduction was employed to transfer the *ycdW*::FRT-Km^R-FRT mutation into *E. coli* KIT16 to produce KIT17 (W3110 $\Delta yjhH\Delta yagEserAxylAB::xdh-adhP-P_{lac}$ -FRT*yiaE*::FRT-Cm^R-FRT*ycdW*::FRT-Km^R-FRT). The phage (W3110-*ycdW*::FRT-Km^R-FRT) was obtained in the Frost group strain collection. Disruption of *ycdW* in *E. coli* KIT17 was confirmed by PCR analysis using the following primers: 5'-CGGAATTCAGTATGGATATCATCTTTT and 5'-CGGAATTCCTGATGCTTTATTAGTAG.

E. coli KIT18

The genomic chloramphenicol and kanamycin markers were excised by the same protocol for generating *E. coli* KIT4. Thermal induction of Flp recombinase of KIT17/pCP20 excised both antibiotic markers. The loss of antibiotic marker gene was verified by PCR analysis using 5'-CGGAATTCGAATGGAGAGAAGCATGAAG, 5'-CGGAATTCGATTATCGGGCTTTACTCTA, 5'-CGGAATTCAGTATGGATATCATCTTTT and 5'-CGGAATTCCTGATGCTTTATTAGTAG This W3110 $\Delta yj h H \Delta y a g E$ -*serAxylAB::xdh-adhP-P_{tac}-FRTyiaE::FRTycdW::FRT* strain was designated as KIT18.

Plasmid pML5.179

The *dhaT* gene was amplified by PCR from *K. pneumoniae* genomic DNA using 5'CGGCAGAAAGCTTGAGAAGGTATATTATGAGCT as the forward primer and 5'-GTTAGCAAAGCTTCCTCGTTAACACTCAGAATG as the reverse primer (*Hind*III restriction sites are underlined), and *PfuTurbo* as the DNA polymerase. The resulting 1.2 kb fragment was cloned into pJF118EH to afford a 6.5 kb plasmid pML5.179.

Plasmid pML6.090

A 1.2 kb *dhaT* gene was excised from plasmid pML5.179 by digestion with *Bam*HI. The resulting gene fragment was treated with Klenow, gel purified and ligated to *Eco*RI-digested, Klenow treated and dephosphorylated pKK223-3 to generate the plasmid pML6.090.

Plasmid pML6.128

A 1.5 kb fragment carrying the *tac* promoter region and the *dhaT* gene was excised from plasmid pML6.090 by digestion with *Bam*HI and ligated to pWN5.238A, which had been sequentially treated with *Bam*HI and CIAP, to afford a 8.5 kb plasmid pML6.128.

Plasmid pML6.133

A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the 8.4 kb plasmid pWN7.096B, which had been previously treated with *Sca*I and CIAP. The ligation mixture was transformed into electrocompetent *E. coli* WN13. Transformants carrying the *serA* insert were selected on M9/glucose medium plates. This generated the 10.2 kb plasmid pML6.133.

Plasmid pML6.135

A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the 8.4 kb plasmid pML6.128, which had been previously treated with *Sca*I and CIAP. The ligation mixture was transformed into electrocompetent *E. coli* WN13. Transformants carrying the *serA* insert were selected on M9/glucose medium plates. This generated the 10.2 kb plasmid pML6.135.

Plasmid pML6.166

E. coli adhP gene was amplified using 5'-DAGGGATCCATGCAGAACATCAT-CCGAAAA as the forward primer, 5'-AGGGATCCTTAGTGACGGAAATCAATCAC as the reverse primer (*Bam*HI restriction sites are underlined) using *PfuTurbo* as the DNA polymerase. The resulting 1 kb PCR products was digested with *Bam*HI, Klenow, gel

purified and ligated to *Eco*RI-digested, Klenow treated and dephosphorylated pKK223-3 to generate the plasmid pML6.166.

Plasmid pML6.168

E. coli adhE gene was amplified using 5'-ATGGATCCATGGCTGTTACTAA-TGTCGC as the forward primer, 5'-CGGGATCCTTAAGCGGATTTTTTCGCTT as the reverse primer (*Bam*HI restriction sites are underlined) using *PfuTurbo* as the DNA polymerase. The resulting 2.7 kb PCR products were digested with *Bam*HI, Klenow, gel purified and ligated to *Eco*RI-digested, Klenow treated and dephosphorylated pKK223-3 to generate plasmids pML6.168.

Plasmid pML6.185

A 1.5 kb fragment carrying the *tac* promoter region and the *adhP* gene was excised from plasmid pML6.166 by digestion with *Bam*HI and ligated to pWN5.238A, which had been sequentially treated with *Bam*HI and CIAP, to afford a 8.5 kb plasmid pML6.185.

Plasmid pML6.195

A 1.6 kb *serA* locus was excised from plasmid pRC1.55B by digestion with *Sma*I and ligated to the 8.4 kb plasmid pML6.185, which had been previously treated with *Sca*I and CIAP. The ligation mixture was transformed into electrocompetent *E. coli* WN13. Transformants carrying the *serA* insert were selected on M9/glucose medium plates. This generated the 10.1 kb plasmid pML6.195.

Plasmid pML6.255

A 1.3 kb fragment carrying the *tac* promoter region and the *adhP* gene was amplified by PCR using 5'-CGACGCGTCCCCGTTCTGGATAATGTTTT as the forward primer and 5'-CTACGCGTATCCGCCAAAACAGAAGCTT as the reverse primer (*A**f**III* restriction sites are underlined) from plasmid pML6.166 using Platinum *Taq* High Fidelity DNA polymerase. The PCR product was digested with *A**f**III* and ligated to a 2.8 kb plasmid pKD3, which had been sequentially treated with *A**f**III* and CIAP, to afford the 4 kb plasmid pML6.255.

Plasmid pML6.259

A 1.0 kb fragment encoding the *yhdH* gene was amplified by PCR from *E. coli* W3110 genomic DNA using 5'-AAGGATCCCCACTTATGCAGGCGTTACT as the forward primer and 5'-AGGGATCCGTTAGTTAACCTTCACCAGC as the reverse primer (*Bam*HI restriction sites as underlined) using Platinum *Taq* High Fidelity DNA polymerase (Invitrogen). The resulting gene was cloned into *Bam*HI digested pJF118EH to afford pML6.259. Transcription of *yhdH* gene is in the same orientation as the *P_{lac}* promoter.

Plasmid pML6.261

A 1.2 kb fragment encoding the *yiaY* gene was amplified by PCR from *E. coli* W3110 genomic DNA using 5'-AGGGATCCATGGCAGCTTCAACGTTCTT as the forward primer and 5'-AGGGATCCTGATGATTACCTCGCTGCAT as the reverse primer (*Bam*HI restriction sites as underlined) using Platinum *Taq* High Fidelity DNA

polymerase. The resulting gene was cloned into *Bam*HI digested pJF118EH to afford pML6.261.

Plasmid pML6.263

A 0.8 kb fragment encoding the *ydjO* gene was amplified by PCR from *E. coli* W3110 genomic DNA using 5'-ATGGATCCATGAGGCGGGAGATGTTTTTC as the forward primer and 5'-ATGGATCCTTCTTATGCTGGCAACGGTC as the reverse primer (*Bam*HI restriction sites as underlined) using Platinum *Taq* High Fidelity DNA polymerase. The resulting gene was cloned into *Bam*HI digested pJF118EH to afford pML6.263.

Plasmid pML6.272

A 0.8 kb fragment encoding the *C. crescentus* xylose dehydrogenase gene *xdh* was excised from plasmid pWN9.068 by digestion with *Sph*I and ligated to the 4 kb plasmid pML6.255, which had been previously treated with *Sph*I and CIAP. In the resulting 4.8 kb plasmid pML6.272, the *adhP* gene is transcribed by the *tac* promoter, while the *xdh* gene was cloned alone without any promoter.

Plasmid pML7.042

A fragment containing a phage T5 promoter, *lacO* and ribosome binding site was amplified by PCR from plasmid pQE30 using 5'-GGGAATTCCCATATGCGTCTTCAC-CTCGAGAAATC as the forward primer, 5'-GGGAATTCCCATATGGTGATGCG-ATCCTCTCATAG as the reverse primer (*Nde*I restriction sites are underlined) using Platinum *Taq* High Fidelity DNA polymerase. The resulting 0.2 kb PCR product was

digested with *Nde*I, gel purified and ligated with *Nde*I digested and dephosphorylated pKD3 vector, generating the 3.0 kb plasmid pML7.042.

REFERENCE

- ¹ Niu, W.; Frost, J. W. unpublished results.
- ² (a) Valentin-Hansen, P.; Aiba, H.; Schumperli, D. *EMBO J.* **1982**, *1*, 317. (b) Jorgensen, P.; Collins, J.; Valentin-Hansen, P. *Mol. Gen. Genet.* **1977**, *155*, 93.
- ³ Seifert, C.; Bowien, S.; Gottschalk, G.; Daniel, R. *Eur. J. Biochem.* **2001**, *268*, 2369.
- ⁴ Luers, F.; Seyfried, M.; Daniel, R.; Gottschalk, G. *FEMS Microbiol. Lett.* **1997**, *154*, 337.
- ⁵ Bobik, T. A.; Xu, Y.; Jeter, R. M.; Otto, K. E.; Roth, J. R. *J. Bacteriol.* **1997**, *179*, 6633.
- ⁶ Yanisch-Perron, C.; Vieira, J.; Messing, J. *Gene* **1985**, *33*, 103.
- ⁷ Conway, T.; Sewell, G. W.; Osman, Y. A.; Ingram, L. O. *J. Bacteriol.* **1987**, *169*, 2591.
- ⁸ Keshav, K. F.; Yomano, L. P.; An, H. J.; Ingram, L. O. *J. Bacteriol.* **1990**, *172*, 2491.
- ⁹ Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. USA*, **1987**, *84*, 4767.
- ¹⁰ Brosius, J.; Holy, A. *Proc. Natl. Acad. Sci. USA*, **1984**, *81*, 6929.
- ¹¹ Fürste, J. P.; Pansegrau, W.; Frank, R.; Blocker, H.; Scholz, P.; Bagdasarian, M.; Lanka, E. Molecular cloning of the plasmid Rp4 primase region in a multi-host-range *P_{tac}* expression vector. *Gene* **1986**, *48*, 119-131.
- ¹² Miller, J. H. *Experiments in Molecular Genetics*; Cold Spring Harbor Laboratory: Plainview, NY, 1972.
- ¹³ Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
- ¹⁴ Silhavy, T. J.; Berman, M. L.; Enquist, L. W. *Experiments with Gene Fusions*; Cold Spring Harbor Laboratory: Plainview, NY, 1984.
- ¹⁵ Pitcher, D. G.; Saunders, N. A.; Owen, R. J. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett. Appl. Microbiol.* **1989**, *8*, 151-156.
- ¹⁶ Miller, J. H. *A Short Course in Bacterial Genetics*; Cold Spring Harbor Laboratory: Plainview, NY, 1992.

- ¹⁷ Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248-254.
- ¹⁸ (a) Barbas III, C. F.; Wang, Y.-F.; Wong, C.-H. *J. Am. Chem. Soc.* **1990**, *112*, 2013. (b) Chen, L.; Dumas, D. P.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 741.
- ¹⁹ Toraya, T.; Krodell, E.; Mildvan, A. S.; Abeles, R. H. *Biochemistry* **1979**, *18*, 417.
- ²⁰ (a) Raj, K. C.; Ingram, L. O.; Maupin-Furlow, J. A. *Arch. Microbiol.* **2001**, *176*, 443. (b) Raj, K. C.; Talarico, L. A.; Ingram, L. O.; Maupin-Furlow, J. A. *Appl. Environ. Microbiol.* **2002**, *68*, 2869. (c) Conway, T.; Osman, Y. A.; Konnan, J. I.; Hoffmann, E. M.; Ingram, L. O. *J. Bacteriol.* **1987**, *169*, 949.
- ²¹ Tsou, A. Y.; Ransom, S. C.; Gerlt, J. A. Buechter, D. D.; Babbitt, P. C.; Kenyon, G. L. *Biochemistry* **1990**, *29*, 9856.
- ²² Keshav, K. F.; Yomano, L. P.; An, H. J.; Ingram, L. O. *J. Bacteriol.* **1990**, *172*, 2491.
- ²³ (a) Tong, I. T.; Liao, H. H.; Cameron, D. C. *Appl. Environ. Microbiol.* **1991**, *57*, 3541. (b) Johnson, E. A.; Lin, E. C. C. *J. Bacteriol.* **1987**, *169*, 2050.
- ²⁴ Harris, E. L. V.; Angal, S. In *Protein Purification Methods: A Practical Approach*; Oxford University Press: Oxford, New York, Tokyo, 1989.
- ²⁵ Gijzen, H. J. M.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 8422.
- ²⁶ Synder, J.; Serianni, A. S. *Carbohydr. Res.* **1991**, *210*, 21.
- ²⁷ Demain, A. L.; Davies, J. E. *Manual of Industrial Microbiology and Biotechnology*; ASM Press: NY, 1986.
- ²⁸ Siegert, P.; Pohl, M.; Kneen, M. M.; Pogozheva, I. D.; Kenyon, G. L.; McLeish, M. J. In *Thiamine: Catalytic mechanisms in normal and disease states*; Jordan, F.; Patel, M. S. Eds.; Marcel Dekker, Inc.: NY, 2004; p275.
- ²⁹ Conway, T.; Osman, Y. A.; Konnan, J. I.; Hoffmann, E. M.; Ingram, L. O. *J. Bacteriol.* **1987**, *169*, 949.
- ³⁰ (a) Stemmer, W. P. C. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747-10751. (b) Stemmer, W. P. C. Rapid evolution of a protein *in vitro* by DNA shuffling.

Nature **1994**, 370, 389-391.

³¹(a) Zhao, H.; Arnold, F. H. Optimization of shuffling for high fidelity recombination. *Nucleic Acids Res.* **1997**, 25, 1307-1308. (b) Arnold, F. H.; Georgiou, G. *Directed Evolution Library Creation: Methods and Protocols*; Humana Press: Totowa, NJ, 2003.

³²Joern, J. M.; Meinhold, P.; Arnold, F. H. *J. Mol. Biol.* **2002**, 316, 643.

³³Niu, W. *PhD Thesis*, Michigan State University, 2004.

³⁴Cherepanov, P. P.; Wackernagel, W. *Gene* **1995**, 158, 9.

³⁵Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 6640.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03062 4476