

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Aldolase-catalyzed synthesis of chiral organofluorines

Permalink

<https://escholarship.org/uc/item/00v9550q>

Author

Fang, Jason

Publication Date

2022

Peer reviewed|Thesis/dissertation

Aldolase-catalyzed synthesis of chiral organofluorines

by

Jason Fang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Michelle C. Y. Chang, Chair

Professor Matthew B. Francis

Professor Wenjun Zhang

Spring 2022

Aldolase-catalyzed synthesis of chiral organofluorines

© 2022

by Jason Fang

Abstract

Aldolase-catalyzed synthesis of chiral organofluorines

by

Jason Fang

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Fluorine is a critically important element for the design of safe and effective drugs in the modern pharmaceutical industry. As the field of synthetic methodology for fluorination of organic compounds has flourished, biocatalysis has also been rapidly adopted as a means to achieve high reaction selectivity and process sustainability. However, the areas of biocatalysis and fluorine chemistry have rarely intersected directly, owing to the rarity of fluorine-processing enzymes in Nature. We envision that this methodology gap can be bridged by the action of carbon-carbon bond forming enzymes upon non-native fluorinated substrates, thus transforming simple organofluorine building blocks into more complex organofluorines. In this dissertation, we describe the application of enzymatic aldol addition to fluoro-aldol reactions that furnish value-added organofluorine products bearing fluorine stereocenters.

Our development of a novel platform for biocatalytic organofluorine synthesis began with the discovery that Type II pyruvate aldolases of the HpcH family use fluoropyruvate as an alternative nucleophilic substrate. After studying the stereoselectivity, kinetics, and mechanism of the fluoropyruvate aldol reaction, the afforded products were converted to diversely functionalized organofluorines through downstream reactions. Next, the HpcH system was expanded beyond fluoropyruvate to generalized β -fluoro- α -ketoacids as nucleophiles. The investigation of these challenging substrates, aided by rational active site engineering, resulted in the first synthesis of tertiary fluorides via biocatalytic C-C bond formation. Usefulness of the HpcH platform was demonstrated with preparative syntheses of products relevant to drug fragments and bioactive natural products. Finally, we studied Type I aldolases with the unique ability to use simple aldehydes and ketones as nucleophiles. Their activity on fluorinated ketones may enable the efficient synthesis of fluorinated sugars, which are applicable to pharmaceuticals and chemical biology. Taken together, this dissertation delineates a selective and sustainable biocatalytic approach to access the unique properties conferred by fluorination, which may have broad implications for improved process design towards important organofluorines.

Table of Contents

<i>Table of Contents</i>	i
<i>List of Figures</i>	iii
<i>List of Abbreviations</i>	vii
<i>Acknowledgments</i>	xi

Chapter 1: Introduction

1.1 <i>Biocatalysis in the pharmaceutical industry</i>	2
1.2 <i>Organofluorines: properties, synthesis, and biosynthesis</i>	4
1.3 <i>The aldol reaction in synthesis and in nature</i>	7
1.4 <i>Thesis organization</i>	10
1.5 <i>References</i>	10

Chapter 2: Developing a pyruvate aldolase platform for the synthesis of chiral organofluorines

2.1 <i>Introduction</i>	18
2.2 <i>Materials and methods</i>	20
2.3 <i>Results and discussion</i>	29
2.4 <i>Conclusion</i>	57
2.5 <i>References</i>	57

Chapter 3: Expanding the substrate scope of HpCH aldolase for construction of tertiary fluorides

3.1 <i>Introduction</i>	62
3.2 <i>Materials and methods</i>	63
3.3 <i>Results and discussion</i>	76
3.4 <i>Conclusion</i>	96
3.5 <i>References</i>	96

Chapter 4: Investigating the aldolase-catalyzed synthesis of fluorinated sugars from fluoroketones

<i>4.1 Introduction</i>	102
<i>4.2 Materials and methods</i>	103
<i>4.3 Results and discussion</i>	109
<i>4.4 Conclusion</i>	118
<i>4.5 References</i>	119

Appendices

<i>Appendix 1: Summary of miscellaneous experiments</i>	123
<i>Appendix 2: Reaction tables for ^{19}F NMR assays</i>	131
<i>Appendix 3: Plasmids, oligonucleotides, and gBlocks</i>	145

List of Figures

Chapter 1

<i>Figure 1.1</i>	<i>Commonly used enzymes in biocatalysis</i>	2
<i>Figure 1.2</i>	<i>Examples of biocatalysis in the pharmaceutical industry</i>	4
<i>Figure 1.3</i>	<i>Properties of fluorine and organofluorines</i>	5
<i>Figure 1.4</i>	<i>Reactions and reagents for chemical fluorination</i>	6
<i>Figure 1.5</i>	<i>Fluorine-containing pharmaceuticals</i>	6
<i>Figure 1.6</i>	<i>Nucleophilic fluorination catalyzed by the fluorinase</i>	7
<i>Figure 1.7</i>	<i>Stereoselective aldol reaction methodology</i>	8
<i>Figure 1.8</i>	<i>Mechanism and classification of aldolases</i>	9

Chapter 2

<i>Figure 2.1</i>	<i>Methodology for the stereoselective fluoro-aldol reaction</i>	19
<i>Figure 2.2</i>	<i>Enzyme information for Hpch aldolases</i>	31
<i>Figure 2.3</i>	<i>SDS-PAGE analysis of Hpch aldolases</i>	31
<i>Figure 2.4</i>	<i>¹⁹F NMR analysis and buffer screening for enzymatic fluoropyruvate addition</i>	32
<i>Figure 2.5</i>	<i>Acceptor scope of Hpch aldolases for fluoropyruvate addition</i>	32
<i>Figure 2.6</i>	<i>¹⁹F NMR data for fluoropyruvate-derived α-fluoroacids</i>	33
<i>Figure 2.7</i>	<i>LC-MS data for β-fluoro-α-ketoacids and α-fluoroacids</i>	33
<i>Figure 2.8</i>	<i>Effect of enzyme loading on Hpch-catalyzed fluoropyruvate addition</i>	34
<i>Figure 2.9</i>	<i>Determination of product relative stereochemistry by oxidation with nitric acid</i>	37
<i>Figure 2.10</i>	<i>Determination of product absolute stereochemistry by reduction with lactate dehydrogenase</i>	37
<i>Figure 2.11</i>	<i>NMR coupling constants supporting the absolute stereochemistry of fluoropyruvate-derived products</i>	38
<i>Figure 2.12</i>	<i>Kinetics of the Hpch-catalyzed aldol additions of pyruvate and fluoropyruvate</i>	39

<i>Figure 2.13</i>	<i>DFT calculations on a model of the fluoroenolate intermediate at the active site of HpCH</i>	40
<i>Figure 2.14</i>	<i>Mechanism and stereoselectivity of HpCH aldolases</i>	40
<i>Figure 2.15</i>	<i>Bioinformatic analysis of active site residue frequencies in the HpCH family</i>	41
<i>Figure 2.16</i>	<i>Effect of phenylalanine mutations on diastereoselectivity of fluoropyruvate aldol addition</i>	41
<i>Figure 2.17</i>	<i>Chemoenzymatic synthesis and isolation of organofluorine products</i>	41
<i>Figure 2.18</i>	<i>LC-MS data for isolated organofluorine products</i>	43
<i>Figure 2.19</i>	<i>NMR spectra of (2S,3R,4R)-2-fluoro-3,4,5-trihydropentanoate</i>	44
<i>Figure 2.20</i>	<i>NMR spectra of (2S,3S,4R)-2-fluoro-3,4,5-trihydropentanoate</i>	45
<i>Figure 2.21</i>	<i>NMR spectra of (2S,3S,4S)-2-fluoro-3,4,5-trihydropentanoate</i>	46
<i>Figure 2.22</i>	<i>NMR spectra of (2S,3R,4R)-2-fluoro-3,4-dihydropentanoate</i>	47
<i>Figure 2.23</i>	<i>NMR spectra of (2S,3S,4R)-2-fluoro-3,4-dihydropentanoate</i>	48
<i>Figure 2.24</i>	<i>NMR spectra of (2S,3S,4S)-2-fluoro-3,4-dihydropentanoate</i>	49
<i>Figure 2.25</i>	<i>NMR spectra of methyl (2S,3S)-2-fluoro-3,4-O-isopropylidene-3,4-dihydroxybutanoate</i>	50
<i>Figure 2.26</i>	<i>NMR spectra of methyl (2S,3S)-2-fluoro-3-hydroxy-4,4-dimethoxybutanoate</i>	51
<i>Figure 2.27</i>	<i>NMR spectra of dimethyl (2S,3R)-2-fluoro-3-hydroxysuccinate</i>	52
<i>Figure 2.28</i>	<i>NMR spectra of dimethyl (2S,3S)-2-fluoro-3-hydroxysuccinate</i>	53
<i>Figure 2.29</i>	<i>Schematic of the HpCH platform for organofluorine synthesis with downstream reactions</i>	54
<i>Figure 2.30</i>	<i>LC-MS data for β-fluoro-α-hydroxyacids and β-fluoro-α-aminoacids</i>	55
<i>Figure 2.31</i>	<i>Preparation of fluorinated sugars by chemical reduction</i>	56
<i>Figure 2.32</i>	<i>LC-MS data for fluorinated sugars</i>	56

Chapter 3

Figure 3.1	Motivation and methodology for synthesis of secondary and tertiary fluoride stereocenters	63
Figure 3.2	Synthesis of β -fluoro- α -ketoacid donor substrates	77
Figure 3.3	SDS-PAGE analysis of EchPcH mutants and MBP-HpCh	78
Figure 3.4	Location of hydrophobic residues targeted for mutation in the active site of EchPcH	78
Figure 3.5	Screening of aldolase variants for aldol addition of β -fluoro- α -ketoacids to formaldehyde	79
Figure 3.6	Verification of kinetic resolution of β -fluoro- α -ketobutyrate by reduction with lactate dehydrogenase	80
Figure 3.7	Active site schematic of EchPcH and rationalization of beneficial mutations	81
Figure 3.8	Acceptor scope of EchPcH for addition of β -fluoro- α -ketoacids	83
Figure 3.9	^{19}F NMR data for α -fluoroacids derived from aldol addition of β -fluoro- α -ketoacids	84
Figure 3.10	LC-MS data for α -fluoroacids	85
Figure 3.11	Isolated organofluorine products relevant to bioactive compounds and pharmaceuticals	87
Figure 3.12	Chemoenzymatic asymmetric synthesis of organofluorines and synthesis of corresponding racemic standards	87
Figure 3.13	NMR spectra of methyl (S)-2-fluoro-3-hydroxypropanoate	88
Figure 3.14	NMR spectra of methyl (S)-2-fluoro-3-hydroxy-2-methylpropanoate	89
Figure 3.15	NMR spectra of methyl (S)-2-fluoro-2-(hydroxymethyl)butanoate	90
Figure 3.16	NMR spectra of ethyl (2S,3RS)-2-fluoro-3-hydroxy-3-phenylpropanoate	91
Figure 3.17	NMR spectra of ethyl (2S,3RS)-2-fluoro-3-hydroxy-2-methyl-3-(2-pyridyl)propanoate	92
Figure 3.18	NMR spectra of ammonium (2S,3S,4R)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate	93
Figure 3.19	NMR spectra of ammonium (2S,3RS,4S)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate	94
Figure 3.20	Mosher ester analysis of organofluorine products	95

Chapter 4

Figure 4.1	Importance of fluorinated sugars and the native reactions of privileged aldolases DERA and FSA	103
Figure 4.2	Previous synthetic routes towards FHA	110
Figure 4.3	Novel synthesis of FHA by selective enzymatic oxidation of FPDO with glycerol dehydrogenase	111
Figure 4.4	SDS-PAGE analysis of glycerol dehydrogenase GldA	111
Figure 4.5	Optimization of conversion for enzymatic oxidation of FPDO to FHA	112
Figure 4.6	Initial rate measurements of GldA-catalyzed oxidation	113
Figure 4.7	Aldol addition assay and two-step derivatization protocol for LC-MS analysis	113
Figure 4.8	LC-MS data for DERA- and FSA-catalyzed aldol additions	114
Figure 4.9	^{19}F NMR spectra for DERA-catalyzed aldol addition of MFA and DFA	115
Figure 4.10	LC-MS assessment of DERA Cys-47 mutant activity in MFA and DFA aldol additions	116
Figure 4.11	Location of Cys-47 and hydrophobic residues targeted for mutation in the active site of DERA	117
Figure 4.12	Qualitative ^{19}F NMR assessment of DERA mutant activity in MFA and DFA aldol additions	117
Figure 4.13	SDS-PAGE analysis of DERA variants, fructose-1,6-bisphosphate aldolase, and triose phosphate isomerase	118
Figure 4.14	Quantitative ^{19}F NMR comparison of DERA variant activity in MFA and DFA aldol additions	118

List of Abbreviations

2FP	2-fluoropropionate
2KA	2-ketoadipate
2PCA	2-pyridinecarboxaldehyde
4HB	4-hydroxybutyrate
ACE	acetone
ADH	alcohol dehydrogenase
ALD	acetaldehyde
ATP	adenosine triphosphate
bp/kb	base pair / kilobase
BSA	bovine serum albumin
CAPS	<i>N</i> -cyclohexyl-3-aminopropanesulfonic acid
Car	carboxylic acid reductase
CHES	<i>N</i> -cyclohexyl-2-aminoethanesulfonic acid
DAST	diethylaminosulfur trifluoride
DERA	deoxyribose-5-phosphate aldolase
DFA	1,3-difluoroacetone
D-GA	D-glyceraldehyde
DHA	1,3-dihydroxyacetone
DHAP	1,3-dihydroxyacetone phosphate
DHB	3,4-dihydroxybutyrate
DHDPS	dihydrodipicolinate synthase
DIBAL-H	diisobutylaluminum hydride
DIPA	diisopropylamine
D-LA	D-lactaldehyde
DMA	dimethoxyacetaldehyde
DMAP	4-dimethylaminopyridine
dNTP	deoxynucleotides
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
FAL	formaldehyde

FDA	5'-fluoro-5'-deoxyadenosine
FDHB	2-fluoro-3,4-dihydroxybutyrate
FDHOP	3-fluoro-4,5-dihydroxy-2-oxopentanoate
FDHP	2-fluoro-3,4-dihydroxypentanoate
FHA	1-fluoro-3-hydroxyacetone
FHAP	1-fluoro-3-hydroxyacetone phosphate
FHS	2-fluoro-3-hydroxysuccinate
FKB	β -fluoro- α -ketobutyrate
FKMV	β -fluoro- α -keto- γ -methylvalerate
FKPB	β -fluoro- α -keto- γ -phenylbutyrate
FKV	β -fluoro- α -ketovalerate
FP	fluoropyruvate
FPA	2-fluoro-2-phenylacetate
FPDO	3-fluoro-1,2-propanediol
FPP	fluoro(phenyl)pyruvate
FSA	fructose-6-phosphate aldolase
FTHP	2-fluoro-3,4,5-trihydroxypentanoate
G3P	D-glyceraldehyde-3-phosphate
GA (HAL)	glycolaldehyde (hydroxyacetaldehyde)
GX	glyoxylic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEPPS (EPPS)	4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid
HILIC	hydrophilic interaction chromatography
HKHD	4-hydroxy-2-oxoheptane-1,7-dioate
HPLC	high-performance liquid chromatography
IBX	2-iodoxybenzoic acid
IPTG	isopropyl β -D-1-thiogalactopyranoside
KCM	KCl, CaCl ₂ , and MgCl ₂
KDPG	2-keto-3-deoxy-6-phosphogluconate
LB	lysogeny (Luria-Bertani) broth
LC-MS	liquid chromatography – mass spectrometry
LDA	lithium diisopropylamide

LDH	lactate dehydrogenase
L-GA	L-glyceraldehyde
L-LA	L-lactaldehyde
MBP	maltose binding protein
MCS	multiple cloning site
MDH	malate dehydrogenase
MFA	fluoroacetone (monofluoroacetone)
MHA	hydroxyacetone (monohydroxyacetone)
MTPA	methoxy(trifluoromethyl)phenylacetic acid (Mosher's acid)
MVK	methyl vinyl ketone
NAD(P)+	nicotinamide adenine dinucleotide (phosphate)
NAD(P)H	nicotinamide adenine dinucleotide (phosphate) reduced
NFSI	<i>N</i> -fluorobenzenesulfonimide
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
OBHA	<i>O</i> -benzylhydroxylamine
PA	phenylacetate
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEI	polyethyleneimine
PET	positron emission tomography
PLP	pyridoxal phosphate
PMSF	phenylmethylsulfonyl fluoride
Pro	propionaldehyde
RP	reversed-phase
SAM	<i>S</i> -adenosyl methionine
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SMBA	(<i>S</i>)- α -methylbenzylamine
TAPS	<i>N</i> -tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TB	Terrific broth
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)phosphine

TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TLC	thin-layer chromatography
TMS-Cl	trimethylsilyl chloride
TPP	thiamine pyrophosphate

Acknowledgments

First and foremost, I would like to thank Prof. Michelle Chang for taking me into her lab and providing the place I called home for six years. Michelle's extraordinary scientific vision and mentorship style have created a wonderful space to perform and share science. I particularly appreciate Michelle granting me the freedom to learn advanced organic synthesis, a long-standing goal of mine that became reality in the one trusty fume hood of 124 Lewis. Additionally, Michelle's open-mindedness and confidence in her students often counterbalanced my tendencies for hasty judgments, which forged me into a better scientist. Neither will Michelle's generosity and culinary expertise in food-related outings ever be forgotten. I would also like to thank the members of my qualifying exam and thesis committees Prof. Matt Francis, Prof. Anne Baranger, Prof. Michael Marletta, and Prof. Wenjun Zhang for their mentorship and support.

The benefit of a small lab group is getting to interact with everyone. I am especially grateful to the halogen/alkyne subgroup: Chia-I Lin, Ningkun Wang, Omer Ad, Jorge Marchand, Jon McMurry, Monica Neugebauer, and Sasilada Sirirunguang. They were collectively my mentors in the early years, and were instrumental in getting my project off the ground. Hongjun Dong, Joe Gallagher, Kersh Thevasundaram, and Vivian Yu were also role models who inspired me with their work ethic and wisdom. I wish the best to continuing postdocs and PhD students Elizabeth Stone, Ed Koleski, Eli Kissman, Mira Liu, Doug Millar, Max Sosa, and Andrew Whitten, who are all leading scientific projects in exciting and thriving directions. Special thanks to Max, Mira, and Ed, who through their frequent company became my confidants for countless matters both in and out of science, and made Covid-19 lockdowns far more enjoyable. Outside of the Chang lab, fellow chemistry PhD students Hikaru Mizuno and Eugene Kim were my kind housemates for four excellent years and I thank them as well.

I thank Diptarka Hait (Prof. Martin Head-Gordon group), whose assistance with computational chemistry resulted in the most significant mechanistic discovery of my thesis. Additionally, I have been fortunate to mentor three talented scientists: Oxford exchange student Laura Turner, whose trans-Atlantic journey was followed by mastery of lab techniques at lightning pace; undergraduate student Matthew Kojima, whose tenacity in the face of Covid-19 restrictions on his research has convinced me of his future success; and graduate student Gabby Dolgonos, whose rotation project was highly productive and exciting to work on together.

Local restaurants are the best part of Berkeley, and I dined out for the majority of meals in graduate school. I thank these businesses for providing my physical sustenance and hosting our group lunches: Aki's, Alley Kitchens, Bears Ramen House, Boba Ninja, Bongo Burger, Camille's, Celia's, Kimchi Garden, Lotus House, Mandarin House, Momo Masala, Muracci's, Northside Café, Shihlin Taiwan Street Snacks, Steve's Bar-B-Que, Super Duper Burgers, Taco Sinaloa, TC Garden, Tea Press Café, Thai Basil, Thai Noodle, TP Tea, and Yi Fang Fruit Tea.

Graduate school has been a roller coaster of experiences and emotions, and I could not have made it to the end without the moral support of friends and family. I am very grateful for Duyen Pham, Prasiddha Gurung, and Vivian Lin, who by virtue of not being chemistry PhD students provided fresh perspectives that carried me through some of the darkest times. Special thanks to my partner, Karen Shi, who brightens each day and inspires me to take on new challenges with her compassion and perseverance. Finally, special thanks to my family—my sister Michelle, my mom Hylin, and my dad Lian—for their unwavering support of not just my graduate studies, but also everything that led up to it and anything that may follow.

Chapter 1: *Introduction*

1.1. Biocatalysis in the pharmaceutical industry

In the last few decades, biocatalysis has risen from relative obscurity to the forefront of chemical synthesis. Biocatalysis entails the repurposing of enzymes, proteins that sustain life by catalyzing metabolic reactions, for the production of compounds desirable to human society. The proliferation of enzymes as part of the synthetic chemist's toolbox has greatly impacted the pharmaceutical sector, by allowing drugs or drug intermediates to be synthesized with high efficiency and exquisite regioselectivity or stereoselectivity [1-3]. Enzymes have the native capacity to function in mild conditions in aqueous solvent, enabling process chemistry that is more sustainable than chemocatalysis (transition metal catalysis and organocatalysis). Synthetic routes can now be streamlined with the integration of biocatalytic and chemocatalytic steps [4], or more ambitiously with purely biocatalytic multi-enzyme cascades [5,6]. Such approaches are chosen appropriately to improve efficiency benchmarks such as space-time productivity (grams of product per liter per hour) and sustainability benchmarks such as process mass intensity (grams of product per gram of input material) [7,8].

The functional group transformations that could be biocatalyzed have historically been limited to simple hydrolysis or acyl transfer of esters and amides. For example, since the 1960s lipases and proteases have been employed in laundry detergents, and used by organic chemists to catalyze kinetic resolutions and desymmetrizations for enantioenriched products [9,10]. Despite the importance of these specialized applications, difficulties in obtaining proteins from natural sources and modifying proteins to tune their properties resulted in a lack of reaction diversity. Starting in the 1990s, advances in gene synthesis, molecular biology, and bioinformatics have increasingly rendered trivial the production of almost any desired enzyme [11]. Furthermore, the advent of directed evolution has made possible the rapid optimization of enzymes via generating and screening large libraries of mutations [12,13]. These developments, culminating in the recognition of directed evolution with the 2018 Nobel Prize in Chemistry, spearheaded the entry of diverse enzyme types into the mainstream purview of biocatalysis. A broad range of functional groups are now interconvertible with enzymes (*Figure 1.1*). Commonly used enzyme classes include

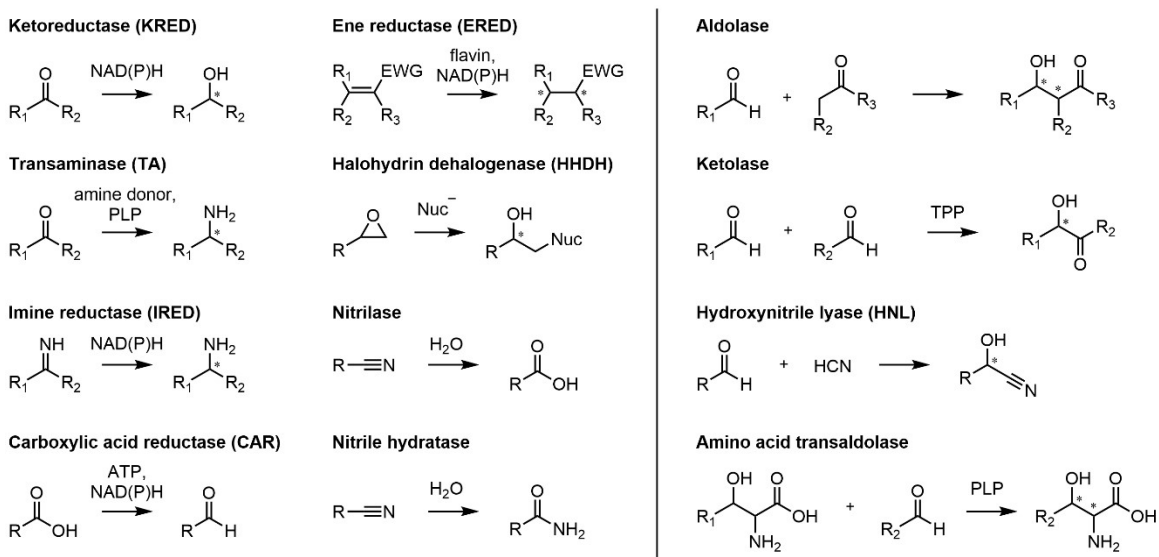


Figure 1.1. Enzyme classes that have emerged as sources of useful biocatalysts. Potential new stereocenters are indicated with an asterisk. Left panel: common functional group transformations; Right panel: carbon-carbon bond-forming transformations.

ketoreductase (KRED), transaminase (TA), imine reductase (IRED), ene reductase (ERED), carboxylic acid reductase (CAR), halohydrin dehalogenase (HHDH), nitrilase, nitrile hydratase, and others [14,15]. Additionally, C-C bond forming enzymes such as aldolase, ketolase, hydroxynitrile lyase (HNL), and amino acid transaldolase can build up complexity from simpler fragments [16-18].

Illustrative of the success of new biocatalytic modes in pharmaceutical manufacturing, the syntheses of asthma medication montelukast (Singulair) and anti-diabetic medication sitagliptin (Januvia) were accomplished with evolved enzymes (*Figure 1.2A*). In the preparation of a chiral alcohol intermediate towards montelukast, a ketoreductase was engineered to reduce a remarkably hydrophobic ketone precursor [19]. The enzymatic method produced far less waste than the previous process, which relied upon stoichiometric amounts of a moisture-sensitive borane reagent. Similarly, the installation of a chiral amine as the last step towards sitagliptin was catalyzed by an extensively engineered transaminase [20]. The single enzymatic step replaced a chemocatalytic sequence that required a rhodium catalyst, high pressures of hydrogen, and multiple recrystallization steps. In the case of both drugs, the biocatalytic route benefited from superior enantioselectivity and yield.

Competing efforts to improve the synthesis of blockbuster cholesterol-lowering drug atorvastatin (Lipitor) have witnessed a confluence of multiple enzymatic approaches, aimed at preparing the 7-amino-3,5-dihydroxyheptanoate side chain bearing two chiral centers (*Figure 1.2B*). One route accessed the three substitutions on the alkyl chain with two applications of ketoreductase and one dehalogenase-catalyzed nitrile installation [21]. A competing strategy used engineered nitrilase to perform desymmetrization of a dinitrile precursor, thus providing a hydroxy stereocenter and masked amine simultaneously [22]. Yet a third route used an aldolase to rapidly assemble the desired 7-amino-3,5-dihydroxyheptanoate product from inexpensive acetaldehyde and chloroacetaldehyde [23]. This strategic progression demonstrates the point to which biocatalysis has matured, where not only can enzymes replace chemical steps, but newer and potentially revolutionary enzymes can replace established ones.

To fully leverage the advantages that biocatalysis can bring to a green economy, the pharmaceutical industry has recently begun to pursue the total enzymatic synthesis of target molecules. As enzymes have high specificity for their intended transformations, multiple enzymes can often be used together in a one-pot cascade process. In proposing a total enzymatic synthesis, one can apply the concept of bioretrosynthesis [4,24]. This entails an extension of retrosynthesis for identifying disconnections in a target molecule, analyzing it as tractable fragments that can be assembled enzymatically. Furthermore, if the proposed pathway enzymes are evolved in the reverse direction, then every directed evolution library can be screened by detection of the final product, saving time and labor that would otherwise be spent on assay development [25]. This strategy was applied to the synthesis of anti-HIV drug islatravir in three steps from 2-ethynylglycerol, acetaldehyde, and 2-fluoroadenine (*Figure 1.2C*) [26]. Overall, the process utilized five evolved enzymes and four auxiliary enzymes in a single aqueous solution. No intermediates were isolated, and the final product automatically crystallized from the reaction mixture. Following the success of this total enzymatic synthesis, a related enzyme cascade was devised for molnupiravir [27], an oral antiviral for treatment of Covid-19, reinforcing the idea that biocatalysis can contribute efficient and sustainable solutions to pressing matters of human health.

Although the diversity of biocatalysis in organic synthesis have grown significantly, it still falls short of that of chemocatalysis, thus a partnership between the two transformative modes is

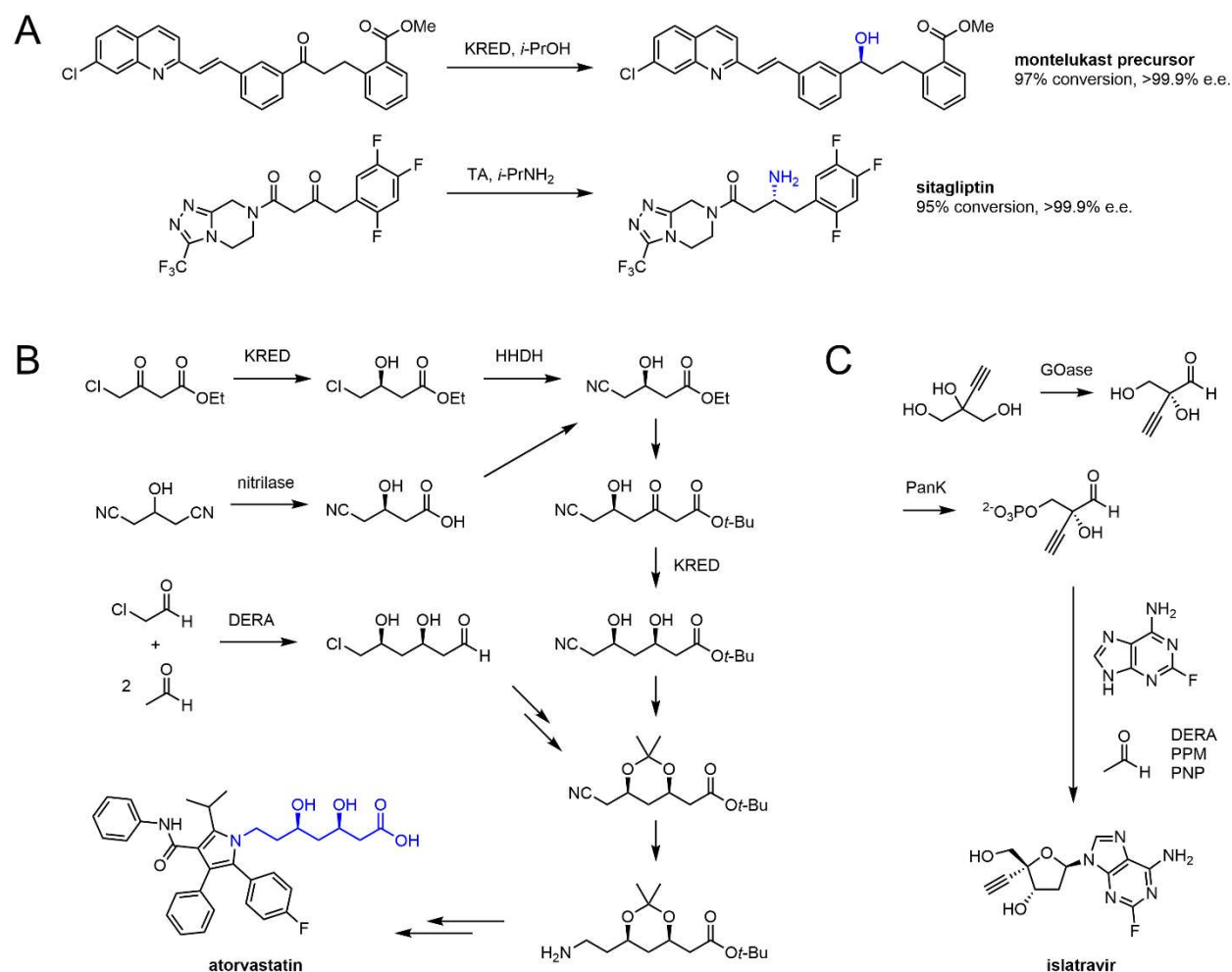


Figure 1.2. Examples of industrial biocatalysis in the pharmaceutical sector. A) Asymmetric ketone reduction and transamination, providing high yield and stereopurity towards montelukast and sitagliptin. B) Network of viable enzymatic routes towards the side chain of atorvastatin. C) Total enzymatic synthesis of islatravir in three steps, using five primary enzymes and four auxiliary enzymes (not shown). DERA: deoxyribose-5-phosphate aldolase, GOase: galactose oxidase, PanK: pantothenate kinase, PPM: phosphopentomutase, PNP: purine nucleoside phosphorylase.

often the most practical [4]. Further research in this area will facilitate the phase-out of chemical processes that burden the environment and the feasibility of more total enzymatic processes. Because directed evolution can transform a poor biocatalyst into an outstanding one, it is crucial to explore both new and known enzymes for potentially unprecedented or unknown promiscuous activities, even if the activity level is very low at first. As the role of biocatalysis in pharmaceutical manufacturing continues to increase, we seek to utilize these strategies to expand the possibilities for robust and scalable drug synthesis.

1.2. Organofluorines: properties, synthesis, and biosynthesis

Fluorine has unique atomic properties that make it a critical design element in the development of safe and effective pharmaceuticals (Figure 1.3). An estimated 20-25% of drugs contain at least one fluorine atom, and this figure increases further to 30-35% for recently approved drugs. [28,29]. As fluorine is the most electronegative element, the C-F bond is highly polarized while at the same time serving as a sterically conservative bioisostere of a C-H or C-OH bond [30]. Strategic

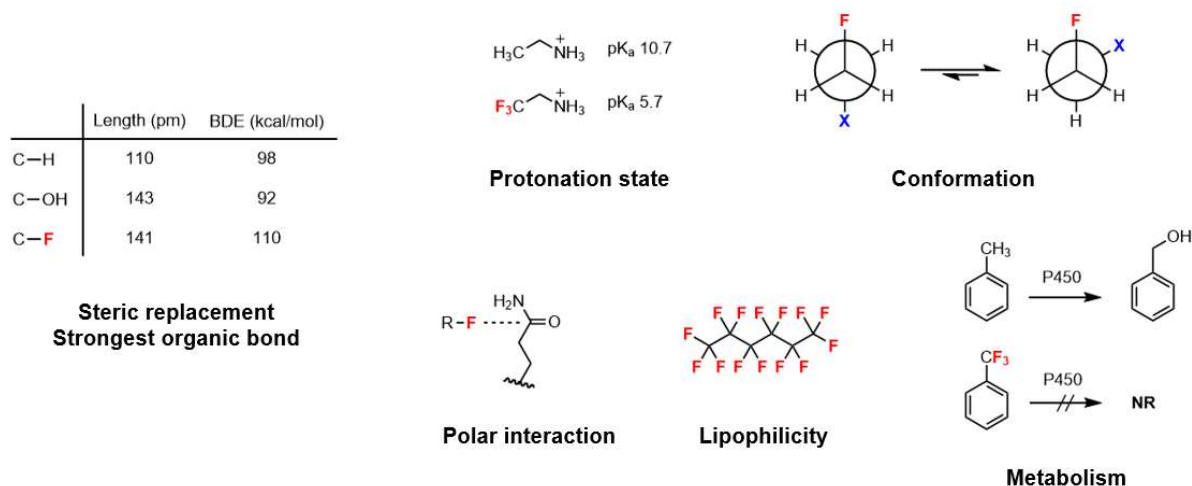


Figure 1.3. Properties of organofluorines used in the design of safe and effective pharmaceuticals. Fluorine is introduced to drug structures to modulate pharmacodynamics, pharmacokinetics, and drug target interactions.

placement of fluorine atoms allows fine tuning of the pK_a of nearby acidic or basic functional groups and modulation of the overall molecule lipophilicity, with important consequences for bioavailability and pharmacodynamic profile [28,31,32].

The biological effects of fluorine are flexible and diverse. Often, fluorine is substituted at a metabolically labile position to improve drug stability, since the high dissociation energy of a C-F bond makes fluorine a generally poor leaving group. On the other hand, fluorine may be an excellent leaving group if the β -carbon can adopt anionic character, which is exploited by mechanism-based inhibitors like the antiviral agent trifluridine [28,31]. Fluorine atoms impose a strong conformational bias on bond angles and dihedral angles. An example of this phenomenon is the gauche effect, whereby hyperconjugation of nearby electron-rich σ -orbitals into electron-poor σ^* -antibonding orbitals leads to a gauche orientation of electronegative substituents. Hyperconjugation effects dictate the conformation and thus the bioactivity of fluorinated cyclic structures, such as medicinally important fluorinated sugars [33]. Lastly, in the context of drug binding at protein targets, C-F bonds are observed to form non-covalent multipolar interactions with the protein backbone or amino acid residues (usually at amide N-H or C=O bonds), which may be crucial for binding affinity or the mechanism of action [32].

Chemical methods for introducing fluorine into molecules can be broadly classified as nucleophilic or electrophilic (Figure 1.4). The field of fluorine chemistry has progressed significantly from an era when harsh and dangerous reagents such as fluorine gas, hydrofluoric acid, and compounds with O-F, S-F, or Xe-F bonds were unavoidable [34]. Rapid and robust methods for nucleophilic fluorination with metal fluoride salts have flourished, especially with regard to the synthesis of ^{18}F -radiopharmaceuticals [35,36]. Direct functional conversion from hydroxy to fluoro, or from carbonyl to difluoro, known as deoxofluorination, is safely performed with reagents such as DAST and PyFluor [37,38]. Meanwhile, electrophilic fluorination was revolutionized with the advent of bench-stable N-F reagents like NFSI and Selectfluor [39,40]. The electrophilic reagents can also function under radical-generating conditions that extends their transformative scope to fluorination at strong and unactivated C-H and C-C bonds [41-43]. Another recent development in general fluorination methodology is the polarity reversal

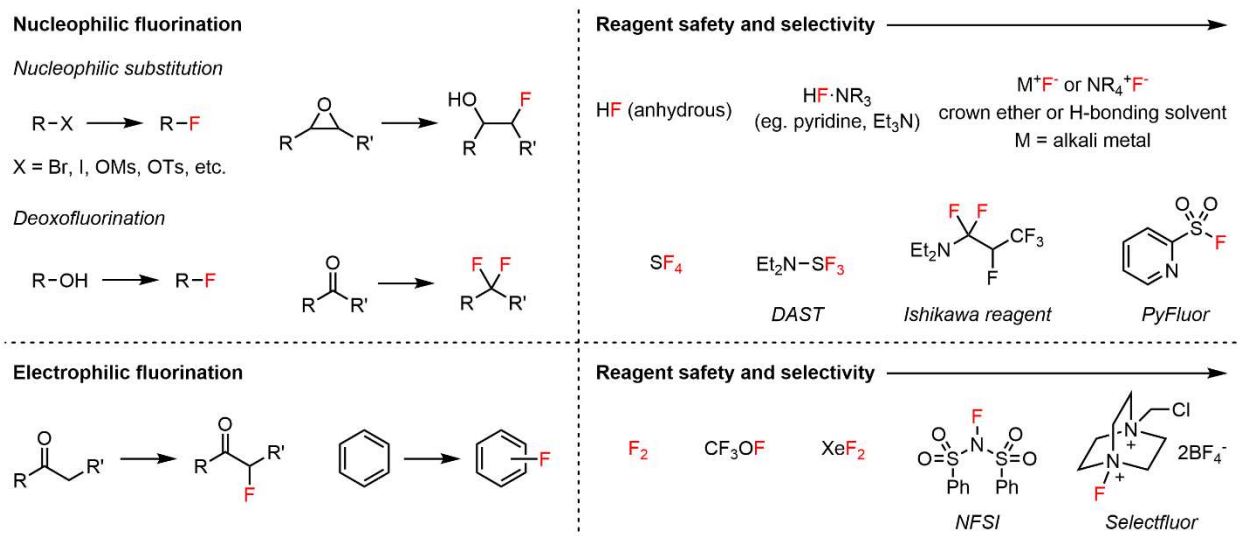


Figure 1.4. Modes of fluorination reactions, and progression of safety and selectivity in commercially available reagents for performing nucleophilic fluorination and electrophilic fluorination.

(umpolung) approach that uses simple fluoride salts to accomplish what would normally require an electrophilic reagent [44].

Aromatic fluorination is simple and widely employed in drug design as the only selectivity issue to address is regiochemistry, but aliphatic fluorination can carry the added complexity of stereochemistry (Figure 1.5). Wide availability of safe fluorinating reagents in the last ten years has led to a wealth of research in asymmetric fluorination [45], which was initially pioneered by combining electrophilic reagents with chiral metal-based catalysts or organocatalysts [46-49]. The concept was successfully applied later to asymmetric nucleophilic fluorination [50-52]. A third mode of construction for complex, chiral organofluorines is to elaborate them from simple organofluorine building blocks and a chiral catalyst. [53-58]. The starting organofluorines are achiral or racemic, and thus inexpensive or easily prepared. Chiral organofluorines can bear either secondary or tertiary fluoride stereocenters (confusingly also called tertiary or quaternary C-F stereocenters respectively—the terminology is fairly inconsistent in literature depending on authors' preference for halogen or carbon descriptors). Tertiary fluoride stereocenters, where the fluorine-bearing carbon is fully substituted, are more difficult to construct than secondary fluoride

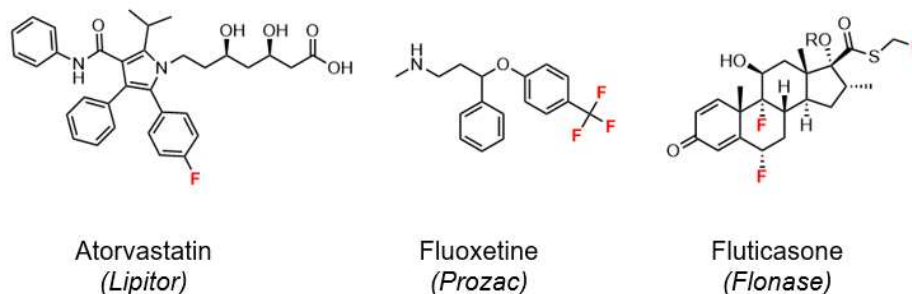


Figure 1.5. Examples of different patterns of fluorination in blockbuster drugs. The cholesterol-lowering drug atorvastatin contains an aryl fluoride. The antidepressant fluoxetine contains a trifluoromethyl group. The asthma and allergy medication fluticasone features primary, secondary, and tertiary aliphatic fluorides.

stereocenters and are thus represented in few approved drugs [59]. Thus, the synthesis of tertiary fluoride stereocenters is a research field of major current interest.

In stark contrast to the abundance of synthetic organofluorine compounds, organofluorines are nearly absent in the biosphere. No enzymes are known to perform electrophilic or radical fluorination. Only one enzyme, the fluorinase, performs nucleophilic fluorination but it requires a highly activated and particular substrate (*Figure 1.6*) [60]. The fluorinase uses fluoride ion to displace a methionine leaving group in S-adenosylmethionine (SAM), producing 5'-fluoro-5'-deoxyadenosine (FDA). Key to the reaction mechanism is desolvation of fluoride by the enzyme,

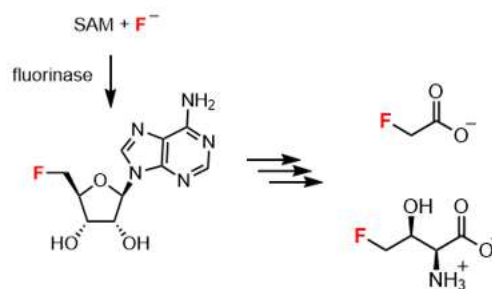


Figure 1.6. The fluorinase forms a C-F bond from inorganic fluoride in the first step of organofluorine biosynthesis pathways.

as the solvation energy of aqueous fluoride is an insurmountable barrier for uncatalyzed substitution. In the soil microbe *Streptomyces cattleya* and other organisms that produce fluorinated natural products, FDA is fed into downstream pathways towards toxic molecules such as fluoroacetate and fluorothreonine. Only about a dozen such fluorinated natural products, thought to be defense mechanisms for their producing organisms, have been isolated [61,62]. All of them are ultimately derived from the fluorinase reaction with the sole exception of nucleocidin, a 4'-fluoro nucleoside whose biosynthesis is not completely understood. Due to this lack of enzymes that natively transform organofluorines, the fields of fluorine chemistry and biocatalysis have hardly shared any overlap.

The gap between fluorine chemistry and biocatalysis may be bridged with the aforementioned “third mode” of asymmetric fluorination, in other words using enzymes to upgrade simple organofluorines. Depending on the exact nature of fluorinated versions of native substrates, these molecules may be excellent substrates, very poor or non-substrates, or even enzyme inhibitors [63]. When using an enzyme as a chiral catalyst to construct fluorine stereocenters from achiral or racemic organofluorines, the fluorine substitution is necessarily close to the reaction center, and strong electronic effects on the reaction rate or impediments to the desired reaction mechanism may arise. Identifying and overcoming these challenges is key to the continued evolution of fluorine chemistry into the next generation of safety, selectivity, and sustainability.

1.3. The aldol reaction in synthesis and in nature

The aldol reaction is a powerful method for forming a C-C bond between the nucleophilic α -carbon of a carbonyl compound and the electrophilic carbonyl carbon of another. It allows the union of two simple fragments to a more complex, valuable product with high atom economy, and the potential creation of up to two new stereocenters. In the last fifty years, many seminal advances made it feasible to selectively produce desired diastereomers or enantiomers of aldol products, as opposed to mixtures [64,65]. Stereoselective aldol reactions opened the door to the total synthesis of natural products replete with stereocenters. In pharmaceutical synthesis, where stereochemical impurities are not only wasteful but might be dangerous to patients, aldol methodology has contributed to increase the proportion of *sp*³-carbons in drugs or drug candidates, with corresponding positive impact on clinical success [66].

Reliable methods to configure the stereocenters of aldol products began to emerge in the 1970s, a century after the fundamental aldol reaction itself was discovered in the 1870s (Figure 1.7) [67]. Diastereoselectivity of aldol reactions between *syn* or *anti* products can be dictated by the *cis* or *trans* geometry of enolate intermediates activated by Lewis acids, as opposed to the traditional use of strong bases [68,69]. Enantioselectivity is introduced with the use of chiral auxiliaries or chiral catalysts. For example, oxazolidinones are widely available auxiliaries that are easily attached prior to an aldol reaction, and cleaved to various functional groups afterward [70]. But catalytic methods, which obviate these extra steps, are ultimately preferred over auxiliaries. Among numerous

developments in direct catalytic aldol reactions, organocatalysis based on enamines emerged as a powerhouse [71], a reputation solidified by the 2021 Nobel Prize in Chemistry.

Despite tremendous advances in methodology, the ability to perform stereoselective aldol reactions often relies upon steric exclusion interactions between bulky substrates and bulky auxiliaries or catalysts. This has caused considerable challenges with regard to small substrates. For example, the “acetate aldol problem” refers to a lack of β -hydroxyl stereoselectivity when using α -unsubstituted nucleophiles [70], and it extends to a “fluoroacetate aldol problem” due to the small size of a fluorine substitution [72,73]. The latter problem is further exacerbated by the need to control both the α -fluoro and β -hydroxyl stereocenters, and by the lower nucleophilicity of fluorinated carbons. Diastereoselective fluoro-aldol reactions have been achieved by roundabout measures, such as compensating for the lack of bulk at the α -carbon by incorporating extremely bulky groups elsewhere on the nucleophile [74-76]. The first enantioselective fluoro-aldol methodology was only reported in 2016, an organocatalytic strategy with fluoromalonic acid half thioesters that represents a pioneering achievement towards tackling an underdeveloped area of the aldol reaction [55].

Aldolases are enzymes that catalyze aldol reactions between nucleophile and electrophile substrates, which in biological literature are commonly referred to as donor and acceptor substrates respectively [16,17]. The native reaction of most aldolases is actually C-C bond cleavage in the retro-aldol direction, related to catabolism of sugars (e.g. the well-known fructose-1,6-bisphosphate aldolase in glycolysis) or to the detoxification of xenobiotics. However, since it was demonstrated in the 1990s that aldolases readily function in the reverse-physiological C-C bond forming direction often with high stereoselectivity, these enzymes have taken a foothold in the biocatalytic toolbox [77]. Aldolases are classified by whether they use the Type I or Type II mechanism for donor activation (Figure 1.8A). In the Type I mechanism, the donor substrate forms a Schiff base with a catalytic lysine residue, allowing α -deprotonation to an enamine intermediate that attacks a bound acceptor substrate. In the Type II mechanism, the donor substrate coordinates

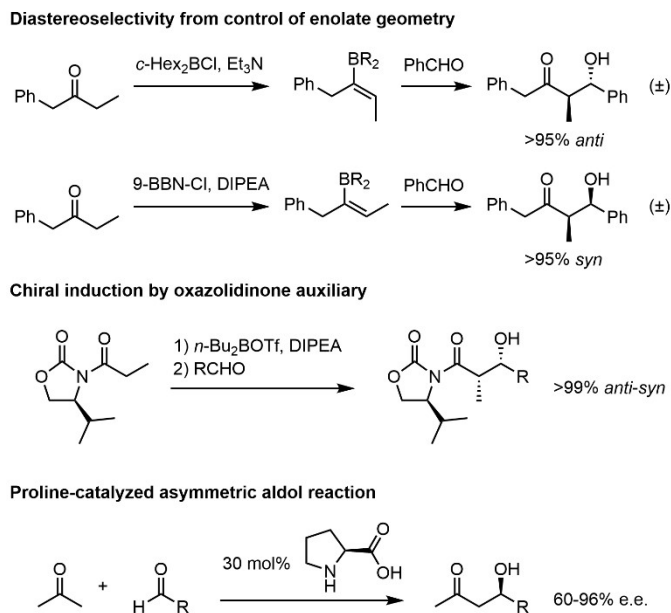
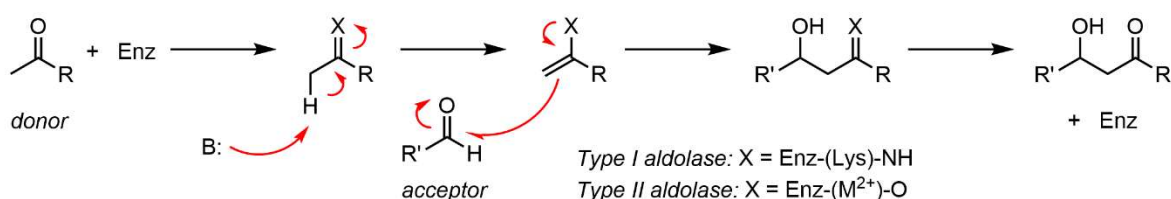


Figure 1.7. Landmark developments in synthetic methodology for stereoselective aldol reactions.

to a divalent metal ion bound by active site residues, and α -deprotonation gives rise to an enolate intermediate that attacks the acceptor. The Type I and II aldolase mechanisms are analogous to enamine-based organocatalysis and Lewis acid catalysis. However, the enzymatic mechanisms also surround the reaction center with a chiral three-dimensional architecture of precisely positioned amino acids, which is crucial to high turnover rates and exquisite stereoselectivity.

Another important classification of aldolases is by their native donor substrate (*Figure 1.8B*). A general paradigm of substrate scope has been observed that aldolases are strict for their native donors but relatively promiscuous for acceptors, such that the broadening of donor scope is a topic of current interest [16,17,77]. The vast majority of known aldolases use either dihydroxyacetone phosphate (DHAP) or pyruvate as the native donor. Both donors are negatively charged and highly functionalized. DHAP-dependent aldolases are indispensable for the synthesis of rare sugars [78], and well-studied to the extent that enzymes can be chosen to selectively access any of four possible stereoisomeric products. A drawback is that DHAP is expensive and unstable, which can be remedied with *in situ* enzymatic generation of DHAP [79]. In comparison, pyruvate is an inexpensive donor but the resulting α -ketoacid aldol adducts are less valuable as final products, although they are amenable to downstream transformations. The donor substrate classification is largely orthogonal to the mechanistic classification, such that Type I and Type II enzyme families exist for both DHAP-dependent and pyruvate-dependent aldolases. Additionally, two exceptional Type I aldolases are known to utilize uncharged carbonyl compounds as donor. The deoxyribose-5-phosphate aldolase (DERA) of deoxynucleotide metabolism uses acetaldehyde, its function having diverged from structurally related glycolytic aldolases [80,81]. Similarly, the fructose-6-phosphate aldolase (FSA) uses dihydroxyacetone and appears to have evolved from photosynthetic transaldolases [82]. Breaking from the donor specificity paradigm, DERA and FSA are unusually tolerant of many simple aldehyde or ketone donors, and therefore enjoy a privileged position in

A Mechanism of Type I and Type II aldolase



B Classification of aldolases by donor substrate

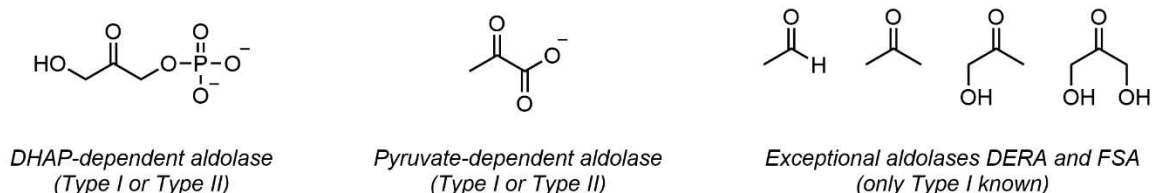


Figure 1.8. Classification of aldolases by mechanism or by donor substrate. A) Type I aldolases activate the donor by using a catalytic lysine residue to form a Schiff base, which is deprotonated to give a nucleophilic enamine. Type II aldolases activate the donor by chelation to a divalent metal ion and subsequent deprotonation to a nucleophilic enolate. B) Most structural families of aldolase have a strict requirement for DHAP or pyruvate as donor substrates. Deoxyribose-5-phosphate aldolase (DERA) and fructose-6-phosphate aldolase (FSA) are unique in their ability to use a variety of simple aldehydes and ketones as donors.

research and industry. Notably, engineered DERA is employed in biocatalytic syntheses of the cholesterol-lowering drug atorvastatin (Lipitor) and the anti-HIV drug islatravir [23,26].

The incorporation of aldol reactions in process chemistry has great potential to reduce step count and waste, by virtue of their high atom economy. Significant advances in chemocatalytic methods for stereoselective aldol reactions have alleviated the main challenge of dealing with product mixtures, and biocatalysts for the aldol reaction such as DERA have also broken into the mainstream. Nevertheless, robust and general fluoro-aldol reactions remain underdeveloped, thus necessitating continued research in chemocatalytic approaches and also exploration of biocatalytic approaches.

1.4. Thesis organization

The experiments and results described in this thesis set out to unify all three broad areas of biocatalysis, fluorine chemistry, and aldol reactions. These areas have individually seen tremendous growth, but their pairwise intersections are comparatively sparse. Biocatalytic aldol reactions emerged in the 1990s but only gained industrial acceptance more recently. Enantioselective fluoro-aldol reactions are few and far between, requiring further development in all modes of catalysis. Biocatalytic fluorination is nearly non-existent due to the rarity of fluorine in nature, but the use of promiscuous biocatalysts to upgrade from simple to complex organofluorines is increasingly recognized as a workaround to this knowledge gap.

We proposed to investigate the fluoro-donor scope of aldolases as a way to integrate the advantages of all three aforementioned areas. In the process this would challenge the paradigm of strict donor specificity in aldolases and probe the true limits of non-native substrates in both steric and electronic disruptions. The goal is to establish methodologies constituting a repertoire of biocatalytic, enantioselective fluoro-aldol reactions. In Chapter 2, we describe our discovery that the HpcH family of Type II pyruvate aldolase efficiently catalyzes aldol reactions of fluoropyruvate. The biochemical characterization of this non-native reaction is accompanied by a platform of downstream reactions for tailoring the fluoropyruvate aldol adducts into diverse organofluorines. The HpcH methodology is further extended in Chapter 3, where we push the frontiers of aldolase substrate scope by synthesizing and testing various β -fluoro- α -ketoacids as donor substrates, in combination with active site engineering of HpcH. Chapters 2 and 3 each culminate in proof-of-concept preparative syntheses of various organofluorines with relevance to bioactive natural products and drug fragments. In Chapter 4 we shift focus to the exceptional aldolases DERA and FSA, interrogating their potential to synthesize valuable fluorosugars directly from fluoroketone donors. The results of these studies represent a significant advance towards expanding the knowledge base for asymmetric fluorination and bridging the chasm between fluorine chemistry and biocatalysis, which we envision will positively impact the sustainable synthesis of fluorinated pharmaceuticals.

1.5. References

1. Koeller, K. M., Wong, C.-H. Enzymes for chemical synthesis. *Nature* **2001**, *409* (6817) 232-240.
2. Devine, P. N., Howard, R. M., Kumar, R., Thompson, M. P., Truppo, M. D., Turner, N. J. Extending the application of biocatalysis to meet the challenges of drug development. *Nat Rev Chem* **2018**, *2* 409-421.

3. Sheldon, R. A., Brady, D., Bode, M. L. The Hitchhiker's guide to biocatalysis: Recent advances in the use of enzymes in organic synthesis. *Chem Sci* **2020**, *11* (10) 2587-2605.
4. Höinig, M., Sondermann, P., Turner, N. J., Carreira, E. M. Enantioselective chemo- and biocatalysis: Partners in retrosynthesis. *Angew Chem Int Ed* **2017**, *56* (31) 8942-8973.
5. Sperl, J. M., Sieber, V. Multienzyme cascade reactions—Status and recent advances. *ACS Catal* **2018**, *8* (3) 2385-2396.
6. Yi, D., Bayer, T., Badenhorst, C. P. S., Wu, S., Doerr, M., Höhne, M., Bornscheuer, U. T. Recent trends in biocatalysis. *Chem Soc Rev* **2021**, *50* (14) 8003-8049.
7. Sheldon, R. A., Woodley, J. M. Role of biocatalysis in sustainable chemistry. *Chem Rev* **2018**, *118* (2) 801-838.
8. Wu, S., Snajdrova, R., Moore, J. C., Baldenius, K., Bornscheuer, U. T. Biocatalysis: Enzymatic synthesis for industrial applications. *Angew Chem Int Ed* **2020**, *60* (1) 88-119.
9. Davis, B. G., Boyer, V. Biocatalysis and enzymes in organic synthesis. *Nat Prod Rep* **2001**, *18* (6) 618-640.
10. Busto, E., Gotor-Fernández, V., Gotor, V. Hydrolases: Catalytically promiscuous enzymes for non-conventional reactions in organic synthesis. *Chem Soc Rev* **2010**, *39* (11) 4504-4523.
11. Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C., Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012**, *485* (7397) 185-194.
12. Giver, L., Gershenson, A., Freskgard, P. O., Arnold, F. H. Directed evolution of a thermostable esterase. *Proc Natl Acad Sci U S A* **1998**, *95* (22) 12809-12813.
13. Packer, M. S., Liu, D. R. Methods for the directed evolution of proteins. *Nat Rev Genet* **2015**, *16* (7) 379-394.
14. Sheldon, R. A., Brady, D. Broadening the scope of biocatalysis in sustainable organic synthesis. *ChemSusChem* **2019**, *12* (13) 2859-2881.
15. Bell, E. L., Finnigan, W., France, S. P., Green, A. P., Hayes, M. A., Hepworth, L. J., Lovelock, S. L., Niikura, H., Osuna, S., Romero, E., Ryan, K. S., Turner, N. J., Flitsch, S. L. Biocatalysis. *Nat Rev Methods Primers* **2021**, *1* (46).
16. Brovotto, M., Gamemara, D., Méndez, P. S., Seoane, G. A. C-C bond-forming lyases in organic synthesis. *Chem Rev* **2011**, *111* (7) 4346-4403.
17. Fesko, K., Gruber-Khadjawi, M. Biocatalytic methods for C-C bond formation. *ChemCatChem* **2013**, *5* (6) 1248-1272.
18. Schmidt, N. G., Eger, E., Kroutil, W. Building bridges: Biocatalytic C-C bond formation toward multifunctional products. *ACS Catal* **2016**, *6* (7) 4286-4311.
19. Liang, J., Lalonde, J., Borup, B., Mitchell, V., Mundorff, E., Trinh, N., Kochrekar, D. A., Cherat, R. N., Pai, G. G. Development of a biocatalytic process as an alternative to the (-)-DIP-Cl-mediated asymmetric reduction of a key intermediate of montelukast. *Org Process Res Dev* **2010**, *14* (1) 193-198.

20. Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W., Hughes, G. J. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **2010**, *329* (5989) 305-309.
21. Ma, S. K., Gruber, J., Davis, C., Newman, L., Gray, D., Wang, A., Grate, J., Huisman, G. W., Sheldon, R. A. A green-by-design biocatalytic process for atorvastatin intermediate. *Green Chem* **2010**, *12* 81-86.
22. DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., Kretz, K., Burk, M. J. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *J Am Chem Soc* **2003**, *125* (38) 11476-11477.
23. Greenberg, W. A., Varvak, A., Hanson, S. R., Wong, K., Huang, H., Chen, P., Burk, M. J. Development of an efficient, scalable, aldolase-catalyzed process for enantioselective synthesis of statin intermediates. *Proc Natl Acad Sci U S A* **2004**, *101* (16) 5788-5793.
24. De Souza, R. O. M. A., Miranda, L. S. M., Bornscheuer, U. T. A retrosynthesis approach for biocatalysis in organic synthesis. *Chemistry* **2017**, *23* (50) 12040-12063.
25. Birmingham, W. R., Starbird, C. A., Panosian, T. D., Nannemann, D. P., Iverson, T. M., Bachmann, B. O. Bioretrosynthetic construction of a didanosine biosynthetic pathway. *Nat Chem Biol* **2014**, *10* (5) 392-399.
26. Huffman, M. A., Fryszkowska, A., Alvizo, O., Borra-Garske, M., Campos, K. R., Canada, K. A., Devine, P. N., Duan, D., Forstater, J. H., Grosser, S. T., Halsey, H. M., Hughes, G. J., Jo, J., Joyce, L. A., Kolev, J. N., Liang, J., Maloney, K. M., Mann, B. F., Marshall, N. M., McLaughlin, M., Moore, J. C., Murphy, G. S., Nawrat, C. C., Nazor, J., Novick, S., Patel, N. R., Rodriguez-Granillo, A., Robaire, S. A., Sherer, E. C., Truppo, M. D., Whittaker, A. M., Verma, D., Xiao, L., Xu, Y., Yang, H. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **2019**, *366* (6470) 1255-1259.
27. McIntosh, J. A., Benkovics, T., Silverman, S. M., Huffman, M. A., Kong, J., Maligres, P. E., Itoh, T., Yang, H., Verma, D., Pan, W., Ho, H.-I., Vroom, J., Knight, A. M., Hurtak, J. A., Klapars, A., Fryszkowska, A., Morris, W. J., Strotman, N. A., Murphy, G. S., Maloney, K. M., Fier, P. S. Engineered ribosyl-1-kinase enables concise synthesis of Molnupiravir, an antiviral for Covid-19. *ACS Cent Sci* **2021**, *7* (12) 1980-1985.
28. Purser, S., Moore, P. R., Swallow, S., Gouverneur, V. Fluorine in medicinal chemistry. *Chem Soc Rev* **2008**, *37* (2) 320-330.
29. Zhou, Y., Wang, J., Gu, Z., Wang, S., Zhu, W., Aceña, J. L., Soloshonok, V. A., Izawa, K., Liu, H. Next generation of fluorine-containing pharmaceuticals, compounds currently in phase II-III clinical trials of major pharmaceutical companies: New structural trends and therapeutic areas. *Chem Rev* **2016**, *116* (2) 422-518.
30. Patani, G. A., LaVoie, E. J. Bioisosterism: A rational approach in drug design. *Chem Rev* **1996**, *96* (8) 3147-3176.
31. O'Hagan, D. Understanding organofluorine chemistry. An introduction to the C-F bond. *Chem Soc Rev* **2008**, *37* (2) 308-319.

32. Müller, K., Faeh, C., Diederich, F. Fluorine in pharmaceuticals: Looking beyond intuition. *Science* **2007**, *317* (5846) 1881-1886.
33. Gillis, E. P., Eastman, K. J., Hill, M. D., Donnelly, D. J., Meanwell, N. A. Applications of fluorine in medicinal chemistry. *J Med Chem* **2015**, *58* (21) 8315-8359.
34. Liang, T., Neumann, C. N., Ritter, T. Introduction of fluorine and fluorine-containing functional groups. *Angew Chem Int Ed* **2013**, *52* (32) 8214-8264.
35. Wu, J. Review of recent advances in nucleophilic C-F bond-forming reactions at sp³ centers. *Tetrahedron Lett* **2014**, *55* (31) 4289-4294.
36. Lee, J.-W., Oliveira, M. T., Jang, H. B., Lee, S., Chi, D. Y., Kim, D. W., Song, C. E. Hydrogen-bond promoted nucleophilic fluorination: concept, mechanism, and applications in positron emission tomography. *Chem Soc Rev* **2016**, *45* (17) 4638-4650.
37. Middleton, W. J. New fluorinating reagents. Dialkylaminosulfur fluorides. *J Org Chem* **1975**, *40* (5) 574-578.
38. Nielsen, M. K., Ugaz, C. R., Li, W., Doyle, A. G. PyFluor: A low-cost, stable, and selective deoxyfluorination reagent. *J Am Chem Soc* **2015**, *137* (30) 9571-9574.
39. Lal, G. S., Pez, G. P., Syvret, R. G. Electrophilic NF fluorinating agents. *Chem Rev* **1996**, *96* (5) 1737-1756.
40. Nyffeler, P. T., Durón, S. G., Burkart, M. D., Vincent, S. P., Wong, C.-H. Selectfluor: Mechanistic insight and applications. *Angew Chem Int Ed* **2004**, *44* (2) 192-212.
41. Rueda-Becerril, M., Sazepin, C. C., Leung, J. C. T., Okbinoglu, T., Kennepohl, P., Paquin, J.-F., Sammis, G. M. Fluorine transfer to alkyl radicals. *J Am Chem Soc* **2012**, *134* (9) 4026-4029.
42. Meyer, D., Jangra, H., Walther, F., Zipse, H., Renaud, P. A third generation of radical fluorinating agents based on *N*-fluoro-*N*-arylsulfonamides. *Nat Commun* **2018**, *9* (1) 4888.
43. Roque, J. B., Kuroda, Y., Göttemann, L. T., Sarpong, R. Deconstructive fluorination of cyclic amines by carbon-carbon cleavage. *Science* **2018**, *361* (6398) 171-174.
44. Adler, P., Teskey, C. J., Kaiser, D., Holy, M., Sitte, H. H., Maulide, N. α -Fluorination of carbonyls with nucleophilic fluorine. *Nat Chem* **2019**, *11* (4) 329-334.
45. Yang, X., Wu, T., Phipps, R. J., Toste, F. D. Advances in catalytic enantioselective fluorination, mono-, di-, and trifluoromethylation, and trifluoromethylthiolation reactions. *Chem Rev* **2015**, *115* (2) 826-870.
46. Althaus, M., Becker, C., Togni, A., Mezzetti, A. Ruthenium-catalyzed asymmetric electrophilic fluorination of 1,3-dicarbonyl compounds. *Organometallics* **2007**, *26* (24) 5902-5911.
47. Rauniyar, V., Lackner, A. D., Hamilton, G. L., Toste, F. D. Asymmetric electrophilic fluorination using an anionic chiral phase-transfer catalyst. *Science* **2011**, *334* (6063) 1681-1684.
48. Meng, W.-T., Zheng, Y., Nie, J., Xiong, H.-Y., Ma, J.-A. Organocatalytic asymmetric one-pot sequential conjugate addition/dearomative fluorination: Synthesis of chiral fluorinated isoxazol-5(4H)-ones. *J Org Chem* **2013**, *78* (2) 559-567.

49. Xu, J., Hu, Y., Huang, D., Wang, K.-H., Xu, C., Niu, T. Thiourea-catalyzed enantioselective fluorination of β -keto esters. *Adv Synth Catal* **2012**, *354* (2-3) 515-526.
50. Kalow, J. A., Doyle, A. G. Enantioselective ring opening of epoxides by fluoride anion promoted by a cooperative dual-catalyst system. *J Am Chem Soc* **2010**, *132* (10) 3268-3269.
51. Pupo, G., Ibba, F., Ascough, D. M. H., Vicini, A. C., Ricci, P., Christensen, K. E., Pfeifer, L., Morphy, J. R., Brown, J. M., Paton, R. S., Gouverneur, V. Asymmetric nucleophilic fluorination under hydrogen bonding phase-transfer catalysis. *Science* **2018**, *360* (6389) 638-642.
52. Wang, Q., Lübcke, M., Biosca, M., Hedberg, M., Eriksson, L., Himo, F., Szabó, K. J. Enantioselective construction of tertiary fluoride stereocenters by organocatalytic fluorocyclization. *J Am Chem Soc* **2020**, *142* (47) 20048-20057.
53. Han, X., Luo, J., Liu, C., Lu, Y. Asymmetric generation of fluorine-containing quaternary carbons adjacent to tertiary stereocenters: Uses of fluorinated methines as nucleophiles. *Chem Commun* **2009**, *15* 2044-2046.
54. Liang, Y., Fu, G. C. Catalytic asymmetric synthesis of tertiary alkyl fluorides: Negishi cross-couplings of racemic α,α -dihaloketones. *J Am Chem Soc* **2014**, *136* (14) 5520-5524.
55. Saadi, J., Wennemers, H. Enantioselective aldol reactions with masked fluoroacetates. *Nat Chem* **2016**, *8* (3) 276-280.
56. Cuadros, S., Dell'Amico, L., Melchiorre, P. Forging fluorine-containing quaternary stereocenters by a light-driven organocatalytic aldol desymmetrization process. *Angew Chem Int Ed* **2017**, *56* (39) 11875-11879.
57. He, Z.-T., Jiang, X., Hartwig, J. F. Stereodivergent construction of tertiary fluorides in vicinal stereogenic pairs by allylic substitution with iridium and copper catalysts. *J Am Chem Soc* **2019**, *141* (33) 13066-13073.
58. Liu, J., Yuan, Q., Toste, F. D., Sigman, M. S. Enantioselective construction of remote tertiary carbon-fluorine bonds. *Nat Chem* **2019**, *11* (8) 710-715.
59. Zhu, Y., Han, J., Wang, J., Shibata, N., Sodeoka, M., Soloshonok, V. A., Coelho, J. A. S., Toste, F. D. Modern approaches for asymmetric construction of carbon-fluorine quaternary stereogenic centers: Synthetic challenges and pharmaceutical needs. *Chem Rev* **2018**, *118* (7) 3887-3964.
60. Zhu, X., Robinson, D. A., McEwan, A. R., O'Hagan, D., Naismith, J. H. Mechanism of enzymatic fluorination in *Streptomyces cattleya*. *J Am Chem Soc* **2007**, *129* (47) 14597-14604.
61. O'Hagan, D., Harper, D. B. Fluorine-containing natural products. *J Fluor Chem* **1999**, *100* (1-2) 127-133.
62. Chan, K. K. J., O'Hagan, D. The rare fluorinated natural products and biotechnological prospects for fluorine enzymology. *Methods Enzymol* **2012**, *516* 219-235.

63. Berkowitz, D. B., Karukurichi, K. R., de la Salud-Bea, R., Nelson, D. L., McCune, C. D. Use of fluorinated functionality in enzyme inhibitor development: Mechanistic and analytical advantages. *J Fluor Chem* **2008**, *129* (9) 731-742.
64. Trost, B. M., Brindle, C. S. The direct catalytic asymmetric aldol reaction. *Chem Soc Rev* **2010**, *39* (5) 1600-1632.
65. Yamashita, Y., Yasukawa, T., Yoo, W.-J., Kitanosono, T., Kobayashi, S. Catalytic enantioselective aldol reactions. *Chem Soc Rev* **2018**, *47* (12) 4388-4480.
66. Lovering, F., Bikker, J., Humblet, C. Escape from Flatland: Increasing saturation as an approach to improving clinical success. *J Med Chem* **2009**, *52* (21) 6752-6756.
67. Wurtz, C. A. Sur un aldéhyde-alcool (On an aldehyde alcohol). *Bulletin de la Société Chimique de Paris* **1872**, *17* 436-442.
68. Evans, D. A., Nelson, J. V., Vogel, E., Taber, T. R. Stereoselective aldol condensations via boron enolates. *J Am Chem Soc* **1981**, *103* (11) 3099-3111.
69. Brown, H. C., Dhar, R. K., Bakshi, R. K., Pandiarajan, P. K., Singaram, B. Major effect of the leaving group in dialkylboron chlorides and triflates in controlling the stereospecific conversion of ketones into either (*E*)- or (*Z*)-enol borinates. *J Am Chem Soc* **1989**, *111* (9) 3441-3442.
70. Evans, D. A., Bartroli, J., Shih, T. L. Enantioselective aldol condensations. 2. Erythroselective chiral aldol condensations via boron enolates. *J Am Chem Soc* **1981**, *103* (8) 2127-2129.
71. Notz, W., Tanaka, F., Barbas III, C. F. Enamine-based organocatalysis with proline and diamines: The development of direct catalytic asymmetric aldol, Mannich, Michael, and Diels-Alder reactions. *Acc Chem Res* **2004**, *37* (8) 580-591.
72. Welch, J. T., Seper, K., Eswarakrishnan, S., Samartino, J. Preparation of α -fluoro enolates and their use in the directed aldol reaction. *J Org Chem* **1984**, *49* 4720-4721.
73. Huang, X.-T., Chen, Q.-Y. Ethyl α -fluoro silyl enol ether: Stereoselective synthesis and its aldol reaction with aldehydes and ketones. *J Org Chem* **2002**, *67* 3231-3234.
74. Welch, J. T., Plummer, J. S. The stereoselective aldol condensation of α -fluoroacetates. *Synth Commun* **1989**, *19* (5-6) 1081-1090.
75. Saadi, J., Akakura, M., Yamamoto, H. Rapid, one-pot synthesis of β -siloxy- α -haloaldehydes. *J Am Chem Soc* **2011**, *133* (36) 14248-14251.
76. Oda, S., Yamamoto, H. Synthesis of β -hydroxy- α -haloesters through super silyl ester directed *syn*-selective aldol reaction. *Org Lett* **2013**, *15* (23) 6030-6033.
77. Dean, S. M., Greenberg, W. A., Wong, C.-H. Recent advances in aldolase-catalyzed asymmetric synthesis. *Adv Synth Catal* **2007**, *349* (8-9) 1308-1320.
78. Li, A., Cai, L., Chen, Z., Wang, M., Wang, N., Nakanishi, H., Gao, X.-D., Li, Z. Recent advances in the synthesis of rare sugars using DHAP-dependent aldolases. *Carbohydr Res* **2017**, *452* 108-115.
79. Schümperli, M., Pellaux, R., Panke, S. Chemical and enzymatic routes to dihydroxyacetone phosphate. *Appl Microbiol Biotechnol* **2007**, *75* (1) 33-45.

80. Barbas III, C. F., Wang, Y. F., Wong, C.-H. Deoxyribose-5-phosphate aldolase as a synthetic catalyst. *J Am Chem Soc* **1990**, *112* (5) 2013-2014.
81. Chambre, D., Guérard-Hélaine, C., Darii, E., Mariage, A., Petit, J.-L., Salanoubat, M., de Berardinis, V., Lemaire, M., Hélaine, V. 2-Deoxyribose-5-phosphate aldolase, a remarkably tolerant aldolase towards nucleophile substrates. *Chem Commun* **2019**, *55* (52) 7498-7501.
82. Roldán, R., Sanchez-Moreno, I., Scheidt, T., Hélaine, V., Lemaire, M., Parella, T., Clapés, P., Fessner, W.-D., Guérard-Hélaine, C. Breaking the dogma of aldolase specificity: Simple aliphatic ketones and aldehydes are nucleophiles for fructose-6-phosphate aldolase. *Chem Eur J* **2017**, *23* (21) 5005-5009.

Chapter 2: *Developing a pyruvate aldolase platform for the synthesis of chiral organofluorines*

Portions of this work were published in the following scientific journal:

Fang, J., Hait, D., Head-Gordon, M., Chang, M. C. Y. Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angew Chem Int Ed* **2019**, 58 (34) 11841-11845.

Portions of this work were performed in collaboration with the following persons:

Diptarka Hait performed DFT calculations on the energetics of the reaction intermediate. Dr. Jonathan L. McMurry and Dr. Chia-I Lin provided invaluable advice in cloning, fluorine chemistry, and carbohydrate chemistry.

2.1. Introduction

Among a variety of enzyme types that catalyze C-C bond formation, aldolases have emerged as uniquely suited to the sustainable synthesis of complex synthons [1-3]. This is because the aldol addition reaction, when effected without subsequent dehydration (i.e. avoiding the aldol condensation reaction), joins together a nucleophilic substrate and an electrophilic substrate with 100% atom economy and produces up to two new stereocenters. Although synthetic aldol reactions using chiral auxiliaries or chiral catalysts can often control the stereochemical outcome of the nascent chiral centers, these methods still lack in universality and efficiency [4,5]. The freedom to choose one's desired stereoisomer and the conferred degree of stereoselectivity can now be rivalled by aldol biocatalysts. For example, within the family of DHAP-dependent aldolases to which the glycolytic fructose-1,6-bisphosphate aldolase belongs, specific enzymes are known for accessing each of the four possible configurations of vicinal hydroxy stereocenters [6,7].

The use of α -fluoro substituted nucleophiles in what is termed the fluoro-aldol reaction poses a particular challenge to diastereo- and enantioselectivity, since the small size of fluorine provides little steric contribution to differentiating the energy of diastereomeric products or transition states. Older methods for inducing diastereoselectivity in the fluoro-aldol reaction have relied upon unusual and extremely bulky pendant groups on a fluoroacetate nucleophile, such as 2,6-di-*tert*-butyl-4-methylphenyl esters or tris(triethylsilyl)silyl esters [8,9]. The first enantioselective fluoro-aldol reaction was only reported in 2016—this organocatalytic strategy based on fluoromalonic acid half thioesters represents a critical study in the unsolved problem of fluoro-aldol reactions (*Figure 2.1A*) [10]. Considering this breakthrough and others in the active field of asymmetric fluorination, a few groups including ours were encouraged to study biocatalytic fluoro-aldol reactions, a project which would begin with interrogating wild-type aldolases for promiscuous activity on fluorinated donors. Not only could such endeavors have general benefits in pharmaceutical synthesis where fluorine incorporation and stereochemical fidelity are key to safety and efficacy, but the origin of most aldolases in carbohydrate metabolism makes this strategy attractive for the synthesis of fluorinated sugars. Nucleoside-mimicking anticancer and antiviral drugs in addition to chemical and radiochemical probes frequently contain fluorinated sugars, and their vicinal fluoro-hydroxy motifs appear suitably poised for construction via biocatalyzed fluoro-aldol reaction [11,12].

This chapter focuses on our efforts to build a chemoenzymatic platform towards chiral organofluorines, using pyruvate aldolases as the entryway to molecules with fluorine stereocenters. Unlike DHAP-dependent aldolases which require an expensive phosphorylated donor, pyruvate aldolases use α -ketoacid donors which are cheaply available. Aldol adducts constructed by pyruvate aldolases still contain a densely-functionalized α -ketoacid motif, which is useful for diversification with downstream reactions. Furthermore, fluoropyruvate is a commercially available substance used as biochemical inhibitor of pyruvate dehydrogenase complex [13], and it is relatively non-toxic (in comparison to the extremely toxic fluoroacetate). These characteristics make fluoropyruvate attractive as a potential non-native substrate of pyruvate aldolases.

Pyruvate aldolases fall into four main structural families using the organization of the Pfam database, with two families each belonging to the Type I (donor activation by catalytic lysine) and Type II (donor activation by divalent metal) mechanistic types. Some of these aldolases have been investigated by previous groups for fluoropyruvate incorporation (*Figure 2.1B*). Type I KDPG aldolases (PF01801) are excellent biocatalysts for addition of pyruvate to diverse aldehydes, but have almost no tolerance for fluoropyruvate [14,15]. Type I DHDPS aldolases (PF00701) have

much greater ability to utilize fluoropyruvate, with resulting (3*R*)-configuration of the fluorine stereocenter under kinetic control [16-18]. However, these enzymes tend to have a narrow scope of the acceptor substrate, such as long-chain aldose sugars or heteroaromatic aldehydes. Type I aldolases are also sensitive to inhibition of the catalytic lysine under high aldehyde concentrations [19]. We were drawn to the promise of lesser-studied Type II aldolases being capable of functioning under such high substrate load conditions, which is desirable for industrial biocatalysis. The Type II DmpG aldolases (PF07836/00782) have evolved to perform retro-aldol reactions via complexation with dehydrogenases, which poses biochemical difficulties to repurposing them for synthetic reactions [20]. Thus we turned our attention to the remaining Type II HpcH aldolase family (PF03328) for selection of enzyme candidates with synthetic potential in the desired fluoropyruvate aldol reaction [21].

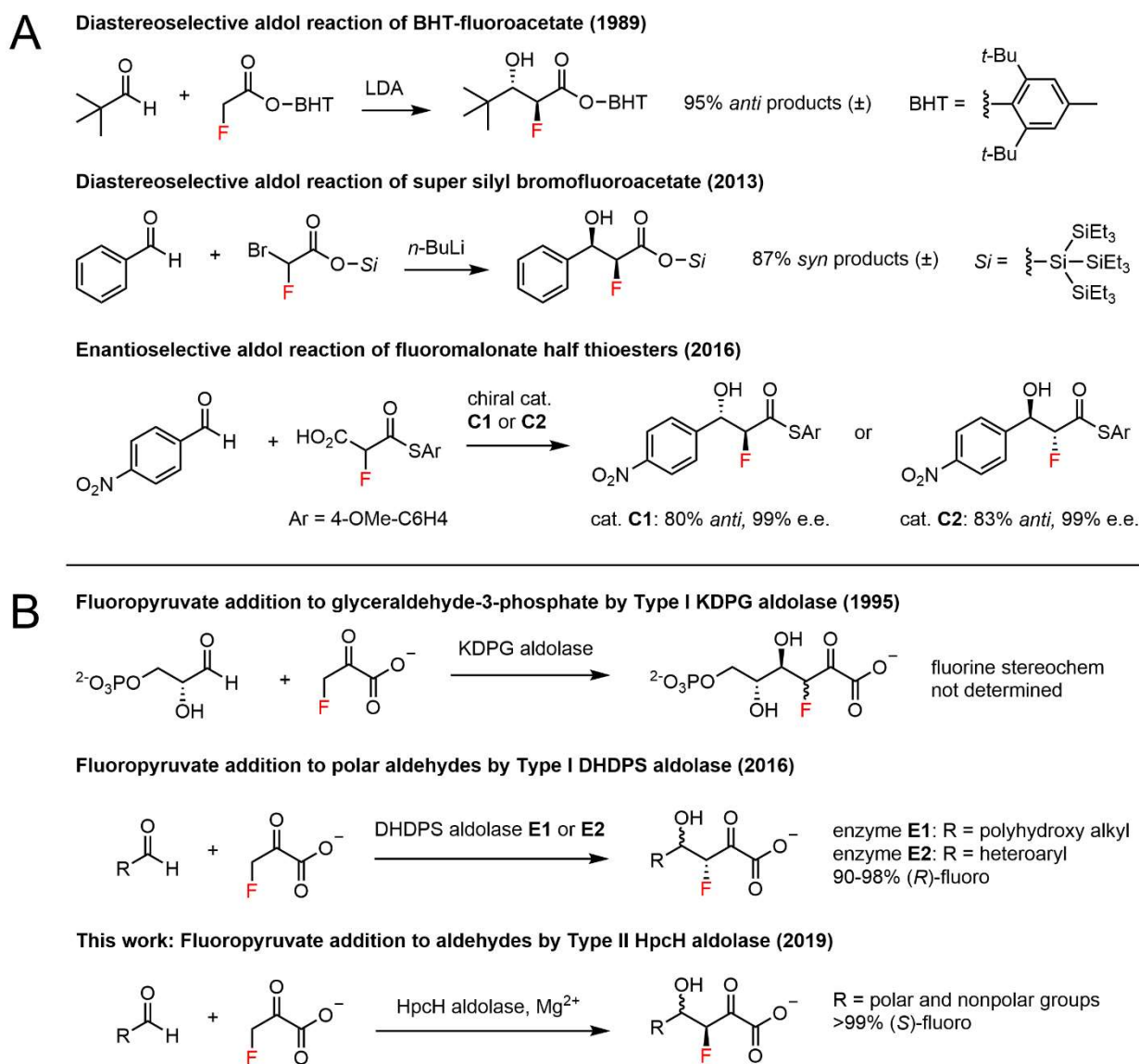


Figure 2.1. Fluoro-aldol reactions. A) Early examples of diastereoselective fluoro-aldol reactions relying upon extremely bulky esters. First enantioselective fluoro-aldol reaction, using fluoromalonate half thioesters and cinchona alkaloid-derived catalysts. B) Enzymatic aldol addition of fluoropyruvate catalyzed by various pyruvate aldolases.

2.2. Materials and methods

Commercial materials. Acetaldehyde, ammonium formate (LC-MS grade), ammonium persulfate, benzaldehyde, *O*-benzylhydroxylamine hydrochloride, bovine serum albumin (BSA), butyraldehyde, catalase from bovine liver, diethyl *L*-tartrate, diethylaminosulfur trifluoride (DAST), diisobutylaluminum hydride (DIBAL-H, 1 M in CH₂Cl₂), dimethoxyacetaldehyde (60 wt. % in H₂O), 2,2-dimethoxypropane, 4,4-dimethyl-4-silapentane-1-sulfonic acid sodium salt (DSS), Dowex 50WX8 hydrogen form, β -fluoropyruvic acid sodium salt monohydrate, fluorouracil, *D*-glyceraldehyde, *L*-glyceraldehyde, glycolaldehyde dimer, glyoxylic acid monohydrate, 2-iodoxybenzoic acid (stabilized, 45 wt. %), *D*-lactaldehyde (1 M in H₂O), *L*-lactaldehyde (1 M in H₂O), *L*-lactate dehydrogenase from rabbit muscle (LDH), lysozyme, 2-mercaptoethanol, 3 M methanolic HCl, (*S*)- α -methylbenzylamine (SMBA), β -nicotinamide adenine dinucleotide hydrate (NAD⁺), β -nicotinamide adenine dinucleotide reduced disodium salt hydrate (NADH), oxalyl chloride (2 M in CH₂Cl₂), phenylmethanesulfonyl fluoride (PMSF), polyethyleneimine (PEI), propionaldehyde, 2-pyridinecarboxaldehyde, pyridoxal 5'-phosphate (PLP), sodium pyruvate, sodium dodecyl sulfate, and *p*-toluenesulfonic acid monohydrate were purchased from Sigma-Aldrich (St. Louis, MO). Agarose, ammonium bicarbonate (LC-MS grade), ammonium chloride, bromophenol blue, calcium chloride dihydrate, carbenicillin disodium salt, dithiothreitol (DTT), deoxynucleotides (dNTPs), ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA), formic acid (LC-MS grade), glycine, 30 wt. % hydrogen peroxide, HEPES, imidazole, kanamycin sulfate, O'GeneRuler 1 kb Plus DNA Ladder, PageRuler Plus Prestained Protein Ladder, polyethylene glycol 3350, polyethylene glycol 8000, potassium chloride, potassium sodium tartrate, sodium bicarbonate, sodium chloride, sodium hydroxide, sodium phosphate monobasic monohydrate, trichloroacetic acid, Tris base, tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and urea were purchased from Thermo Fisher Scientific (Waltham, MA). Glycerol, LB Miller agar, LB Miller broth, magnesium chloride hexahydrate, anhydrous magnesium sulfate, anhydrous sodium sulfate, Terrific broth, and triethylamine were purchased from EMD-Millipore (Burlington, MA). InstantBlue Protein Stain was from Expedeon (San Diego, CA). Isopropyl β -*D*-1-thiogalactopyranoside (IPTG) was from Santa Cruz Biotechnology (Dallas, TX). Acrylamide-bis 30% solution, Bradford assay reagent concentrate, ethidium bromide, and tetramethylethylenediamine (TEMED) were purchased from Bio-Rad Laboratories (Hercules, CA). Chloroform-*d*, deuterium oxide, and dimethylsulfoxide-*d*₆ were from Cambridge Isotope Laboratories (Tewksbury, MA). Phusion polymerase, Phusion HF buffer, all restriction enzymes, restriction enzyme buffer (CutSmart), and Taq ligase were from New England Biolabs (Ipswich, MA). T5 exonuclease was from Epicentre (Madison, WI). Ni-NTA agarose resin and DNA purification kits were purchased from Qiagen (Redwood City, CA). Oligonucleotides and gBlocks gene fragments were synthesized by Integrated DNA Technologies (Coralville, IA). All chemicals were used as purchased without further purification.

Bacterial strains. *Escherichia coli* DH10B-T1^R was used for plasmid construction and *E. coli* BL21(DE3)-T1^R was used for protein production. The strains were made chemically competent by the method of [22].

Construction of expression plasmids. Plasmids were constructed by Gibson isothermal assembly and verified by sequencing (Quintara Biosciences; South San Francisco, CA and Genewiz; South Plainfield, NJ). *In silico* sequence alignments were performed on Benchling (Benchling, San Francisco, CA). All DNA sequences used in this study are listed in *Appendix 3*. Plasmids encoding N-terminally His₁₀-tagged Type II aldolases were derived from pET16hp-

IMDH which was linearized with NdeI/BamHI. For the plasmid encoding the C-terminally His₆-tagged transaminase from *Vibrio fluvialis* (VfAT), the pET24a vector was linearized with NdeI/XhoI. The vector fragments were purified by extraction from 1% agarose gel. The genes *garL* and *rhmA*, encoding Type II aldolases GarL and RhmA from *E. coli*, were amplified by polymerase chain reaction (PCR) from the genome of *E. coli* DH10B-T1^R. Primers were reconstituted to 100 μM in water. The 50 μL PCR mixtures contained 5× Phusion HF buffer, 0.2 mM dNTPs, 1 μL cell suspension (a single colony in 50 μL water), 2 units Phusion HF polymerase, and 0.5 μM each primer. Thermal cycling was as follows: 3 min at 98 °C, followed by 35 cycles of 30 s at 98 °C, 30 s at (T_m-5) °C, and 1 min at 72 °C, followed by a final extension of 5 min at 72 °C. A 5 μL sample was visualized on 1% agarose gel and the remaining material was purified by spin column. The genes encoding EchPcH from *E. coli* C strain, SwHpcH1 and SwHpcH2 from *Sphingomonas wittichi*, and VfAT were obtained as codon-optimized gBlocks containing vector-homologous overhangs and were resuspended in water to 50 ng/μL.

Purified vectors and inserts were combined (9:1 v/v) in 5 μL volume, to which 15 μL of Gibson master mix, prepared according to [23], was added. The mixture was incubated in a thermocycler at 50 °C for 1 h, then added to chemically competent *E. coli* DH10B-T1^R along with 20 μL of 5× KCM solution (0.5 M KCl, 0.15 M CaCl₂, 0.25 M MgCl₂) and water to 200 μL total. After incubation on ice for 20 min, the cells were heat-shocked at 42 °C for 90 s and returned to ice for 2 min. Cells were diluted with 1 mL of LB, recovered with shaking at 37 °C for 45 min, and pelleted by centrifugation (7,000 × g, 1 min). The supernatant was partially removed (0.8 mL), then the cells were resuspended and plated on LB agar with 50 μg/mL antibiotic. Plates were incubated at 37 °C overnight. Liquid cultures (5 mL LB with 50 μg/mL antibiotic) were inoculated with single colonies and grown at 37 °C overnight, following which the plasmids were isolated and confirmed by sequencing from the T7 promoter and T7 terminator.

Point mutations in the plasmid pET16hp-HpcH were introduced by the two-step modification of the Quikchange site-directed mutagenesis method [24]. Mutagenic primers were reconstituted to 100 μM in water. Single primer extension reactions (50 μL) were prepared containing 5X Phusion HF buffer, 0.2 mM dNTPs, 20 ng template, 1 unit Phusion HF polymerase, and 0.5 μM forward or reverse primer. Thermal cycling was as follows: 3 min at 98 °C, followed by 10 cycles of 30 s at 98 °C, 30 s at 65 °C, and 3 min at 72 °C, followed by a final extension of 5 min at 72 °C. Then, 25 μL of each primer extension reaction was combined and supplemented with 1 unit Phusion HF polymerase. Thermal cycling was continued for 25 cycles. DpnI (2 units) was added to digest the template for 1 h at 37 °C. *E. coli* DH10B-T1^R was transformed with 5 μL of the mixture (KCM method), then plated on LB agar with 50 μg/mL carbenicillin and incubated at 37 °C overnight. Liquid cultures were inoculated with single colonies and grown at 37 °C overnight. The plasmids were isolated and mutations confirmed by sequencing.

Expression and purification of His-tagged proteins. Expression plasmids (10 ng) were transformed into *E. coli* BL21(DE3)-T1^R (KCM method) and a single colony was used to inoculate 25 mL of TB media with 50 μg/mL of the appropriate antibiotic. The seed culture was grown overnight with shaking at 37 °C (200 RPM). Then, 1 L of TB in an Ultra Yield baffled flask (Thomson Instrument Company; Oceanside, CA) was inoculated with the seed culture and growth was continued to OD₆₀₀ of 0.8-1.2. The culture was chilled in an ice bath for 15 min and protein expression was induced with IPTG (1 mM). Expression proceeded at 16 °C overnight, after which the cells were harvested (7,000 × g, 5 min, 4 °C) and either stored at -80 °C or subjected to protein purification immediately.

Harvested cells were resuspended in 5 mL/(g wet cell wt.) of lysis buffer (20 mM Tris pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% v/v glycerol). To the cells was added PMSF (1 mM) and lysozyme (1 mg/mL). After incubation at r.t. for 30 min, the sample was sonicated (QSonica Q700) with the following program: 10 s on, 20 s off, 1 min total process time, amplitude 50. Cell debris was removed by centrifugation (15,000 × g, 20 min, 4 °C). To the supernatant was added PEI (0.05% w/v) and the precipitated nucleic acids were removed by centrifugation (15,000 × g, 20 min, 4 °C). Ni-NTA agarose resin (50% suspension in 20% EtOH, 0.2 mL/(g wet cell wt.)) was added to the samples, which were shaken gently on an orbital shaker at r.t. for 30 min and then poured into a glass column. The resin was washed with >20 column volumes (CV) of wash buffer (same as lysis buffer with 20 mM imidazole) until the effluent tested negative by the Bradford dye-binding assay. Protein was eluted with >5 CV of elution buffer (same as lysis buffer with 250 mM imidazole) until the effluent tested negative again. The elution fraction was supplemented with β-ME (0.1% v/v) and relevant cofactors (1 mM MgCl₂ for Type II aldolases; 20 μM PLP for VfAT). Fractions and purified proteins were analyzed by SDS-PAGE.

The proteins were concentrated in Amicon Ultra-15 centrifugal filter units with 10 kDa molecular weight cutoff (EMD Millipore; Burlington, MA). Note that the HpcH family Type II aldolases have poor solubility at 4 °C, and concentration past 2 mg/mL should be performed at room temperature. Buffer exchange was conducted with PD-10 columns or Sephadex G-25 Medium (GE Healthcare; Chicago, IL) into HEPES storage buffer (50 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% v/v glycerol, 1 mM TCEP) or phosphate storage buffer (50 mM NaH₂PO₄-NaOH pH 7.5, 150 mM NaCl, 10% v/v glycerol, 1 mM DTT). Desalted samples were supplemented with cofactors and further concentrated. The concentrations of Type II aldolases were determined by the absorbance at 280 nm, using extinction coefficients (GarL: 22460 M⁻¹cm⁻¹, RhmA: 36900 M⁻¹cm⁻¹, EcHpcH: 32430 M⁻¹cm⁻¹, SwHpcH1: 18450 M⁻¹cm⁻¹, SwHpcH2: 25440 M⁻¹cm⁻¹) predicted with ExPASy ProtParam (Swiss Institute of Bioinformatics). The concentration of VfAT was determined by the Bradford dye-binding assay against a BSA standard curve [25].

Safety note on handling of organofluorine compounds. Fluoroacetate is an extremely toxic metabolic poison. The decarboxylative treatment of fluoropyruvate aldol reactions, used throughout this chapter, generates fluoroacetate in solution. Such samples should be handled with great care. Other organofluorines including fluoropyruvate and novel synthesized compounds may have uncharacterized toxicity and warrant general handling precautions.

Assays for aldolase-catalyzed fluoropyruvate addition. Aldol addition reactions of total volume 600 μL were prepared containing 50 mM fluoropyruvate, 55 mM aldehyde, 1 mM MgCl₂, and 0.01-0.05 mol% aldolase (eg. 90-450 μg for 30 kDa protein) in either 20 mM HEPES-NaOH or 50 mM NaH₂PO₄-NaOH pH 7.5 buffers. After screening the aldehydes glycolaldehyde (GA), dimethoxyacetaldehyde (DMA), and glyoxylic acid (GX) for the optimal buffer, HEPES was used for RhmA, EcHpcH, and SwHpcH2; phosphate was used for GarL and SwHpcH1. Substrates were added from 1 M stock solutions in water except for benzaldehyde (1 M in DMSO, final concentration 5% v/v DMSO). The stock solution of GX was neutralized with 3 M NaOH prior to use. The reactions were initiated with aldolase and proceeded at room temperature for 24 h. Reactions were quenched by decarboxylation with addition of 60 μL of 30 wt. % H₂O₂ and incubation for 30 min. D₂O (100 μL) was added and precipitated proteins were removed by centrifugation at 20,000 × g for 5 min at 4 °C. Samples were analyzed by quantitative ¹⁹F NMR (564 MHz). The extent of conversion (see *Appendix 2* for full data) was determined by integration of the decarboxylated products (dd) against fluoroacetate (-218.0 ppm, t).

1D-NMR spectroscopy. All NMR was performed at the UC Berkeley College of Chemistry NMR Facility on a Bruker AVQ-400 (400 MHz, Z-gradient QNP probe), AV-500 (500 MHz, Z-gradient BBO probe) or AV-600 (600 MHz, Z-gradient BBO probe). All experiments were conducted at 298 K. Chemical shifts were referenced to the solvent and expressed relative to tetramethylsilane (^1H and ^{13}C) or trichlorofluoromethane (^{19}F). For all ^{19}F NMR, the spectral center was set to -150 ppm. Quantitative ^{19}F NMR experiments included a recycle delay of 20 s between scans to allow for full relaxation. Data was processed in MestreNova (MestreLab Research). Baseline distortion in the -150 to -200 ppm region due to fluoropolymer in the probes was corrected with backward linear prediction (Toeplitz) of the first 128 data points, followed by manual phase correction and final baseline correction (Whittaker smoother).

Characterization of relative stereochemistry by nitric acid oxidation. An aqueous mixture of L-tartrate and (2*S*,3*S*)-2-fluoro-3-hydroxysuccinate was synthesized by deoxofluorination. In a dry round-bottom flask, diethyl L-tartrate (206 mg, 1 mmol) was dissolved in 5 mL anhydrous CH_2Cl_2 , cooled to -78 °C, and treated with DAST (0.13 mL, 1 mmol). The solution was stirred for 5 h, then quenched with water (1 mL) and warmed to room temperature. The organic layer was washed with water (2×10 mL) and concentrated. A 20 μL sample of the crude residue was hydrolyzed by heating in 0.6 mL 1 M HCl (95 °C, 1 h). The sample was neutralized with 3 M NaOH and washed with CH_2Cl_2 (2×0.5 mL), then D_2O (100 μL) was added and the sample was analyzed by ^{19}F NMR to ascertain the chemical shift. Samples of fluoroacids derived from aldehydes GA (GarL), DMA (RhmA), and GX (GarL and EcHpcH) were prepared in 600 μL volume with the fluoropyruvate addition method from above, using 0.05 mol% aldolase followed by decarboxylation. Then, concentrated nitric acid was added to the GA sample (800 μL , 9.0 M HNO_3) and DMA sample (100 μL , 2.3 M HNO_3). Oxidation was effected by heating at 95 °C for 1 h, followed by neutralization with 10 M NaOH and addition of D_2O (100 μL). The four samples containing 2-fluoro-3-hydroxysuccinate in various diastereomeric ratios were analyzed by quantitative ^{19}F NMR (564 MHz).

Characterization of absolute stereochemistry by enzymatic reduction. Samples of fluoropyruvate adducts derived from GA (GarL and EcHpcH), DMA (RhmA and EcHpcH), GX (GarL and EcHpcH), and propionaldehyde (RhmA and EcHpcH) were prepared in 300 μL volume by the previous fluoropyruvate addition method without decarboxylation, using 0.05 mol% of the indicated aldolases. To the samples was added 100 μL of 100 mM EDTA pH 9.0 to inactivate the aldolase, followed 10 min later by NADH (200 μL , 100 mM stock in 10 mM NaOH) and LDH (100 μL of 2 mg/mL stock). The enzymatic reactions (14.3 mM EDTA, 28.6 mM NADH, maximum of 21.4 mM substrate, and 0.2 mg LDH) were incubated at room temperature for 16 h followed by addition of D_2O (100 μL). The samples containing 3-fluoro-2-hydroxyacids were analyzed by quantitative ^{19}F NMR (564 MHz).

Density functional theory analysis of model systems. The active site intermediate of fluoropyruvate addition by HpcH family aldolases was modeled as the octahedral complex $[\text{Mg}(\text{H}_2\text{O})_2(\text{OAc})_2(\text{Fpyr-enolate})]^{2-}$ with either (*E*)- or (*Z*)-configuration at the fluoroenolate. All calculations were done with the Q-Chem 5.2 package using the $\omega\text{B97M-V}$ range-separated hybrid density functional and the SMD solvent model [26-28]. All geometries were optimized using the def2-SVPD basis set, and single point energies for the resulting optimized structures were recomputed with the larger def2-TZVPD basis set [29]. Nuclear contributions to the free energy were estimated using a modified quasi rigid rotor harmonic oscillator model with $\omega\text{B97M-V/def2-SVPD}$ geometries and frequencies [30]. These contributions to the free energy differences however

were quite small (~1 kcal/mol or smaller) relative to the differences in electronic energies predicted by density functional theory.

Progress curves of aldol addition by EcHpcH. A reaction mixture to generate 2-fluoro-3,4-dihydroxybutyrate (FDHB) was prepared with 50 mM fluoropyruvate, 55 mM glycolaldehyde, 1 mM MgCl₂, and 0.01 mol% EcHpcH, to a total volume of 5 mL in 20 mM HEPES pH 7.5. The reaction was initiated with the aldolase and incubated at room temperature. Aliquots (600 μL) were removed at 0 min, 10 min, 30 min, 1 h, 2 h, 3.5 h, 6 h, 12 h, and 24 h and quenched with 30 wt. % H₂O₂ (60 μL, 30 min). Precipitated proteins were removed by centrifugation (20,000 × g, 5 min) and 100 μL D₂O was added. The extent of conversion was analyzed by quantitative ¹⁹F NMR (564 MHz), integrating the product peaks (major product: *anti*-FDHB, -196.0 ppm, dd; minor product: *syn*-FDHB, -201.6 ppm, dd) against fluoroacetate (-218.0 ppm, t). A time course was also carried out with 50 mM pyruvate instead of fluoropyruvate, to generate 3,4-dihydroxybutyrate (DHB). Aliquots (600 μL) were removed at 0.5, 1, 2.5, 5 min, and 10 min-24 h as above. The extent of conversion was analyzed by ¹H NMR (500 MHz), integrating the C-2 methylene peaks of DHB (2.28 ppm, dd; 2.22 ppm, dd) against the methyl peak of acetate (1.82 ppm, s).

Steady-state kinetic characterization of EcHpcH. Kinetic parameters for aldolase donors were determined by a discontinuous assay in which donor concentrations were varied for a fixed saturated concentration of acceptor. To characterize pyruvate addition and validate the assay, reactions of 100 μL volume in 20 mM HEPES pH 7.5 were prepared in a 96-well plate, containing 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, or 50 mM fluoropyruvate, 333 mM glycolaldehyde, 1 mM MgCl₂, and 50 nM EcHpcH (n = 5). For fluoropyruvate addition, reactions were carried with 5 μM EcHpcH with both 333 mM and 1 M glycolaldehyde (n = 3 for each aldehyde concentration). Since high levels of aldehyde were not observed to inhibit enzyme turnover, all six trials were used for analysis. The reactions were incubated for 10 min at room temperature after addition of aldolase, then were quenched with 10 μL of 30 wt. % H₂O₂. After 20 min, 11 μL of each sample was added to 89 μL acetonitrile. Precipitated proteins were removed by centrifugation (20,000 × g, 5 min), and 10 μL of the supernatant was added to 90 μL acetonitrile in a 384-well plate for overall 100-fold dilution of the analyte.

To prepare standard solutions of DHB and FDHB, reaction mixtures were prepared as described for progress curves and incubated at room temperature for 16 h. Decarboxylation was effected with 500 μL 30 wt. % H₂O₂ and 30 min of vigorous stirring. Precipitated proteins were removed (20,000 × g, 5 min) after which the samples were lyophilized with a vacuum concentrator. Ethanol (2 mL) was added to precipitate buffer salts which were removed by centrifugation (5,000 × g, 1 min). Ethanol was removed with a vacuum concentrator and the residue was re-dissolved in water (0.5 mL). The solutions were basified to pH ~13 with 3 M NaOH to hydrolyze lactones, then neutralized with 3 M HCl and brought to 1 mL final volume with water. The concentrations of DHB and FDHB were determined by quantitative ¹H NMR (500 MHz) or ¹⁹F NMR (564 MHz), using DSS or fluorouracil as standards respectively. Concentrated DHB standards (2.5, 5.0, 7.5, 10, 25, 50, and 100 μM) and FDHB standards (0.025, 0.05, 0.075, 0.1, 0.25, 0.5, and 1 mM) were prepared in HEPES buffer and diluted twice into acetonitrile as described for the assay samples (final standard curve concentrations 0.25-10 μM).

Samples from the 384-well plate were injected with an Agilent RapidFire High-Throughput MS System into the mass spectrometer module of an Agilent 6460 Triple Quadrupole LC-MS System. The following parameters were used for the RapidFire autosampler: HILIC type H1 cartridge; solvents A and B: 90% acetonitrile, 10 mM ammonium formate; solvent C: 40%

acetonitrile, 10 mM ammonium formate; flow rate of 1.5, 1.25, and 1.25 mL/min for pumps A, B, and C respectively; cycle durations: aspirate 0.6 s, load/wash 2.5 s, no extra wash, elute 5 s, reequilibrate 3 s. DHB was detected by tandem MS, using a collision energy of 15 V and following the 119.0 \rightarrow 59.0 m/z (ESI-) transition. For FDHB, the 137.0 \rightarrow 87.0 m/z (ESI-) transition was monitored. The standard curve was constructed with three replicate measurements of each standard. Data were fitted to the standard Michaelis-Menten equation for pyruvate addition and to the modified Michaelis-Menten equation with substrate inhibition for fluoropyruvate addition using Solver of Microsoft Excel.

EcHpcH sequence alignments. The sequence of EcHpcH (UniProt ID: B1IS70) was used as query for the UniProt BLAST tool to find homologous sequences in the UniRef50 database. Consideration of only bacterial sequences from known organisms provided a total of 459 representative homologs. Multiple sequence alignment was performed with MUSCLE (EMBL-EBI). The amino acid frequencies at residues of interest were analyzed in Microsoft Excel by extraction of the n -th character from the aligned strings.

Synthesis of 2-fluoro sugar acids. Fluoropyruvate addition reactions (0.15 mmol) were prepared in 3 mL total volume with 50 mM fluoropyruvate, 55 mM aldehyde (D-glyceraldehyde, L-glyceraldehyde, D-lactaldehyde, or L-lactaldehyde), 1 mM MgCl₂, and 0.01 mol% of the aldolase required to obtain the desired stereoisomer. The reaction catalyzed by GarL took place in 50 mM phosphate buffer pH 7.5 while the remainder used 20 mM HEPES pH 7.5. The reactions were incubated at room temperature without stirring for 16 h, then treated with 30 wt. % H₂O₂ (150 μ L) and stirred vigorously for 30 min. Particulates were removed by centrifugation (20,000 \times g, 5 min). The polyhydroxy 2-fluoroacid products were purified by semi-preparative HPLC on an Agilent 1200 Series HPLC System equipped with a SeQuant ZIC-HILIC column (5 μ m, 200 Å , 21.2 \times 150 mm). The chromatography conditions were as follows. Solvent A: 10% acetonitrile, 10 mM ammonium formate; solvent B: 90% acetonitrile, 10 mM ammonium formate; flow rate 5 mL/min; maximum pressure 400 bar; timetable (%A): 0% at 0 min, 0% at 8 min, 33% at 68 min, 60% at 80 min, 60% at 88 min, 70% at 93 min, 25% at 97 min, 0% at 105 min, 0% at 120 min; fraction collection: 384 fractions in deep 2 mL 96-well plates. Fractions were screened by LC-MS (ESI-) with a HILIC column as described in the final section, and those containing the product were combined and evaporated to 20 mL. The remaining sample was lyophilized in a vacuum concentrator to reveal white solids, consisting of ammonium formate and the ammonium pentonate salts. Yields were estimated by addition of DSS (0.15 mmol of DSS methyl protons) to the NMR samples. Due to co-elution of HEPES-derived impurities with the *syn*-product of D-glyceraldehyde, the sample was diluted with water (10 mL), added to Dowex 50WX8 (1 g, sodium form), and stirred for 30 min. The sample was filtered from the resin and lyophilized in a vacuum concentrator to provide the product as the sodium pentonate salt.

(2*S*,3*R*,4*R*)-2-fluoro-3,4,5-trihydroxypentanoate. ¹H NMR (600 MHz, DMSO-*d*₆) δ 4.82 (dd, J = 49.4, 1.7 Hz, 1H), 3.64 (ddd, J = 28.4, 8.8, 2.0 Hz, 1H), 3.56 (dt, J = 8.8, 4.6 Hz, 1H), 3.39-3.35 (m, 2H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ 171.6 (d, J = 22.3 Hz), 90.1 (d, J = 185.3 Hz), 71.9 (d, J = 19.0 Hz), 71.1 (d, J = 3.3 Hz), 63.7. ¹⁹F NMR (565 MHz, DMSO-*d*₆) δ -200.9 (dd, J = 50.2, 29.0 Hz).

(2*S*,3*S*,4*R*)-2-fluoro-3,4,5-trihydroxypentanoate. ¹H NMR (600 MHz, DMSO-*d*₆) δ 4.51 (dd, J = 49.5, 5.3 Hz, 1H), 3.70 (ddd, J = 15.4, 5.3, 1.9 Hz, 1H), 3.51 (td, J = 6.4, 1.5 Hz, 1H), 3.35 (dd, J = 10.6, 6.9 Hz, 1H), 3.07 (dd, J = 10.6, 5.9 Hz, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆) δ

173.0 (d, $J = 20.5$ Hz), 90.8 (d, $J = 184.6$ Hz), 71.2 (d, $J = 5.5$ Hz), 70.5 (d, $J = 22.9$ Hz), 62.6. ^{19}F NMR (565 MHz, DMSO- d_6) δ -184.0 (dd, $J = 49.5, 15.5$ Hz).

(2*S*,3*S*,4*S*)-2-fluoro-3,4,5-trihydroxypentanoate. ^1H NMR (600 MHz, DMSO- d_6) δ 4.66 (dd, $J = 50.3$ Hz, 2.4 Hz, 1H), 3.75 (ddd, $J = 20.9, 6.4, 2.6$ Hz, 1H), 3.54 (q, $J = 5.5$ Hz, 1H), 3.47 (dd, $J = 11.1, 3.9$ Hz, 1H), 3.37 (dd, $J = 11.1, 6.0$ Hz, 1H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 172.5 (d, $J = 20.7$ Hz), 92.1 (d, $J = 182.1$ Hz), 72.49 (d, $J = 21.1$ Hz), 72.36 (d, $J = 6.6$ Hz), 63.4. ^{19}F NMR (565 MHz, DMSO- d_6) δ -188.3 (dd, $J = 50.3, 20.9$ Hz).

(2*S*,3*R*,4*R*)-2-fluoro-3,4-dihydroxypentanoate. ^1H NMR (600 MHz, DMSO- d_6) δ 4.74 (dd, $J = 50.2, 1.4$ Hz, 1H), 3.47 (m, 1H), 3.42 (ddd, $J = 28.0, 8.7, 1.3$ Hz, 1H), 1.07 (d, $J = 6.0$ Hz, 3H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 173.5 (d, $J = 21.0$ Hz), 90.5 (d, $J = 185.8$ Hz), 76.2 (d, $J = 18.7$ Hz), 66.3 (d, $J = 3.8$ Hz), 21.0. ^{19}F NMR (565 MHz, DMSO- d_6) δ -200.5 (dd, $J = 50.3, 28.1$ Hz).

(2*S*,3*S*,4*R*)-2-fluoro-3,4-dihydroxypentanoate. ^1H NMR (600 MHz, DMSO- d_6) δ 4.45 (dd, $J = 49.6, 5.7$ Hz, 1H), 3.69 (qd, $J = 6.3, 2.7$ Hz, 1H), 3.39 (ddd, $J = 14.3, 5.7, 2.9$ Hz, 1H), 1.05 (d, $J = 6.5$ Hz, 3H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 172.8 (d, $J = 20.1$ Hz), 90.4 (d, $J = 184.6$ Hz), 74.8 (d, $J = 22.5$ Hz), 66.1 (d, $J = 5.4$ Hz), 20.4. ^{19}F NMR (565 MHz, DMSO- d_6) δ -186.3 (dd, $J = 49.7, 14.2$ Hz).

(2*S*,3*S*,4*S*)-2-fluoro-3,4-dihydroxypentanoate. ^1H NMR (600 MHz, DMSO- d_6) δ 4.57 (dd, $J = 50.6, 3.8$ Hz, 1H), 3.66 (quint, $J = 6.2$ Hz, 1H), 3.52 (ddd, $J = 19.2, 5.8, 3.9$ Hz, 1H), 1.03 (d, $J = 6.3$ Hz, 3H). ^{13}C NMR (151 MHz, DMSO- d_6) δ 172.6 (d, $J = 20.7$ Hz), 91.7 (d, $J = 184.3$ Hz), 75.8 (d, $J = 20.7$ Hz), 67.2 (d, $J = 6.0$ Hz), 19.9. ^{19}F NMR (565 MHz, DMSO- d_6) δ -189.2 (dd, $J = 50.9, 19.1$ Hz).

Synthesis of methyl (2*S*,3*S*)-2-fluoro-3,4-*O*-isopropylidene-3,4-dihydroxybutanoate. In a round-bottom flask, sodium fluoropyruvate (1.02 g, 7 mmol), glycolaldehyde dimer (462 mg, 7.7 mmol monomer), MgCl_2 (140 μL of 1 M), and EcHpcH V118F (21 mg, 0.01 mol%, exchanged into HEPES storage buffer without glycerol) were mixed in a total volume of 140 mL of 20 mM HEPES pH 7.5. The reaction proceeded at room temperature without stirring for 16 h, then was terminated with 30 wt. % H_2O_2 (7 mL) and vigorous stirring for 30 min. Catalase (15 mg) was added and stirring was continued for 15 min to decompose excess H_2O_2 . The solution was concentrated *in vacuo* to ~ 20 mL, filtered, and lyophilized in a vacuum concentrator to a viscous residue which was dissolved in methanolic HCl (80 mL of 1.5 M). Precipitates were filtered off and the solution was heated to reflux for 16 h. The solution was concentrated to ~ 3 mL, then acetone (35 mL) was added, precipitates were filtered off, and 2,2-dimethoxypropane (5 mL) and *p*-toluenesulfonic acid monohydrate (50 mg) were added. The reaction mixture was stirred at room temperature for 16 h. The solvent was removed *in vacuo* and replaced with diethyl ether (50 mL). The organic layer was washed with saturated NaHCO_3 (3 x 50 mL) and the aqueous layers were back-extracted with diethyl ether (3 x 20 mL). The combined organic layers were washed with water (50 mL) and brine (50 mL), then dried over MgSO_4 and concentrated to provide the product as a yellow oil (348 mg, 26%). ^1H NMR (400 MHz, CDCl_3) δ 4.94 (dd, $J = 48.8, 4.4$ Hz, 1H), 4.46 (dtd, $J = 18.3, 5.7, 4.4$ Hz, 1H), 4.10-4.07 (m, 2H), 1.44 (s, 3H), 1.37 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 167.7 (d, $J = 23.1$ Hz), 110.5, 87.9 (d, $J = 190.1$ Hz), 75.0 (d, $J = 23.7$ Hz), 64.6 (d, $J = 5.0$ Hz), 52.6, 26.2, 25.2. ^{19}F NMR (376 MHz, CDCl_3) δ -199.3 (dd, $J = 48.8, 18.3$ Hz).

Synthesis of methyl (2*S*,3*S*)-2-fluoro-3-hydroxy-4,4-dimethoxybutanoate. In a round-bottom flask, sodium fluoropyruvate (1.02 g, 7 mmol), dimethoxyacetaldehyde solution (60 wt. %, 1.34 mL, 7.7 mmol), MgCl_2 (140 μL of 1 M), and SwHpcH1 (21 mg, 0.01 mol%) were mixed

in a total volume of 140 mL of 20 mM HEPES pH 7.5. The reaction proceeded at room temperature without stirring for 16 h. Decarboxylation, concentration, and esterification in methanol were carried out as in the previous synthesis. Methanol was removed under reduced pressure and EtOAc (50 mL) was added. The organic layer was washed with saturated NaHCO₃ (3 x 50 mL) and the aqueous layers were back-extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with water (50 mL) and brine (50 mL), then dried over Na₂SO₄ and concentrated to provide the product as a pale yellow oil (230 mg, 17%). ¹H NMR (400 MHz, CDCl₃) δ 5.10 (dd, *J* = 48.0, 2.4 Hz, 1H), 4.51 (d, *J* = 7.2 Hz, 1H), 4.13-4.02 (m, 1H), 3.81 (s, 3H), 3.47 (s, 3H), 3.43 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 167.7 (d, *J* = 23.7 Hz), 102.7 (d, *J* = 8.2 Hz), 89.0 (d, *J* = 189.7 Hz), 71.9 (d, *J* = 19.9 Hz), 54.8, 54.6, 52.3. ¹⁹F NMR (376 MHz, CDCl₃) δ -204.2 (dd, *J* = 48.1, 23.8 Hz).

Synthesis of dimethyl (2S,3R)-2-fluoro-3-hydroxysuccinate. A solution of glyoxylic acid monohydrate (506 mg, 5.5 mmol) in water (10 mL) was neutralized with 3 M NaOH. A round-bottom flask was charged with sodium fluoropyruvate (730 mg, 5 mmol), the solution of sodium glyoxylate, MgCl₂ (100 μL of 1 M), GarL (15 mg, 0.01 mol%), and 50 mM sodium phosphate pH 7.5 to a total volume of 100 mL. The reaction proceeded at room temperature without stirring for 16 h. Decarboxylation (5 mL 30 wt. % H₂O₂, 10 mg catalase), concentration, esterification in methanol, and extractive workup (EtOAc) were performed as in the previous synthesis to provide the product as a colorless oil (252 mg, 28%). ¹H NMR (400 MHz, CDCl₃) δ 5.26 (dd, *J* = 47.0, 1.8 Hz, 1H), 4.69 (dd, *J* = 30.5, 6.1 Hz, 1H), 3.89 (s, 3H), 3.88 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.4 (d, *J* = 3.0 Hz), 166.8 (d, *J* = 24.2 Hz), 88.8 (d, *J* = 193.1 Hz), 71.2 (d, *J* = 20.8 Hz), 53.4, 52.8. ¹⁹F NMR (376 MHz, CDCl₃) δ -206.1 (dd, *J* = 47.1, 30.4 Hz).

Synthesis of dimethyl (2S,3S)-2-fluoro-3-hydroxysuccinate. The previous procedure was repeated with wild-type EchHpcH and 20 mM HEPES pH 7.5 as buffer. After decarboxylation, esterification, and extraction steps, the product was revealed as a yellow oil (221 mg, 25%). ¹H NMR (400 MHz, CDCl₃) δ 5.27 (dd, *J* = 47.3, 2.3 Hz, 1H), 4.75-4.68 (m, 1H), 3.85 (s, 3H), 3.84 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.1 (d, *J* = 9.3 Hz), 166.7 (d, *J* = 23.7 Hz), 89.7 (d, *J* = 194.2 Hz), 71.4 (d, *J* = 21.9 Hz), 53.4, 52.7. ¹⁹F NMR (376 MHz, CDCl₃) δ -200.8 (dd, *J* = 47.3, 22.9 Hz).

Sample preparation of 3-fluoro-2-oxoacids and 2-fluoroacids for LC-MS. 3-fluoro-2-oxoacids (50 mM maximum) were prepared by the method of the fluoropyruvate addition assay using EchHpcH as described above. For analysis of the oxoacids, 10 μL of each oxoacid sample was added to 80 μL water and 10 μL of 100% w/v trichloroacetic acid (TCA) to precipitate proteins. Samples were clarified by centrifugation (20,000 × g, 5 min). For analysis of the decarboxylated 2-fluoroacids, 10 μL of each oxoacid sample was added to 1 μL 30 wt. % H₂O₂ and 79 μL water and incubated for 20 min. TCA (10 μL) was added, then samples were clarified (20,000 × g, 5 min). All sample supernatants were diluted tenfold into acetonitrile and analyzed by LC-MS as described in the final section.

Sample preparation of 3-fluoro-2-hydroxyacids for LC-MS. 3-Fluoro-2-oxoacids formed from all twelve aldehydes in this study were screened for their ability to be reduced by lactate dehydrogenase. To 30 μL of each oxoacid sample generated as described above was added 10 μL of 100 mM EDTA pH 9.0. After 10 min, NADH (20 μL, 100 mM stock in 10 mM NaOH) and LDH (10 μL of 2 mg/mL stock) were added. Reduction proceeded at room temperature for 16 h overnight, then 200 μL water and 30 μL TCA were added. Samples were clarified (20,000 × g, 5 min), diluted tenfold into acetonitrile, and analyzed by LC-MS as described in the final section.

Sample preparation of 3-fluoro-2-aminoacids for LC-MS. 3-Fluoro-2-oxoacids formed from all twelve aldehydes in this study were screened for their ability to be converted to amino acids by transaminase VfAT. To 30 μL of each oxoacid sample was added 20 μL of 20 mM HEPES pH 7.5 with 100 μM PLP, 5 μL of 1 M (*S*)- α -methylbenzylamine (SMBA, stock solution in 50% DMSO), and 45 μL of VfAT (stock 5.0 mg/mL). The final transamination reactions (containing <15 mM oxoacids, 20 mM PLP, 50 mM SMBA, and 225 μg VfAT) were incubated at room temperature for 16 h overnight, then added to 170 μL water and 30 μL TCA. Samples were clarified by centrifugation (20,000 $\times g$, 5 min), then evaporated in a vacuum concentrator to one-tenth of the original volume ($\sim 30 \mu\text{L}$). Samples were diluted tenfold into acetonitrile, clarified again, and analyzed by LC-MS as described in the final section, with the modification that positive ion mode (ESI+) was used.

2F-2d-d-Ara, 2F-2d-d-Lyx, and 2F-2,5-dd-d-Ara. The salts of (2*S*,3*R*,4*R*)-2-fluoro-3,4,5-trihydroxypentanoate, (2*S*,3*S*,4*R*)-2-fluoro-3,4,5-trihydroxypentanoate, and (2*S*,3*R*,4*R*)-2-fluoro-3,4-dihydroxypentanoate were converted to fluorosugars by methyl esterification, acetonide protection (regioselectivity undetermined), reduction to an alcohol at C1, and partial oxidation to the aldehyde. DMSO-*d*₆ was removed from the respective NMR samples with a vacuum concentrator. The solids were dissolved in 1 M methanolic HCl (1.5 mL) and insoluble materials were pelleted by centrifugation (5,000 $\times g$, 1 min). The supernatants were transferred to sealed vials and heated at 50 $^{\circ}\text{C}$ for 16 h. Then, the solvent was removed *in vacuo* and exchanged for acetone (1.3 mL). Insoluble materials were again removed by centrifugation (5,000 $\times g$, 1 min), then 2,2-dimethoxypropane (0.2 mL) and *p*-toluenesulfonic acid monohydrate (1 mg) were added. Samples were stirred at room temperature for 16 h. Acetone was removed *in vacuo* and replaced with CH_2Cl_2 (2 mL). The organic layers were washed with saturated NaHCO_3 (5 mL) and brine (5 mL) and dried over MgSO_4 . The samples, containing acetonides of pentonate esters, were cooled to 0 $^{\circ}\text{C}$ and DIBAL-H solution (0.2 mL, 1 M in CH_2Cl_2) was added dropwise to each. After warming to room temperature over 2 h, the samples were diluted with EtOAc (7 mL) and washed with saturated Rochelle salt (6 mL, agitated every 5 min for 30 min until clear phase separation). The organic solvent was removed under a stream of nitrogen. To each residue was added 45 wt.% stabilized IBX (100 mg) in 1.5 mL DMSO, and the oxidation proceeded with stirring at room temperature for 16 h. DMSO was removed in a vacuum concentrator, and each residue was stirred for 30 min in 1 M HCl (1 mL) for deprotection of acetonide groups, providing 2-fluoro-2-deoxy-D-arabinose, 2-fluoro-2-deoxy-D-lyxose, and 2-fluoro-2,5-dideoxy-D-arabinose in solution.

2F-2d-l-Ery. Methyl (2*S*,3*S*)-2-fluoro-3,4-*O*-isopropylidene-3,4-dihydroxybutanoate was converted into a fluorosugar by reduction to an alcohol and partial oxidation to an aldehyde at C-1. A dry round-bottom flask under nitrogen was charged with anhydrous CH_2Cl_2 (5 mL) and the starting ester (121 mg, 0.63 mmol), then cooled to 0 $^{\circ}\text{C}$. DIBAL-H (1.8 mL, 1 M solution in CH_2Cl_2 , 1.8 mmol, 2.9 eq.) was added dropwise. The reaction was stirred for 15 min at 0 $^{\circ}\text{C}$, then for 4 hours while warming to room temperature. The reaction mixture was quenched with EtOAc (20 mL), then stirred rapidly with saturated potassium sodium tartrate (Rochelle's salt, 20 mL) until clear layer separation was achieved (~ 30 min). The organic layer was separated and the aqueous layer was extracted with EtOAc (3 \times 20 mL). The combined organic layers were washed with water (30 mL) and brine (30 mL), then dried over Na_2SO_4 , and concentrated to provide a pale yellow oil containing (2*S*,3*R*)-3-fluoro-1,2-*O*-isopropylidene-1,2,4-butanetriol; ^{19}F NMR (377 MHz, CDCl_3): δ -197.6 (dtd). The crude protected triol was used directly in a Swern oxidation. At -78 $^{\circ}\text{C}$, DMSO (158 μL , 2.2 mmol, 10 eq.) in CH_2Cl_2 (0.5 mL) was added to oxalyl chloride (0.55 mL of 2 M solution in CH_2Cl_2 , 1.1 mmol, 5 eq.) and stirred for 30 min. The starting material was

slowly added as a solution in CH₂Cl₂ (0.75 mL) followed by triethylamine (0.46 mL, 3.3 mmol, 15 eq.). The reaction mixture was stirred for 15 min at -78 °C, then for 2 h while warming to room temperature. The reaction was carefully quenched with water (1 mL). Saturated NH₄Cl (20 mL) was added and the product was extracted with EtOAc (5 × 15 mL). The combined organic layers were washed with saturated NaHCO₃ (30 mL) and brine (30 mL), dried over Na₂SO₄, and concentrated to provide a crude orange oil containing (2*S*,3*S*)-2-fluoro-3,4-*O*-isopropylidene-3,4-dihydroxybutanal or its aldehyde hydrate. The oil was stirred for 30 min in 1 M HCl (1 mL) for deprotection of the acetonide, providing 2-fluoro-2-deoxy-L-erythrose in solution.

3F-3d-D-Ery. Methyl (2*S*,3*S*)-2-fluoro-3-hydroxy-4,4-dimethoxybutanoate was converted into a fluorosugar by reduction to an alcohol at C-1 and deprotection of the masked aldehyde at C-4. Using the procedure above, the starting ester (230 mg, 1.17 mmol) was reduced with 5 mL of DIBAL-H (5 mmol, 4.3 eq.). Workup gave a yellow oil containing (2*S*,3*S*)-2-fluoro-4,4-dimethoxy-1,3-butanediol; ¹⁹F NMR (377 MHz, CDCl₃): δ -199.1 (dtd). The oil was stirred for 30 min in 1 M HCl (1 mL) for deprotection of the dimethyl acetal, providing 3-fluoro-3-deoxy-D-erythrose in solution.

Sample preparation of fluorosugars for LC-MS. A 10 μL aliquot of each solution of fluorosugar in 1 M HCl was added to 90 μL of 100 mM *O*-benzylhydroxylamine (OBHA) and allowed to sit at room temperature for 16 h. Samples were clarified by centrifugation (20,000 × g, 5 min) and analyzed by LC-QTOF as described in the final section, with the modifications that positive ion mode (ESI+) was used, and that 0.1% formic acid was used as solvent A.

High resolution LC-MS analysis. Characterization of all compounds by electrospray ionization (ESI) and high resolution LC-MS was performed on an Agilent 6530 Accurate-Mass Q-TOF LC-MS in negative ion mode unless otherwise stated. For hydrophobic compounds (fluoroesters and OBHA-derivatized sugars), the LC-MS was equipped with an Agilent Poroshell 120 EC-C18 column (2.7 μm, 2.1 × 50 mm). Reversed-phase chromatography conditions were as follows. Solvent A: 10 mM ammonium bicarbonate; solvent B: acetonitrile; flow rate 0.6 mL/min; maximum pressure 600 bar; timetable (%B): 0% at 0 min, 3% at 2 min, 100% at 5.2 min, 100% at 6 min, 5% at 6.2 min, 5% at 7.2 min. For hydrophilic compounds (all compounds with a free carboxylate), the LC-MS was equipped with an Ascentis Express HILIC column (2 μm, 2.1 × 50 mm). HILIC-mode chromatography conditions were as follows. Solvent A: 10% acetonitrile, 10 mM ammonium formate; solvent B: 90% acetonitrile, 10 mM ammonium formate; flow rate 0.5 mL/min; maximum pressure 800 bar; timetable (%A): 0% at 0 min, 0% at 0.8 min, 33% at 3.5 min, 70% at 4 min, 70% at 4.5 min, 25% at 5 min, 0% at 5.5 min, 0% at 6 min. LC-MS samples of purified compounds were prepared as 1 mM solutions in the equilibration solvent.

2.3. Results and discussion

Five pyruvate aldolases from the Type II HpcH family (PF03328) were chosen as candidates for the desired fluoropyruvate aldol reaction (Figure 2.2). These included GarL from the glucarate/galactarate catabolism pathway of *E. coli* [31,32], RhmA from the rhamnose catabolism pathway of *E. coli* [33-36], EcHpcH from the homoprotocatechuate degradation pathway of *E. coli* C strain [37-40], plus the poorly characterized aldolases SwHpcH1 and SwHpcH2 from *Sphingomonas wittichi*. The *E. coli* enzymes were chosen for their extensive biochemical characterization, while the two *S. wittichi* enzymes were included after a high-throughput study of pyruvate aldolases had reported potential stereocomplementary before [41]. The gene encoding

EcHpcH is curiously absent from common B and K-12 lab strains of *E. coli*. Therefore, only GarL and RhmA were amplified from the genome of *E. coli* DH10B (K-12 based strain) while EcHpcH, SwHpcH1, and SwHpcH2 were obtained as codon-optimized synthetic genes. Cloning, expression, and purification of the recombinant N-terminal His₁₀-tagged proteins from *E. coli* provided the five aldolases in moderate to excellent yields (5-30 mg/L culture), which were visualized by SDS-PAGE analysis (Figure 2.3).

Fluoropyruvate aldol addition assays were generally conducted by incubation of fluoropyruvate (50 mM) with aldehyde acceptor (55 mM), buffer (50 mM HEPES or phosphate, pH 7.5) and MgCl₂ (1 mM) for an overnight period. The β -fluoro- α -ketoacids were then subjected to oxidative decarboxylation with H₂O₂, a process that simplifies product analysis by converting reactants and products into α -fluoroacids and also precipitates enzymes from the reaction. The samples were analyzed by ¹⁹F NMR, in which the extent of conversion could be measured from the relative integrals of α -fluoroacid products and fluoroacetate. Furthermore, ¹⁹F NMR allowed for discernment between the diastereomeric *syn* and *anti* configurations of vicinal fluoro-hydroxy moieties in the products, with *syn* diastereomers known to appear upfield with typically larger ³J_{F-H} coupling constants due to hyperconjugation effects [16,17].

We initially screened the addition of fluoropyruvate to three polar aldehydes—glycolaldehyde (GA), dimethoxyacetaldehyde (DMA), and glyoxylate (GX) in either HEPES or phosphate buffer, and saw promising conversions with most enzyme-substrate-buffer combinations (Figure 2.4). GarL and SwHpcH1 were observed to provide higher conversions with phosphate buffer. This phenomenon is consistent with reports that phosphate plays a role in the catalytic mechanism of some HpcH family aldolases [32,42]. No significant differences between HEPES and phosphate conditions were observed for RhmA, EcHpcH, and SwHpcH2. Proceeding to investigate an expanded substrate scope for the fluoropyruvate aldol addition, the ¹⁹F NMR assays were extended to a panel of twelve aldehydes comprising chiral polar, achiral polar, and nonpolar groups (Figure 2.5 and Figure 2.6). Phosphate buffer was used for GarL and SwHpcH1, while HEPES buffer was used for RhmA, EcHpcH, and SwHpcH2. High to quantitative conversions (80-100%) were achieved for nearly all combinations of an aldolase with a polar aldehyde. Nonpolar aldehydes were converted less efficiently (10-70%), with a preference for the three-carbon propionaldehyde (Pro) over other alkyl chain lengths. Heteroaromatic aldehydes such as 2-pyridinecarboxaldehyde (2PCA) were converted efficiently, indicating the possible importance of hydrogen-bond acceptors. In the absence of any aldehyde, the aldolases were observed to catalyze trace amounts of fluoropyruvate dimerization. All aldol products, pre- and post-decarboxylation, were also characterized by high resolution LC-MS (Figure 2.7).

It was hypothesized and later confirmed that the diastereomeric ratio of *syn* and *anti* products from the fluoropyruvate aldol reaction arose from variable configuration of the hydroxy group, while the fluorine group remained in a fixed configuration. This is consistent with reports of HpcH family enzymes lacking facial discrimination at the aldehyde acceptor [33,39]. Interestingly, the diastereomeric ratio had great variation in our screen of five enzymes with twelve aldehydes, appearing dictated by either the aldehyde or the enzyme in different cases. For example, the L-configured chiral aldehydes L-glyceraldehyde and L-lactaldehyde led to a strong preference for the *anti* isomer regardless of enzyme. No consistent stereochemistry was induced by D-configured chiral aldehydes. However, GarL and EcHpcH were strikingly stereocomplementary when using GX as acceptor (99% *syn* and 92% *anti*, respectively). SwHpcH1 and SwHpcH2 did not produce stereocomplementary results with fluoropyruvate, in contrast to the original motivation for their

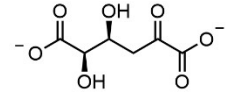
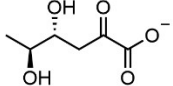
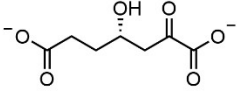
	UniProt ID	Protein annotation and synonyms	Native aldehyde	Native adduct
GarL	P23522	5-keto-4-deoxy-D-glucarate aldolase DDGA, DDG aldolase	tartronate semialdehyde	
RhmA	P76469	2-keto-3-deoxy-L-rhamnonate aldolase KDRA, KDR aldolase, YfaU	L-lactaldehyde	
EcHpcH	B1IS70	4-hydroxy-2-oxo-heptane-1,7-dioate aldolase HKHD aldolase, HHED aldolase, Hpal	succinate semialdehyde	
SwHpcH1	A5VH82			
SwHpcH2	A5VAX1			

Figure 2.2. Enzyme information for Type II HpcH family pyruvate aldolases used in this study.

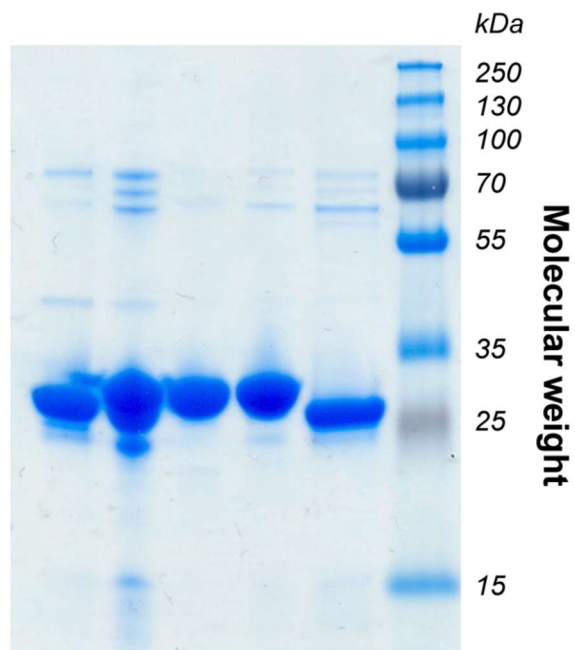


Figure 2.3. SDS-PAGE gel of aldolases. From left to right: GarL, RhmA, EcHpcH, SwHpcH1, SwHpcH2, molecular weight marker.

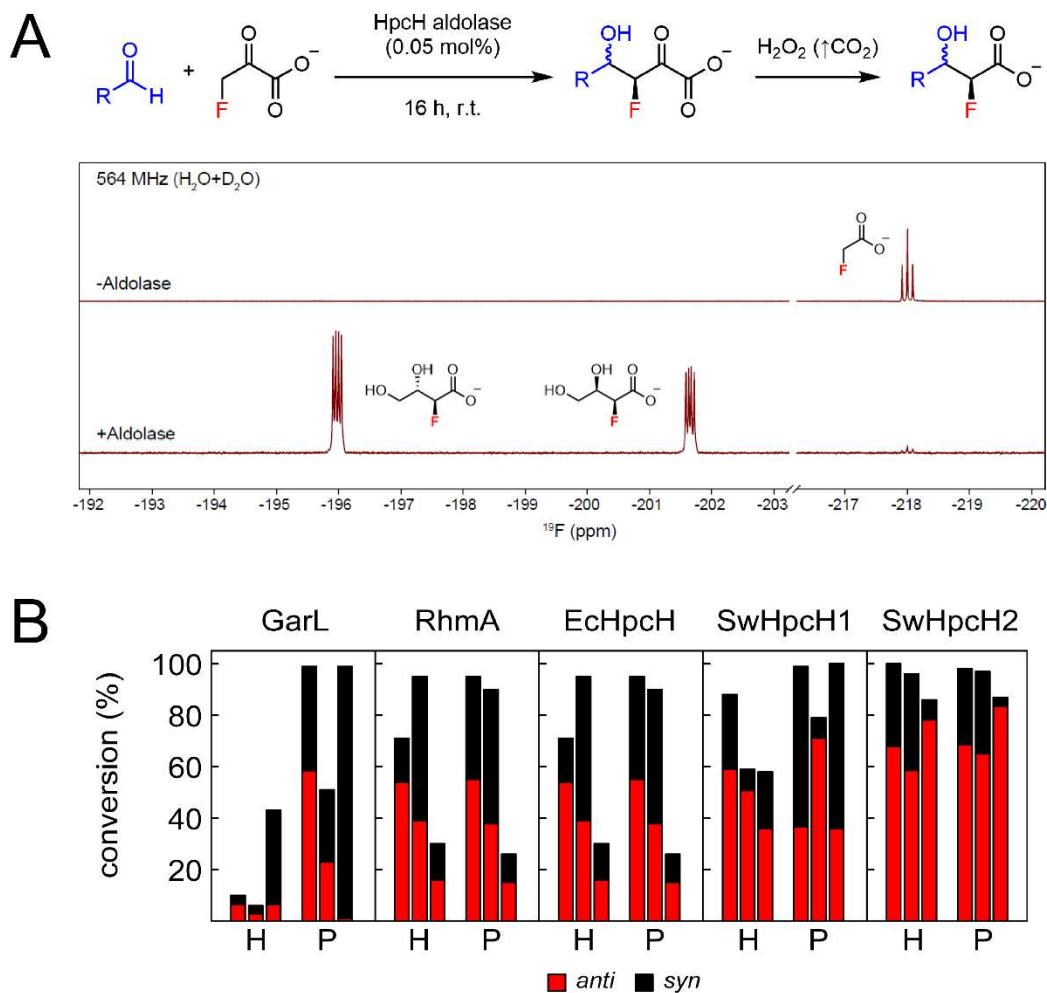


Figure 2.4. Screening for fluoropyruvate aldol adducts. A) Reaction scheme and example of ^{19}F NMR analysis for GA adduct. B) Buffer screen to compare HEPES (H) and phosphate (P). Aldehydes used in each group of three bars: GA, DMA, GX.

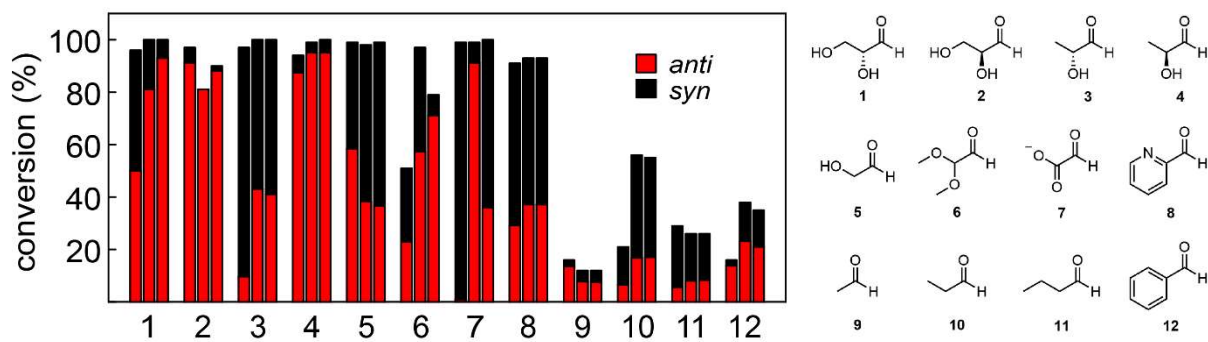


Figure 2.5. Substrate scope of fluoropyruvate aldol addition with 12 aldehydes of various properties. Enzymes used in each group of three bars: GarL, RhmA, EchHpcH. A preference for polar aldehydes is observed. The diastereoselectivity is highly dependent on both the aldehyde and the enzyme homolog.

	<i>anti</i> products			<i>syn</i> products		
	δ (ppm)	$^2J_{F-H}$ (Hz)	$^3J_{F-H}$ (Hz)	δ (ppm)	$^2J_{F-H}$ (Hz)	$^3J_{F-H}$ (Hz)
1	-191.0	49.4	20.6	-203.6	48.9	29.3
2	-194.7	49.9	25.2	-199.4	49.6	27.5
3	-192.8	49.3	22.2	-202.8	49.1	28.8
4	-195.3	50.6	24.7	-201.6	49.0	28.2
5	-196.0	50.5	25.2	-201.6	49.3	27.1
6	-197.7	49.7	26.5	-201.9	48.9	28.7
7	-189.0	49.7	25.0	-197.9	48.4	34.7
8	-193.2	50.8	23.5	-198.0	48.9	26.5
9	-197.5	50.1	24.3	-200.1	52.2	26.8
10	-197.9	51.6	27.2	-199.7	49.4	26.4
11	-198.2	52.0	27.8	-199.2	49.5	26.2
12	-196.0	49.4	24.5	-197.3	51.9	27.2

Figure 2.6. ^{19}F NMR data for all decarboxylated fluoropyruvate adducts. Numbers of aldehydes used correspond to those shown in Figure 2.5.

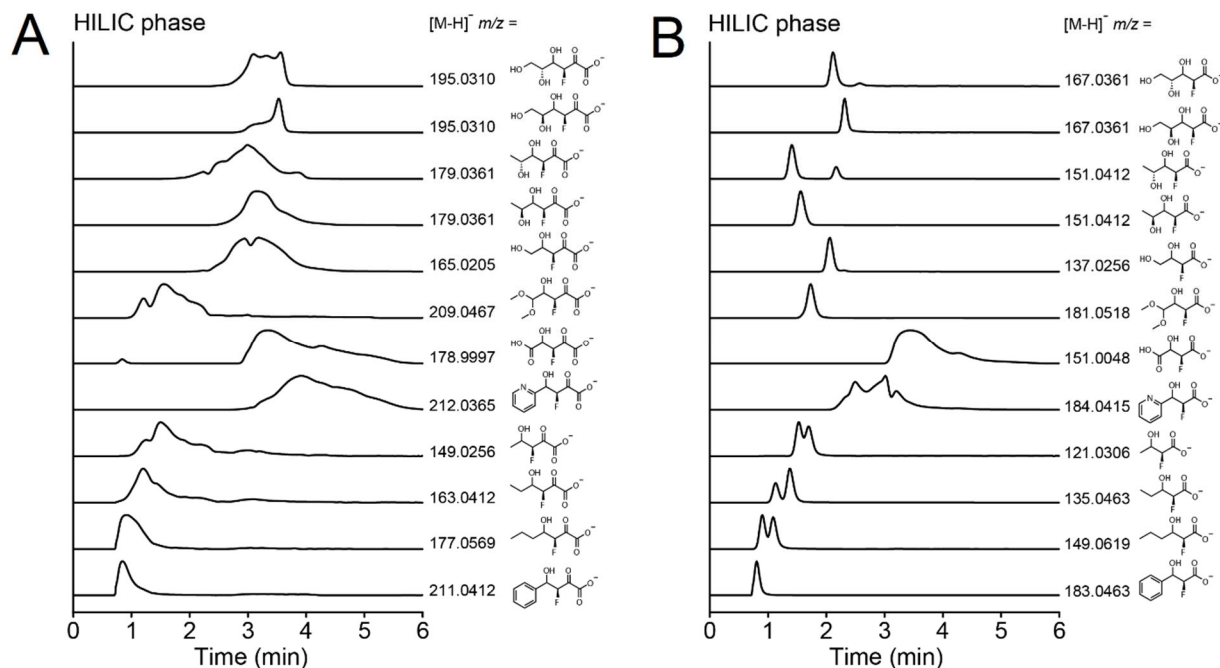


Figure 2.7. High resolution LC-MS of products from fluoropyruvate. The corresponding masses were not detected in blank injection controls. The intensity of each chromatogram is scaled to the largest in each set. A) 3-Fluoro-2-oxoacid adducts. Poor peak shape for most products is likely due to multiple speciation from ketone hydration or cyclization. B) 2-Fluoroacids post-decarboxylation. Poor peak shape for some products may be attributable to complex hydrogen-bonding interactions with the stationary phase.

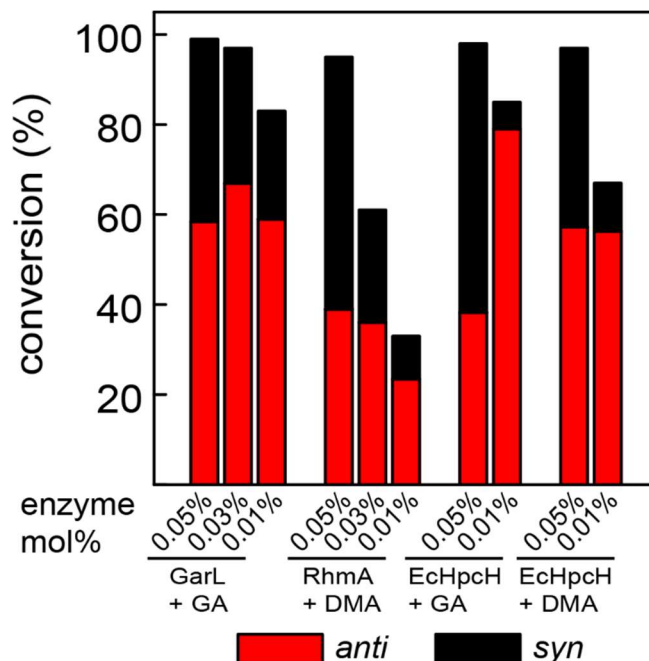


Figure 2.8. Altered diastereoselectivity in fluoropyruvate aldol reactions observed at lower enzyme loading of HpcH.

inclusion in this study. The effect of catalyst loading on relative stereochemistry was also investigated in reactions with GA and DMA (Figure 2.8). The loading of GarL or EcHpcH was varied for GA reactions, and the loading of RhmA or EcHpcH was varied for DMA reactions. These aldehydes led to mixtures of 39-59% *anti* composition using the standard 0.05 mol% enzyme loading, but with 0.01 mol% enzyme loading the diastereoselectivity improved to 71-93% *anti* at the expense of qualitative conversions. These results suggest the interplay of kinetic and thermodynamic control, striking an easily perturbed balance between the kinetically favored *anti* product and the thermodynamically favored *syn* product [43]. Diastereomeric ratios would thus not only be affected by the choice of aldehyde and enzyme homolog, but also by the enzyme loading and reaction time, and these factors might be tuned to achieve the desired hydroxyl group configuration.

With a promising system for fluoropyruvate aldol addition in hand, we sought to clarify the stereochemical course of the reaction. The new fluorine and hydroxyl stereocenters set by fluoropyruvate addition led us to confirm both the relative and absolute stereochemistry of products by indirect NMR-based methods. The relative stereochemistry, which we had preliminarily assigned on the basis of ^{19}F chemical shift and $^3J_{\text{F-H}}$ coupling constants, was determined with a nitric acid oxidation strategy that allowed us to compare several enzymatic products to a synthetic standard (Figure 2.9). First, diethyl *L*-tartrate was treated with DAST and then hydrolyzed to provide an *anti* diastereomer of 2-fluoro-3-hydroxysuccinate (*anti*-FHS), specifically the (2*S*,3*S*)-stereoisomer, which displayed a single ^{19}F NMR peak. This served as a standard for relative stereochemistry, as the deoxyfluorination of tartrate esters with sulfuranes is known to proceed by inversion of configuration [44]. The reaction of fluoropyruvate with GX catalyzed by GarL produced a single ^{19}F NMR peak that differed from the standard, which was thus assigned as the *syn* diastereomer of 2-fluoro-3-hydroxysuccinate (*syn*-FHS). Using EcHpcH instead resulted in both peaks, with *anti*-FHS as major product. Products from the reaction of fluoropyruvate with GA and DMA, catalyzed by GarL and RhmA respectively, were treated with

nitric acid to oxidize the C-4 position to a carboxylic acid, thus converting the species into *anti*- or *syn*-FHS as well. The upfield aldol products were oxidized to *syn*-FHS and the downfield aldol products to *anti*-FHS, thus corroborating our assignment of diastereomers.

With regards to the absolute configuration of fluoropyruvate aldol products, we first observed that only one of the two nascent stereogenic centers could be variable, since reactions with chiral aldehydes (D- or L-configured glyceraldehyde or lactaldehyde) gave no more than two peaks in ^{19}F NMR. To ascertain if the same behavior held true for achiral aldehydes, the β -fluoro- α -ketoacid aldol adducts with GA, DMA, GX, and Pro were subjected to stereoselective enzymatic reduction by L-lactate dehydrogenase, producing the corresponding β -fluoro- α -hydroxyacids (*Figure 2.10*). From this reaction, only two products were observed by ^{19}F NMR out of the four possible diastereomers. Assuming a fixed L- or (2*R*)-hydroxy group, analysis of $^3J_{\text{F-H}}$ coupling constants for these reduced adducts (larger values for *syn* relationships of fluoro and hydroxy) leads to the determination that fluorine is (3*S*)-configured in the original adducts (*Figure 2.11A*). Later on, α -fluoroacids derived from the four chiral aldehydes were purified, enabling ^1H NMR and analysis of $^3J_{\text{H-H}}$ coupling constants (smaller values for *syn* relationships of electronegative groups) (*Figure 2.11B*). This data further supported our conclusion that the fluoropyruvate aldol reaction is highly selective for (3*S*)-fluorine.

Progress curves over 24 h were constructed for the EcHpcH-catalyzed addition of pyruvate or fluoropyruvate to glycolaldehyde, in order to compare the kinetic behavior of the two donors (50 mM donor, 55 mM glycolaldehyde, 0.05 mol% EcHpcH). It was observed that the fluoropyruvate addition slowly approached conversion overnight, while native pyruvate addition under identical conditions was rapidly completed within a few minutes (*Figure 2.12A*). To dissect the factors that could account for this rate difference, steady-state Michaelis-Menten parameters were measured by a discontinuous assay with mass spectrometry detection, varying the concentration of pyruvate or fluoropyruvate (1-50 mM) and keeping the concentration of glycolaldehyde saturated (333 mM or 1 M) in accordance with its previously reported K_m value [39]. Enzymatically synthesized (*R*)-3,4-dihydroxybutyrate (DHB, stereochemistry assumed) and (2*S*,3*S*)-2-fluoro-3,4-dihydroxybutyrate (FDHB) were used to construct standard curves of the expected post-decarboxylation products. The data was fit to the Michaelis-Menten equation for pyruvate, and to the Michaelis-Menten equation with substrate inhibition for fluoropyruvate due to the observation of rates decreasing at higher donor concentrations (*Figure 2.12B*). We found that the fluorine substitution results in approximately a 500-fold decrease in k_{cat} which is largely responsible for the slow rate of fluoropyruvate addition, while the substrate inhibition phenomenon has a smaller contribution to the overall rate defect. Relative to the k_{cat} difference, the K_m difference between pyruvate and fluoropyruvate was small, and in fact the K_m of fluoropyruvate was lower.

Having characterized the substrate scope and kinetic parameters of the fluoropyruvate aldol addition, we then turned to DFT calculations in order to explore the origin of the strong (3*S*)-fluoro stereochemical preference (*Figure 2.13*). In the proposed mechanism of EcHpcH, the divalent metal ion is roughly octahedral, with bidentate chelation of the enolate intermediate and the remaining four sites occupied by a *cis* pair of water molecules and a *cis* pair of Asp/Glu residues [38-40]. This arrangement was used as a basis to model the metal-fluoroenolate complex in the enzyme active site as the simplified species $[\text{Mg}(\text{H}_2\text{O})_2(\text{OAc})_2(\text{Fpyr-enolate})]^{2-}$, which could bear fluorine in the (*Z*)- or (*E*)-fluoroenolate geometry. DFT calculations on the two possible isomeric model complexes revealed that the (*Z*)-isomer is favored by approximately 4.6 kcal/mol. The

instability of the (E)-isomer likely arises from an unfavorable 1,3-interaction between the fluorine and carboxylate oxygens, since the free (E)-fluoroenolate at equilibrium features the carboxylate twisted out of plane. In accordance with the HpcH enzyme architecture, delivery of the aldehyde to a consistent face of the (Z)-fluoroenolate results in a (3*S*)-fluoro product (*Figure 2.14*). Interestingly, a (Z)-fluoroenamine has been observed in the crystal structure of a Type I DHDPS family aldolase [16], but this enzyme leads to the (3*R*)-fluoro product since the aldehyde binding site is on the opposite face of the activated donor. Therefore, consistent fluorine-determined electronic features provide the high driving force for enantioselectivity. The absolute configuration of fluorine that arises depends on which face of the activated donor is poised to attack an incoming aldehyde, which is ultimately dictated by enzymatic architecture, such that the desired fluorine configuration might be achieved by selecting an aldolase from the appropriate structural family.

Given our model for the origin of (3*S*)-fluorine stereoselectivity, which is set by electronic factors and not easily altered, we next explored methods to control the more variable hydroxyl group stereochemistry. Noting that SwHpcH1 had a uniquely high *anti* selectivity when DMA was used as the acceptor, inspection of its active site revealed the residues Phe-117 and Phe-210 (SwHpcH1 numbering) in place of Val-118 and Leu-212 (EcHpcH numbering) found in all four other aldolases. A sequence alignment of 459 EcHpcH homologs was performed, which indicated that variations in non-catalytic active site residues such as these were fairly common (*Figure 2.15*). Hypothesizing that the bulkier Phe residue could modulate the hydroxyl group stereoselectivity by restricting the poses available to an aldehyde, the V118F, A174F, and L212F variants of EcHpcH were purified for screening against the same panel of twelve aldehydes as before. EcHpcH A174F and L212F produced stereocomplementary results in reactions with D-glyceraldehyde and D-lactaldehyde, while EcHpcH V118F had high *anti* selectivity in reactions with GA and DMA even at high conversions (*Figure 2.16*). Such selectivity with wild-type enzymes was only achieved previously by reducing the enzyme loading and sacrificing conversion. In the course of this study, others showed with RhmA that the strategy of removing steric bulk from the active site by mutation of non-conserved residues could broaden the substrate scope from pyruvate to larger α -ketoacids [45,46]. Our preliminary results with EcHpcH mutants bearing increased steric bulk show that the opposite strategy may also have value for improvement or even reversal of hydroxyl group stereoselectivity.

With a collection of wild-type and engineered aldolases in hand, we desired to demonstrate their applicability to preparative synthesis by scaling up several reactions from their ¹⁹F NMR analytical conditions, and isolating the corresponding fluorinated compounds with high enantioselectivity. Where possible, an enzyme and reaction condition was selected from our screening results that would enable high *syn* or *anti* diastereoselectivity as well (*Figure 2.17*). After the decarboxylation of fluoropyruvate adducts, excess H₂O₂ was decomposed with catalase. Due to the expensive nature of chiral aldehydes (D- or L-configured glyceraldehyde or lactaldehyde), the reactions with these substrates were conducted on milligram scale with direct purification by preparative HPLC on a HILIC column. This provided six stereo-enriched α -fluoro sugar acids (d.r. 87-98%), corresponding to three stereoisomers each of 2-fluoro-3,4,5-trihydropentanoate (FTHP) and 2-fluoro-3,4-dihydropentanoate (FDHP). For the inexpensive aldehydes GA, DMA, and GX, reactions were performed on gram scale, and the products esterified in methanolic HCl after removal of water and enzymes. The diol product of GA was additionally protected as the acetonide with 2,2-dimethoxypropane. The transformation of these compounds into relatively non-polar α -fluoroesters allowed routine purification of the stereo-enriched products (d.r. 87-98%)

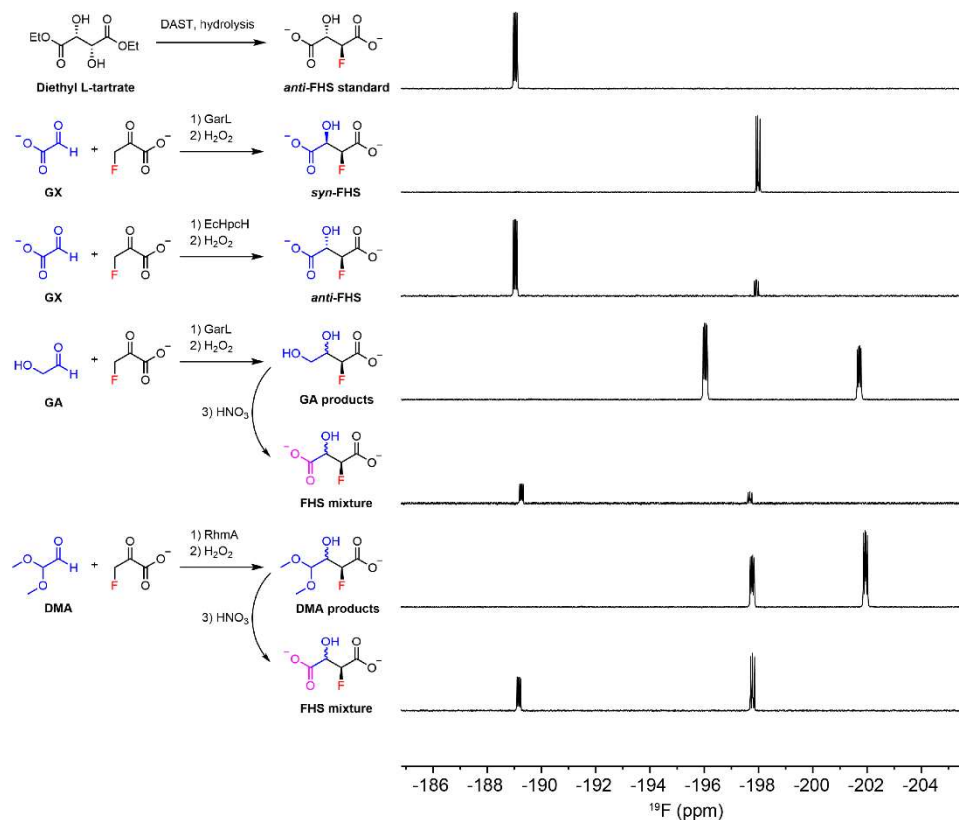


Figure 2.9. Nitric acid experiment to confirm relative stereochemistry of products. After terminal oxidation to a carboxylic acid, ^{19}F NMR peaks corresponded to *syn*- or *anti*-FHS.

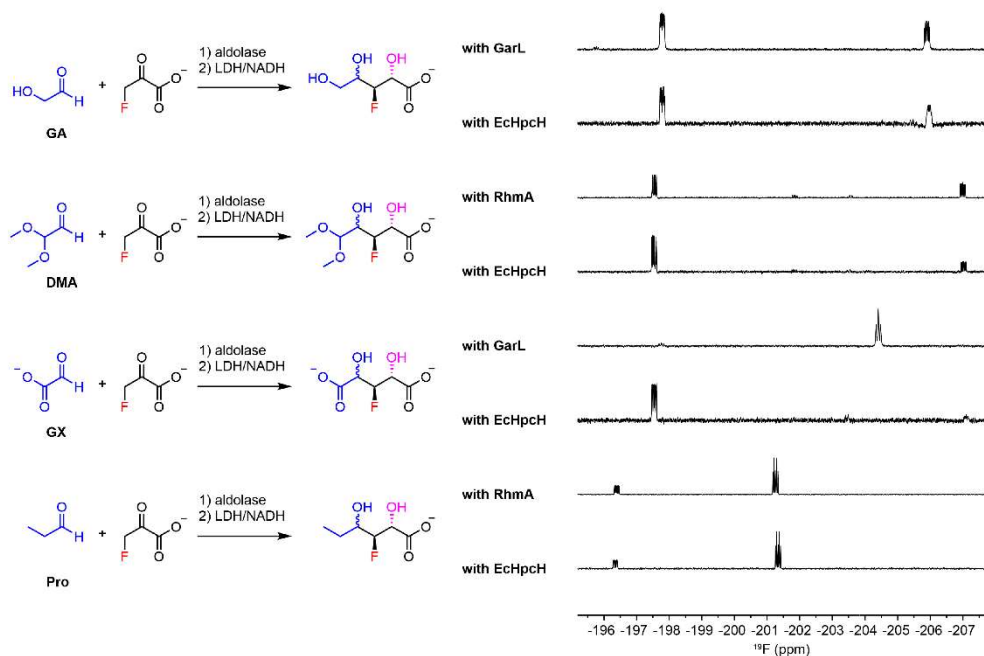


Figure 2.10. Lactate dehydrogenase experiment to confirm absolute stereochemistry of products. After stereoselective reduction by LDH, at most two peaks are observed in ^{19}F NMR.

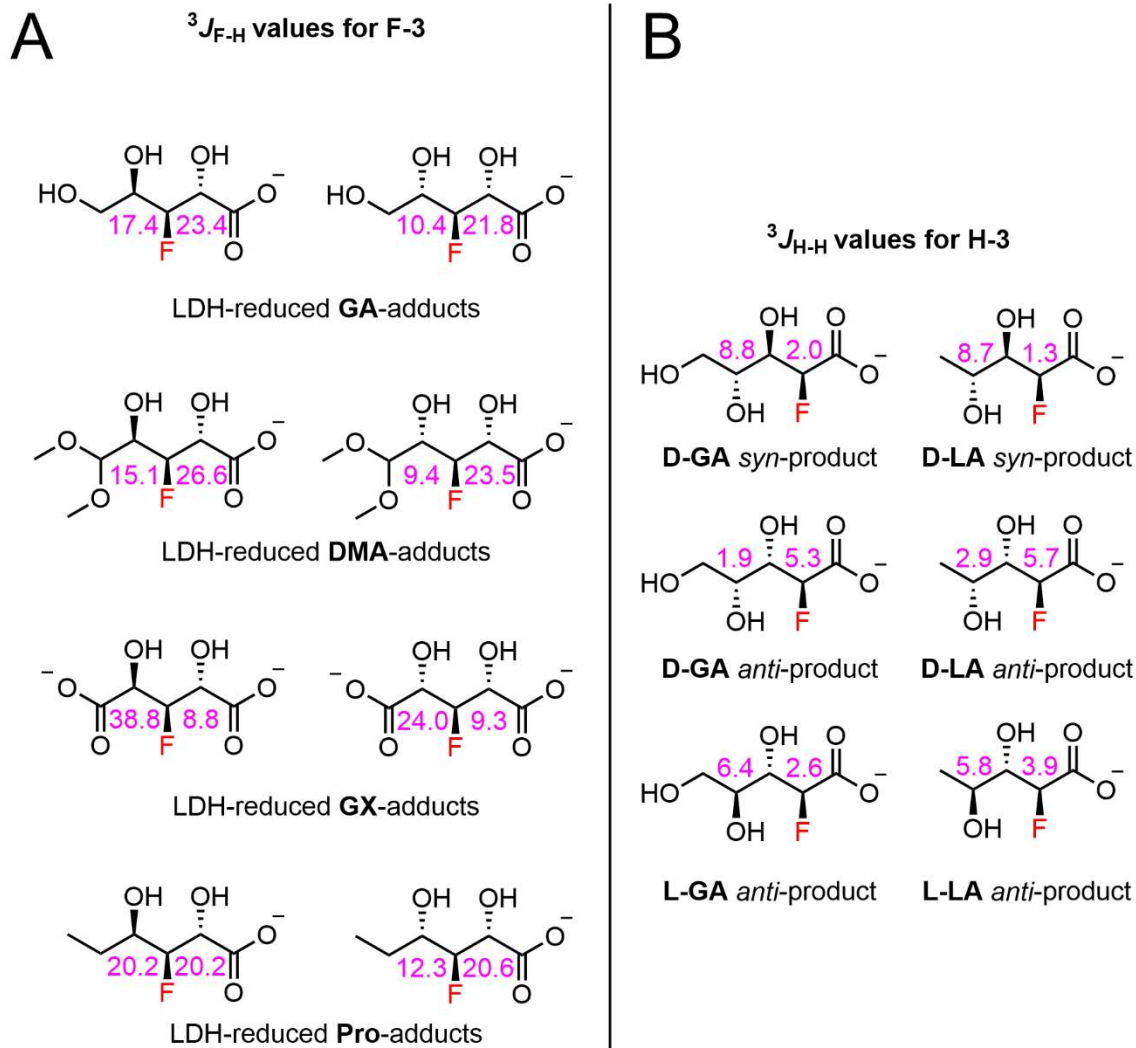
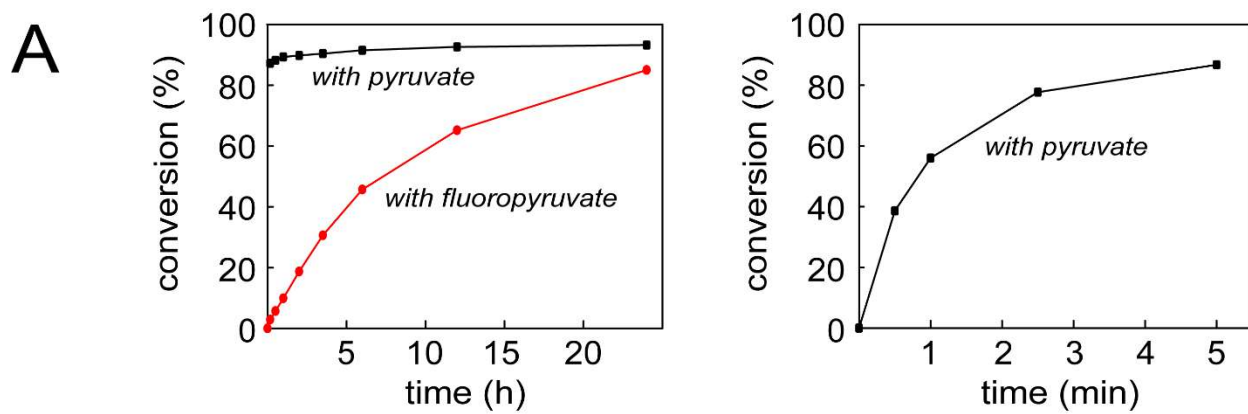


Figure 2.11. NMR coupling constants confirming the stereoselectivity of the HpcH-catalyzed fluoropyruvate aldol reaction as (*S*)-fluoro. A) Vicinal fluorine-proton coupling constants for various adducts reduced by LDH. The assigned structures are consistent with larger coupling values for *syn*-relationships of electronegative groups. B) Vicinal proton-proton coupling constants for H-3 of several 2-fluoroacids that were later isolated. The assigned structures are consistent with smaller coupling values for *syn*-relationships of electronegative groups.



B

Aldolase donor	k_{cat} (s^{-1})	K_m (mM)	K_{si} (mM)	k_{cat}/K_m ($\text{M}^{-1}\text{s}^{-1}$)
Pyruvate	220 ± 20	6.0 ± 0.7	n.d.	$(3.8 \pm 0.3) \cdot 10^4$
Fluoropyruvate	0.46 ± 0.02	1.1 ± 0.1	64 ± 5	$(4.4 \pm 0.5) \cdot 10^2$

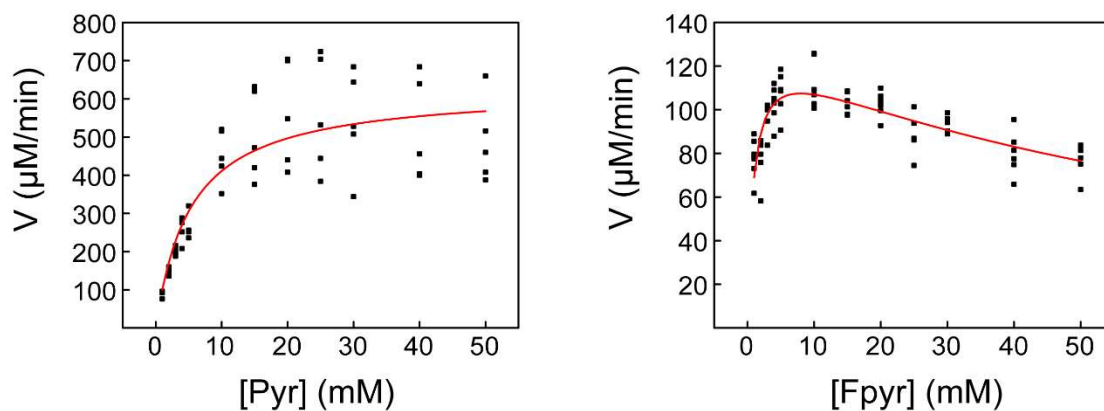


Figure 2.12. Kinetics of EchPcH. A) Progress curves for the addition of pyruvate or fluoropyruvate to GA, catalyzed by 0.05 mol% EchPcH. B) Michaelis-Menten parameters and plots for pyruvate and fluoropyruvate.

Parameter	Free energy difference (kcal/mol)
Energy(full (<i>E</i>)-fluoroenolate complex) – Energy(full (<i>Z</i>)-fluoroenolate complex)	4.62
Energy((<i>E</i>)-fluoroenolate at complex geometry) – Energy((<i>Z</i>)-fluoroenolate at complex geometry)	5.74
Energy(Mg(OAc) ₂ (H ₂ O) ₂ at (<i>E</i>)-complex geometry) – Energy(Mg(OAc) ₂ (H ₂ O) ₂ at (<i>Z</i>)-complex geometry)	-0.79
Energy(binding of fluoroenolate in the (<i>E</i>)-complex) – Energy(binding of fluoroenolate in the (<i>Z</i>)-complex)	0.33
Energy((<i>E</i>)-fluoroenolate at free equilibrium geometry) – Energy((<i>Z</i>)-fluoroenolate at free equilibrium geometry)	4.48
Energy((<i>E</i>)-fluoroenolate at complex geometry) – Energy((<i>E</i>)-fluoroenolate at free equilibrium geometry)	2.84
Energy((<i>Z</i>)-fluoroenolate at complex geometry) – Energy((<i>Z</i>)-fluoroenolate at free equilibrium geometry)	1.58

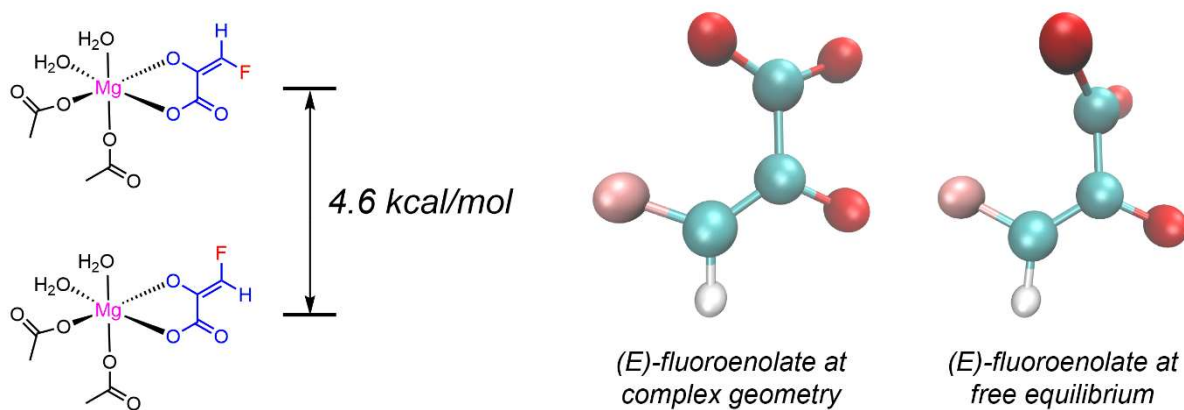


Figure 2.13. DFT calculations on a model complex of the fluoroenolate intermediate in the active site of HpCh.

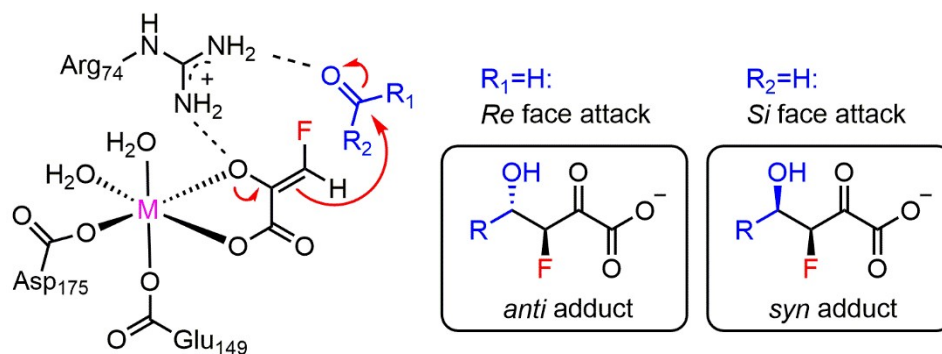


Figure 2.14. Mechanism of the C-C bond formation step for fluoropyruvate aldol addition. The fluorine stereocenter is fixed by the (*Z*)-fluoroenolate geometry, while the hydroxyl stereocenter is variable depending on which face of the aldehyde is attacked.

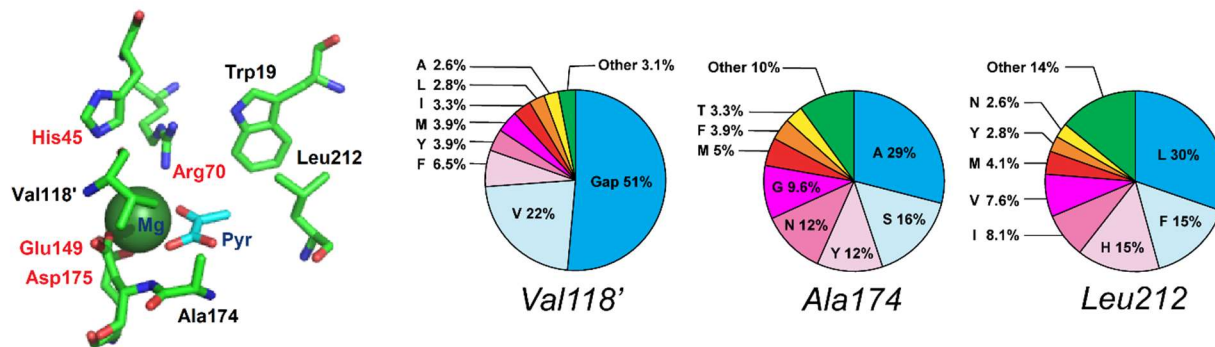


Figure 2.15. Analysis of natural variation for non-catalytic active site residues. Crystal structure is from RCSB PDB 2VWT [Ref. 34]. Prime mark for Val-118 indicates this residue is from a neighboring subunit chain.

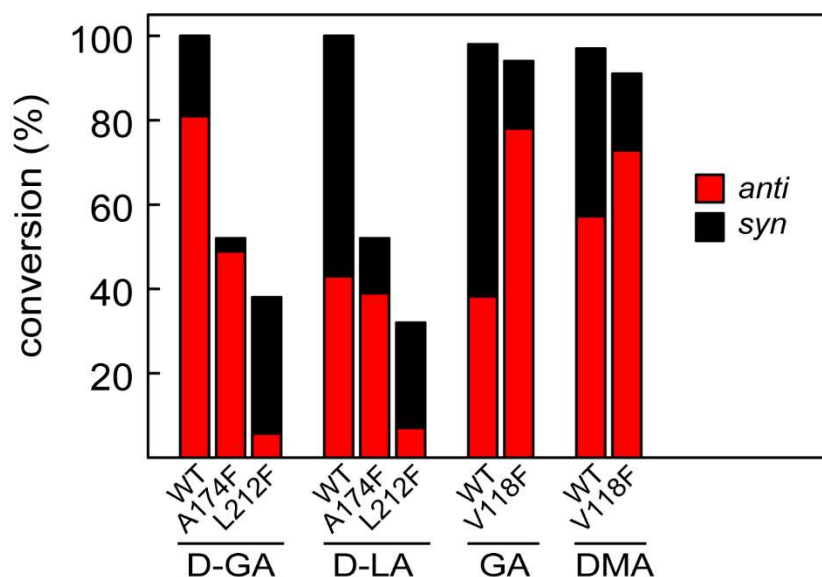


Figure 2.16. Altered diastereoselectivity observed on substrates *D*-glyceraldehyde, *D*-lactaldehyde, glycolaldehyde, and dimethoxyacetaldehyde, when small residues of *EcHpcH* were mutated to phenylalanine.

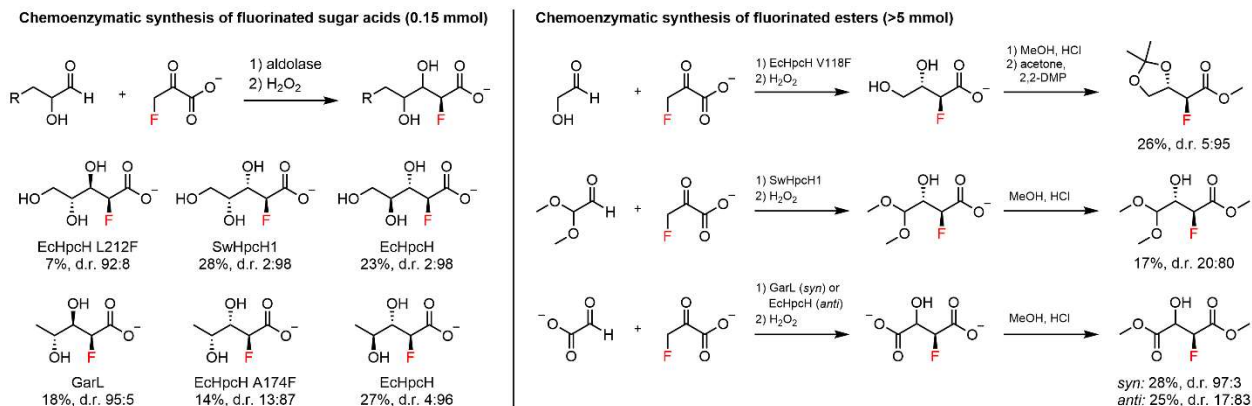


Figure 2.17. Chemoenzymatic synthesis and purification of fluorinated compounds in high diastereomeric purity.

by organic extraction. All purified compounds were characterized fully by high resolution LC-MS (Figure 2.18) and by ^1H , ^{13}C , and ^{19}F NMR (Figure 2.19 through Figure 2.28).

To show that the structures accessed by fluoropyruvate addition can serve as versatile precursors for downstream processing, we used chemical and enzymatic methods to transform products of the aldol reaction into a range of compounds such as β -fluoro- α -hydroxyacids, β -fluoro- α -amino acids, and fluorosugars (Figure 2.29). These compounds were prepared on small scale and detected by high resolution LC-MS. From the initial β -fluoro- α -ketoacid adducts (products of fluoropyruvate addition without decarboxylation), we demonstrated the feasibility of dehydrogenation and transamination reactions. Both reactions were inspired by the native metabolism of pyruvate, which is acted upon by dehydrogenases to form lactate and by transaminases to form alanine. Indeed, commercial L-lactate dehydrogenase, which was employed earlier for verifying absolute stereochemistry, reduced ten different fluoropyruvate adducts to the corresponding β -fluoro- α -hydroxyacids (Figure 2.30A). Only the adducts of D- or L-glyceraldehyde were not reduced. To attempt transamination of fluoropyruvate adducts, we cloned and purified the transaminase from *Vibrio fluvialis* (VfAT) which was known to convert fluoropyruvate to β -fluoroalanine [47]. VfAT with (*S*)- α -methylbenzylamine (SMBA) as amine donor was competent to convert the fluoropyruvate adducts of Pro and 2PCA into β -fluoro- α -amino acids (Figure 2.30B).

Finally, to produce 2'- or 3'-fluorosugars, we subjected several of the purified acids or esters to synthetic routes involving chemical reduction (Figure 2.31). The α -fluoroacids derived from the *syn*-adduct of D-glyceraldehyde, the *anti*-adduct of D-glyceraldehyde, and the *syn*-adduct of D-lactaldehyde were treated with a sequence of methyl esterification, acetonide protection, DIBAL-H reduction, IBX oxidation, and acidic deprotection. This process respectively provided 2-fluoro-2-deoxy-D-arabinose, 2-fluoro-2-deoxy-D-lyxose, and 2-fluoro-2,5-dideoxy-D-arabinose. Notably, 2-fluoro-2-deoxy-D-arabinose is a component of the anticancer drug clofarabine, which has been targeted for chemoenzymatic synthesis [48]. The α -fluoroester derived from GA underwent acetonide protection, DIBAL-H reduction, Swern oxidation, and acidic deprotection to give 2-fluoro-2-deoxy-L-erythrose. An inversion strategy was proposed to convert the α -fluoroester derived from DMA into a fluorosugar, taking advantage of the masked C-4 aldehyde. Thus, the treatment of this compound with DIBAL-H and acidic deprotection led to 3-fluoro-3-deoxy-D-erythrose. All fluorosugars were characterized by high resolution LC-MS following derivatization with *O*-benzylhydroxylamine (OBHA) (Figure 2.32). These efforts show that aldolases can serve as the basis of a platform for stereoselective synthesis of pharmaceutically-relevant organofluorines that may be difficult to access by other means.

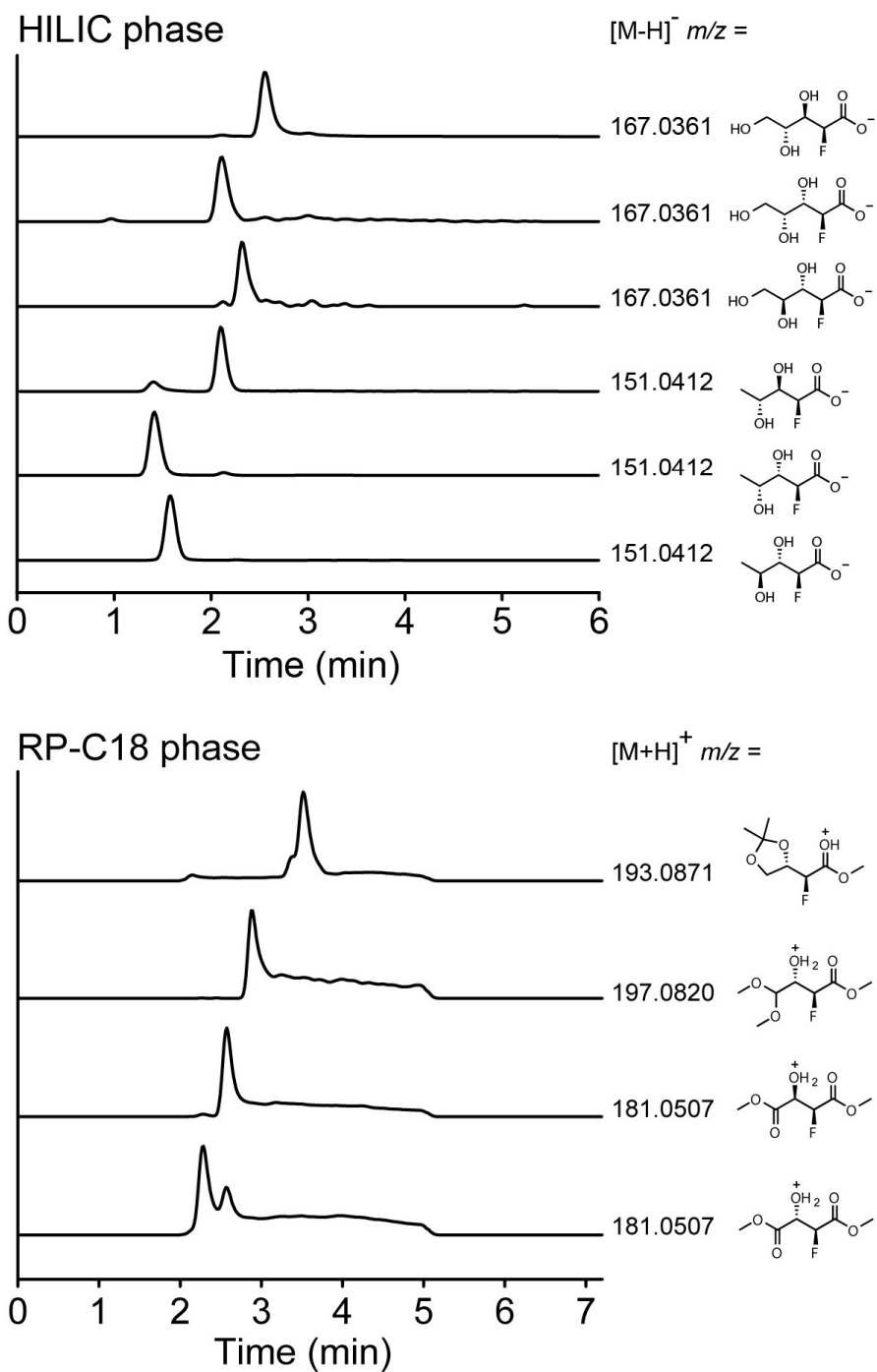


Figure 2.18. High resolution LC-MS of isolated fluorinated compounds. The intensity of each chromatogram is scaled to the largest in each set.

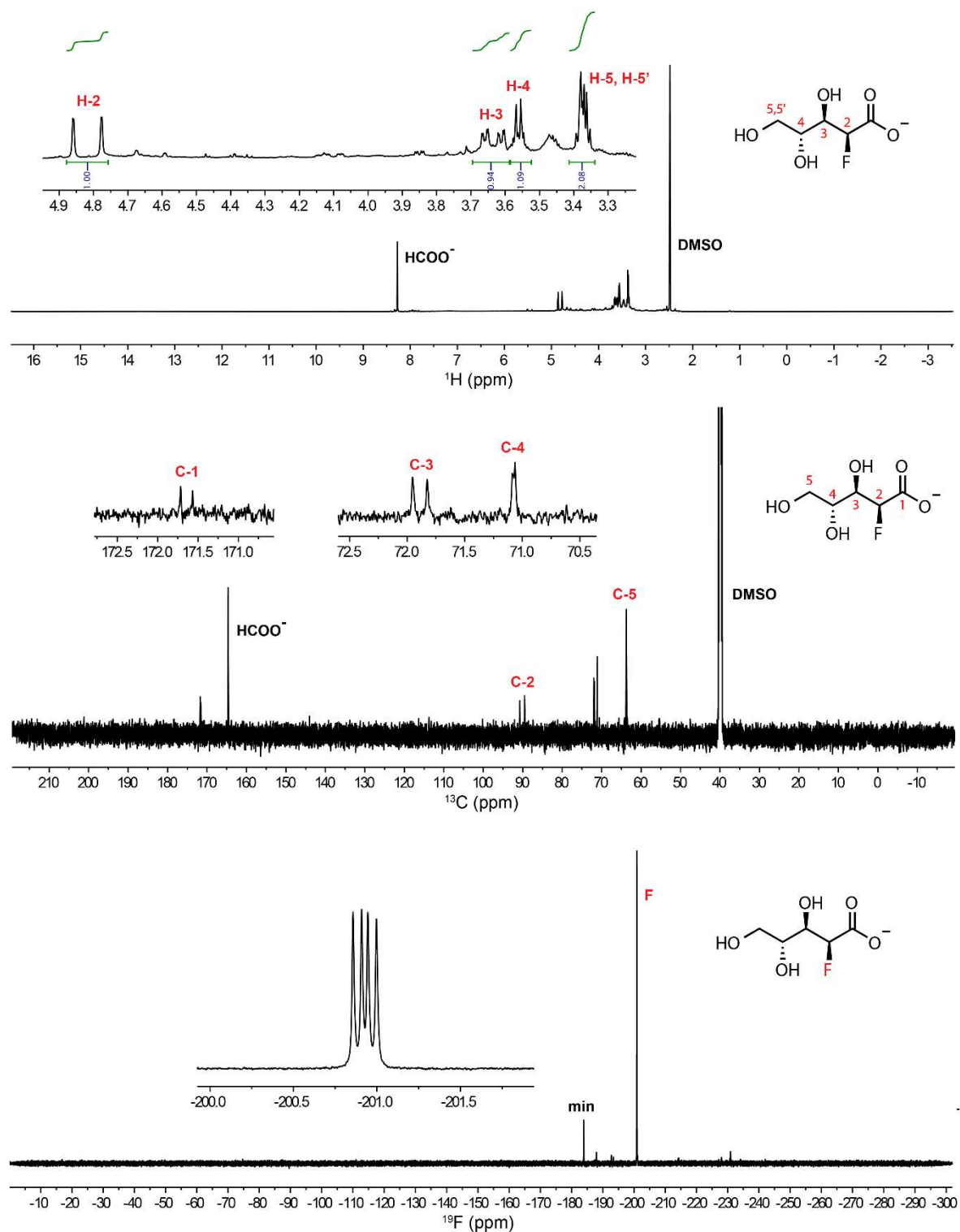


Figure 2.19. NMR spectra of $(2S,3R,4R)$ -2-fluoro-3,4,5-trihydroxypentanoate; min = minor diastereomer.

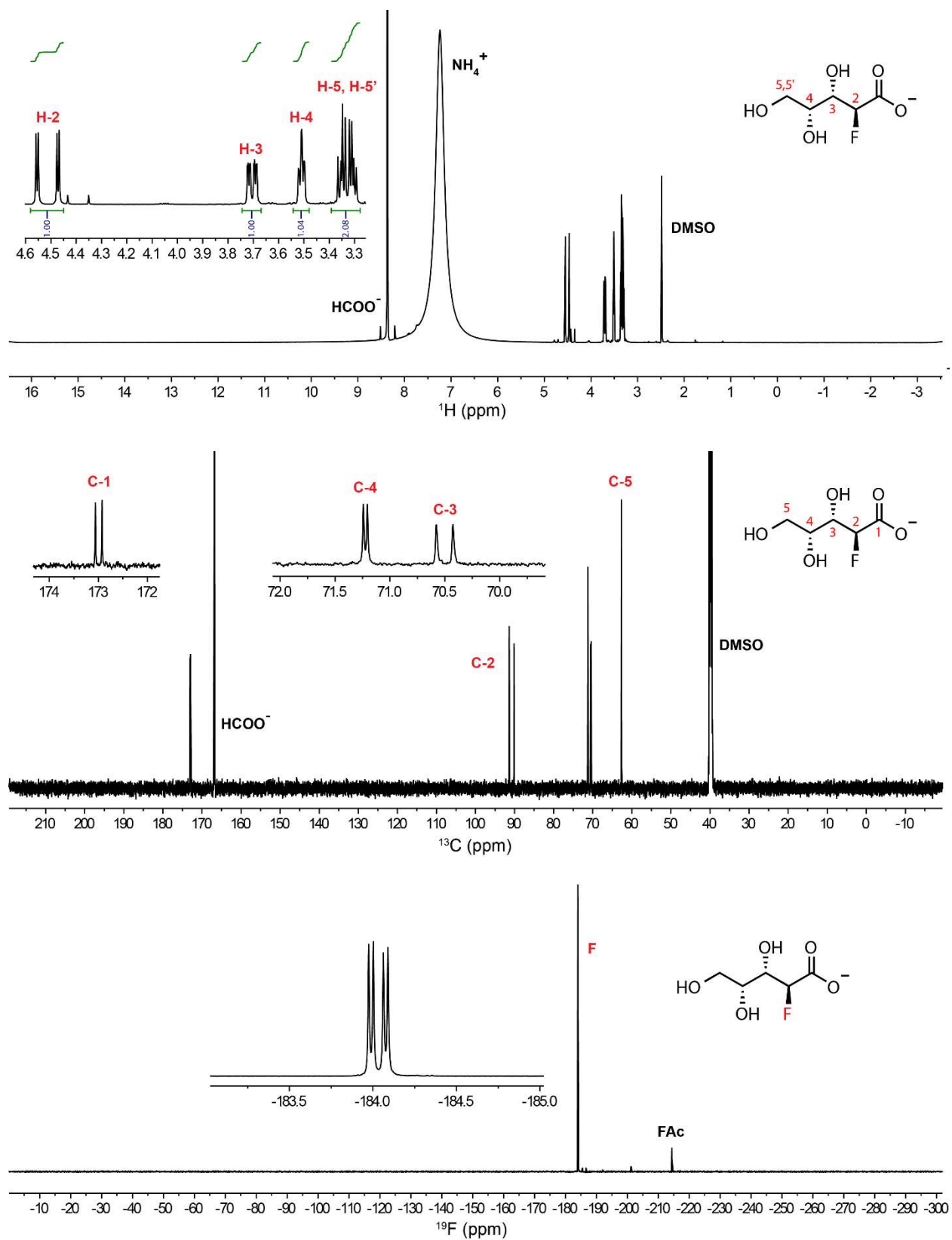


Figure 2.20. NMR spectra of (2S,3S,4R)-2-fluoro-3,4,5-trihydroxypentanoate; FAc = fluoroacetate.

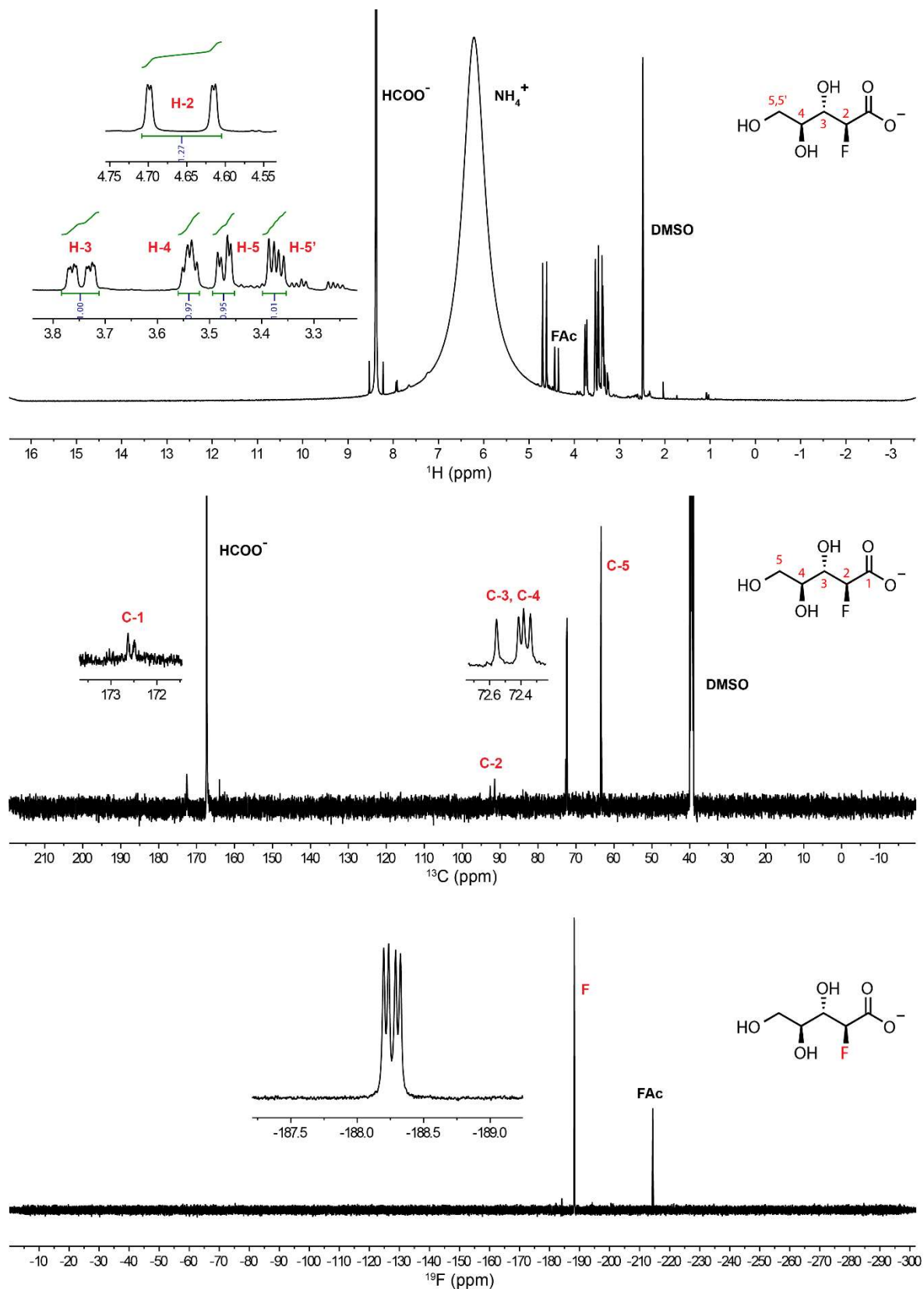


Figure 2.21. NMR spectra of (2S,3S,4S)-2-fluoro-3,4,5-trihydroxypentanoate; FAc = fluoroacetate.

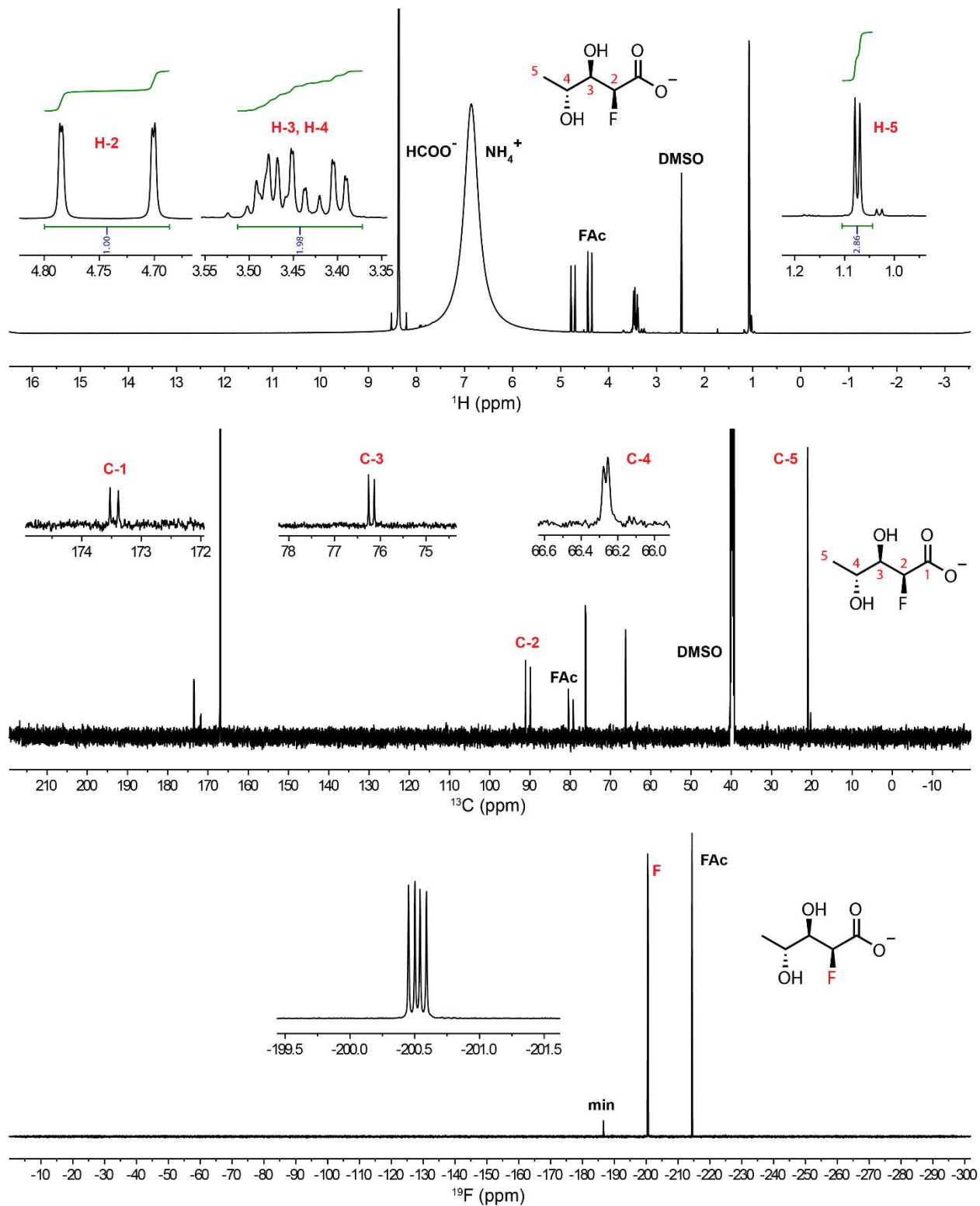


Figure 2.22. NMR spectra of (2S,3R,4R)-2-fluoro-3,4-dihydroxypentanoate; min = minor diastereomer, FAc = fluoroacetate.

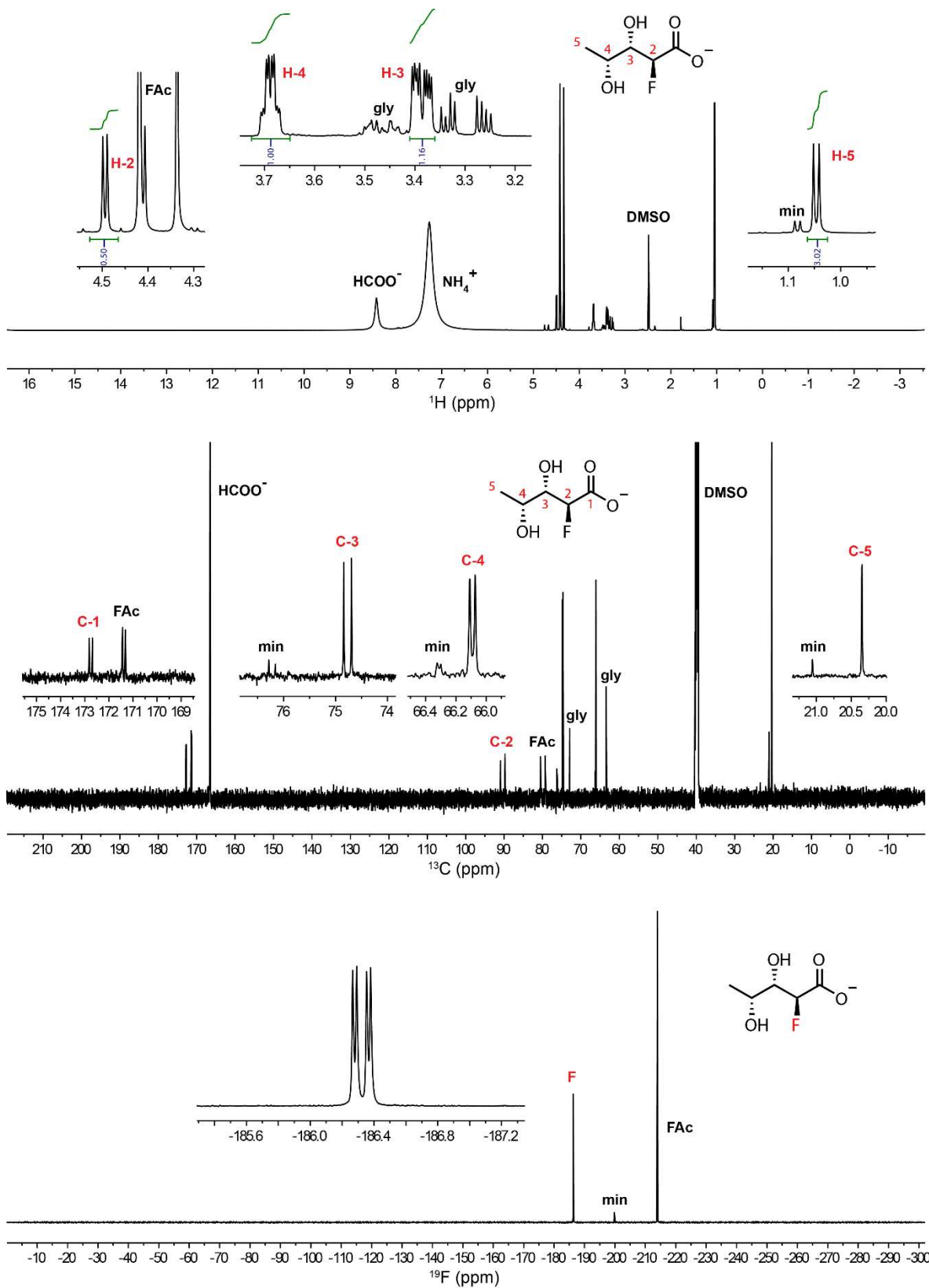


Figure 2.23. NMR spectra of $(2S,3S,4R)$ -2-fluoro-3,4-dihydroxypentanoate; min = minor diastereomer, FAc = fluoroacetate, gly = glycerol.

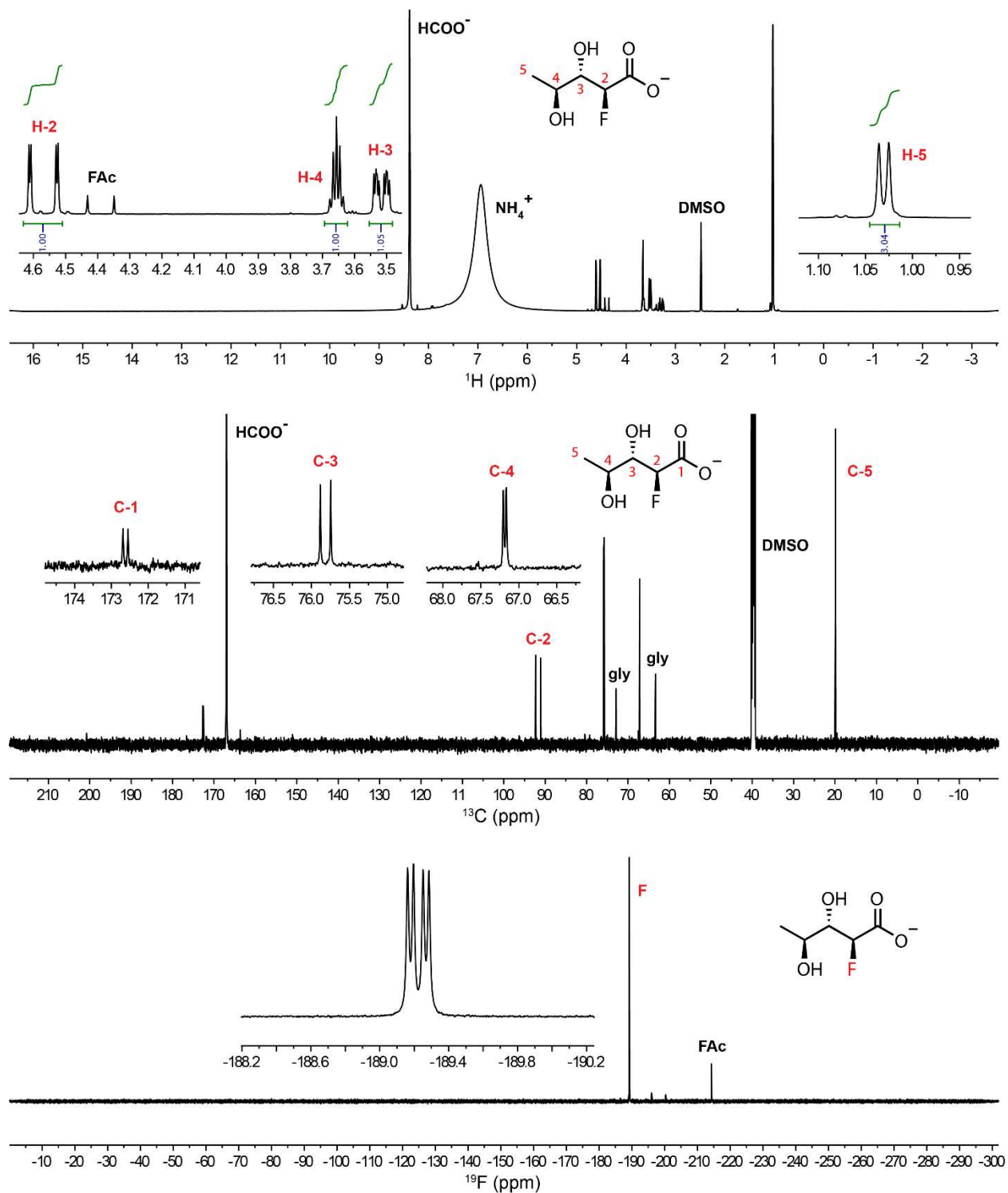


Figure 2.24. NMR spectra of (2S,3S,4S)-2-fluoro-3,4-dihydroxypentanoate; FAc = fluoroacetate, gly = glycerol.

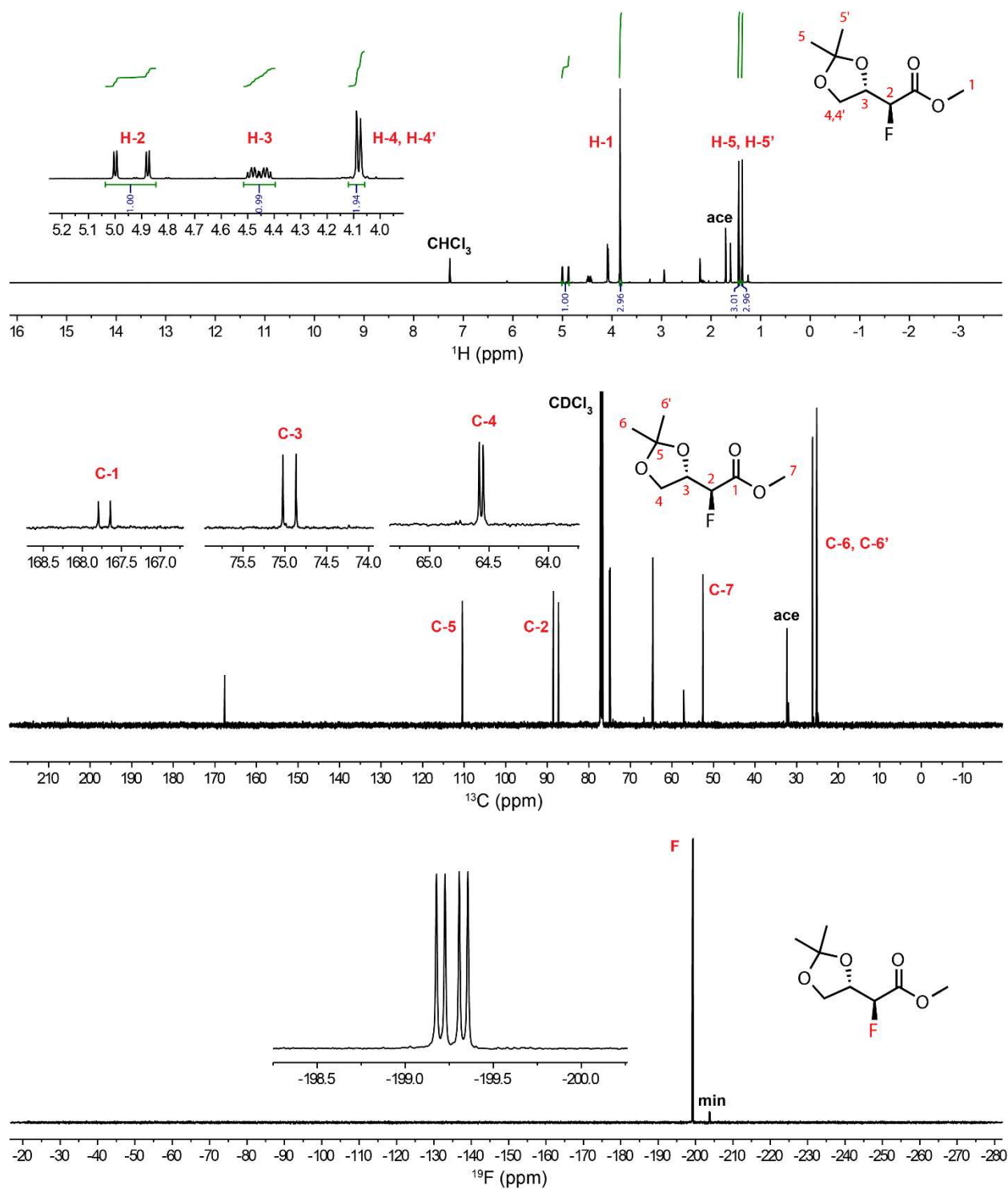


Figure 2.25. NMR spectra of methyl (2S,3S)-2-fluoro-3,4-O-isopropylidene-3,4-dihydroxybutanoate; min = minor diastereomer, ace = acetone.

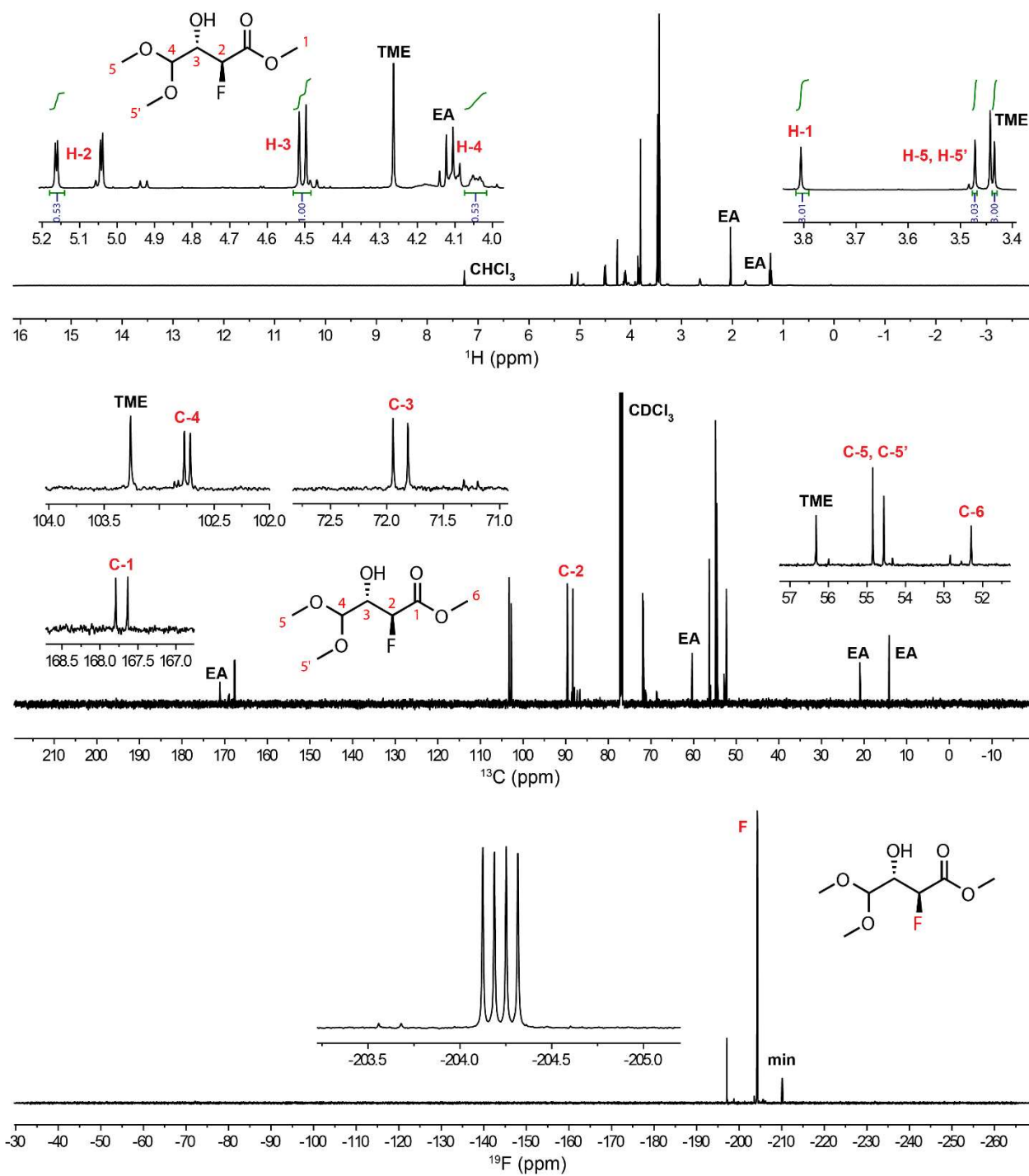


Figure 2.26. NMR spectra of methyl (2S,3S)-2-fluoro-3-hydroxy-4,4-dimethoxybutanoate; min = minor diastereomer, EA = ethyl acetate, TME = 1,1,2,2-tetramethoxyethane.

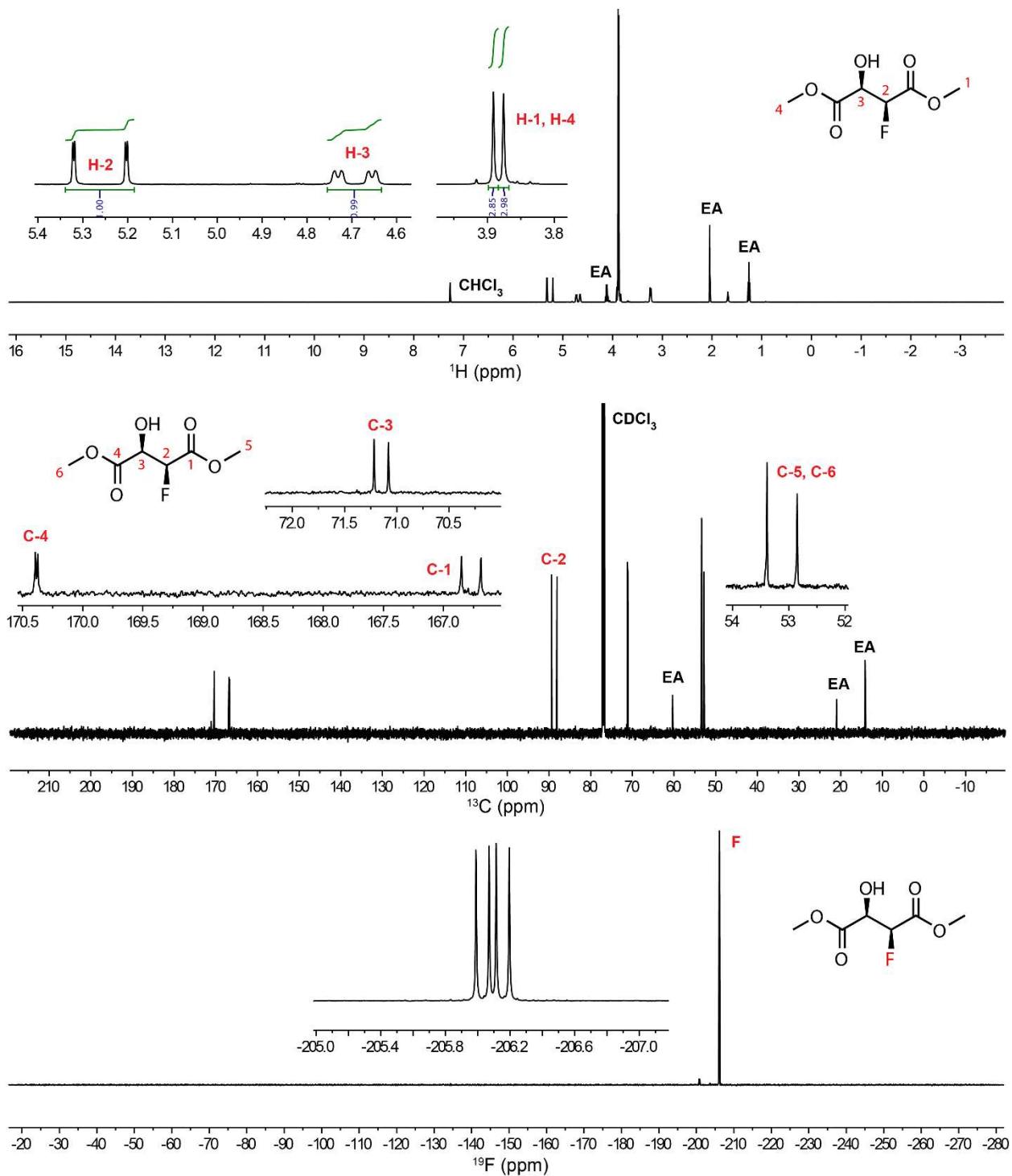


Figure 2.27. NMR spectra of dimethyl (2S,3R)-2-fluoro-3-hydroxysuccinate; EA = ethyl acetate.

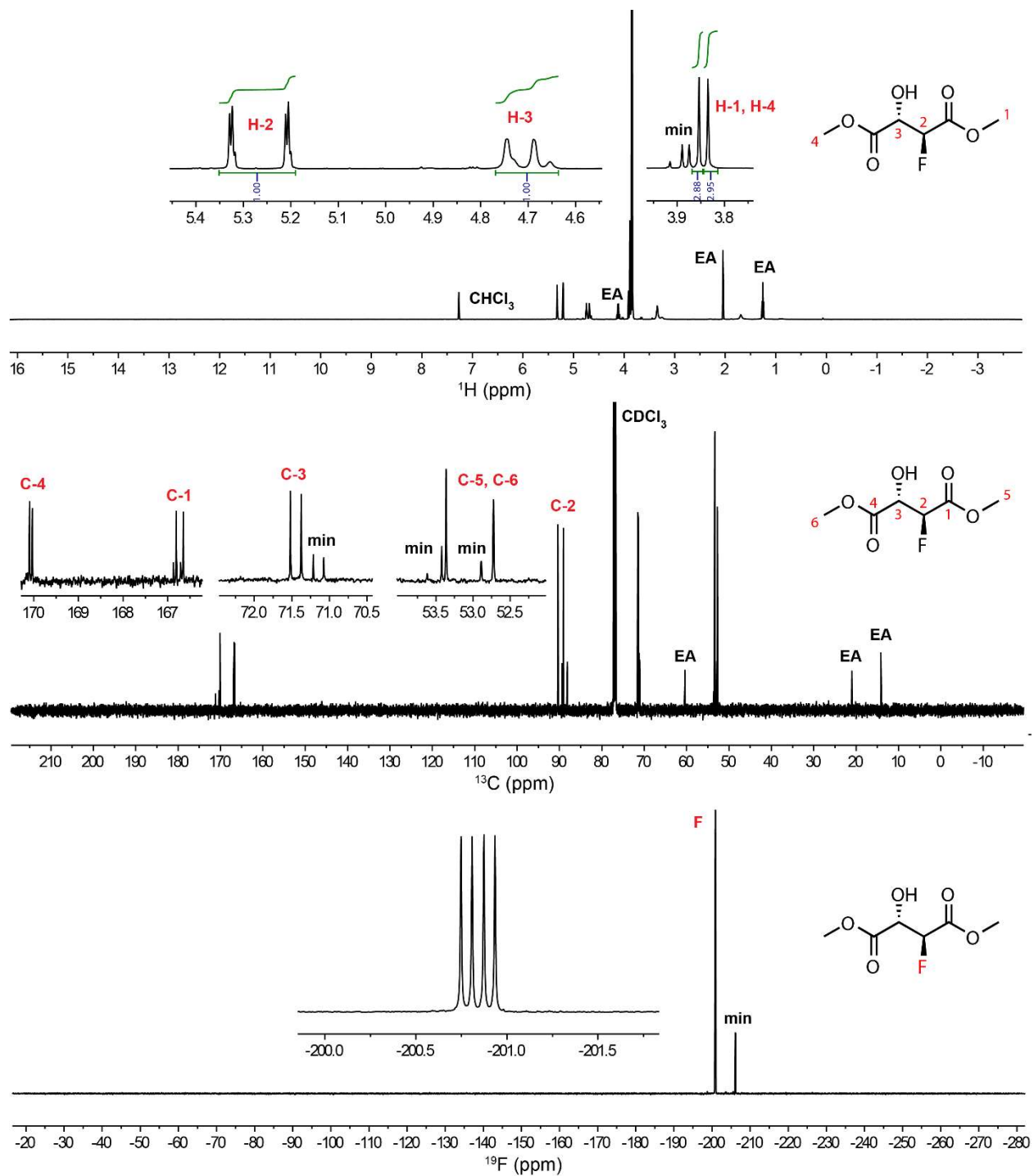


Figure 2.28. NMR spectra of dimethyl (2*S*,3*S*)-2-fluoro-3-hydroxysuccinate; min = minor diastereomer, EA = ethyl acetate.

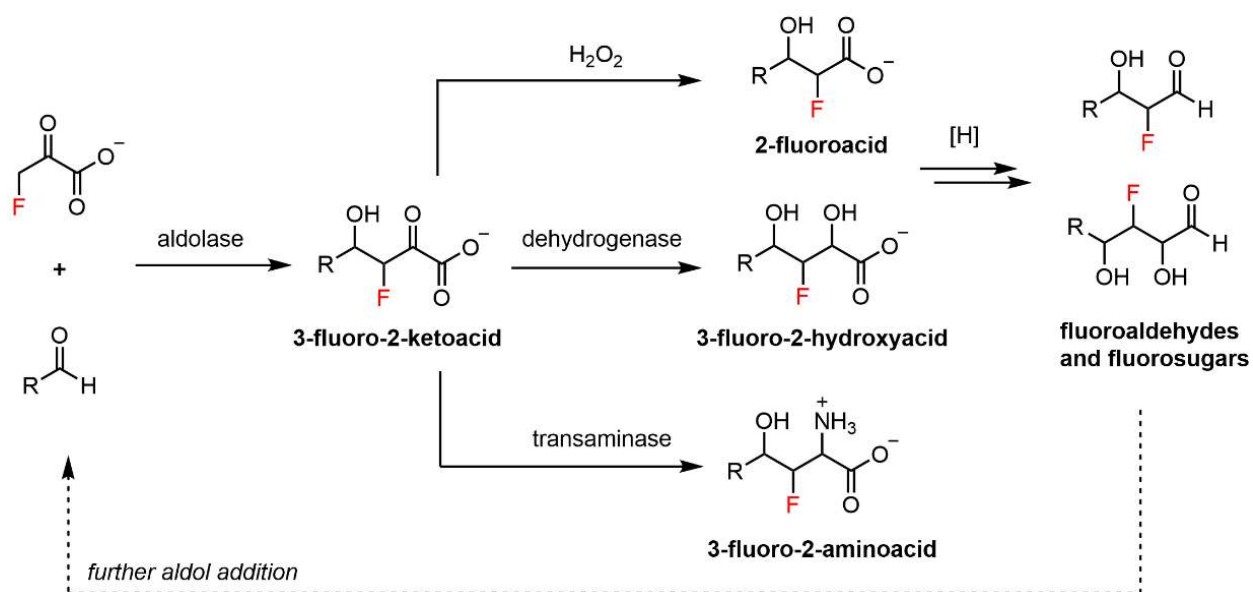


Figure 2.29. Roadmap of downstream reactions for transforming fluoropyruvate aldol adducts into other organofluorine compounds such as fluorinated hydroxyacids, amino acids, and sugars.

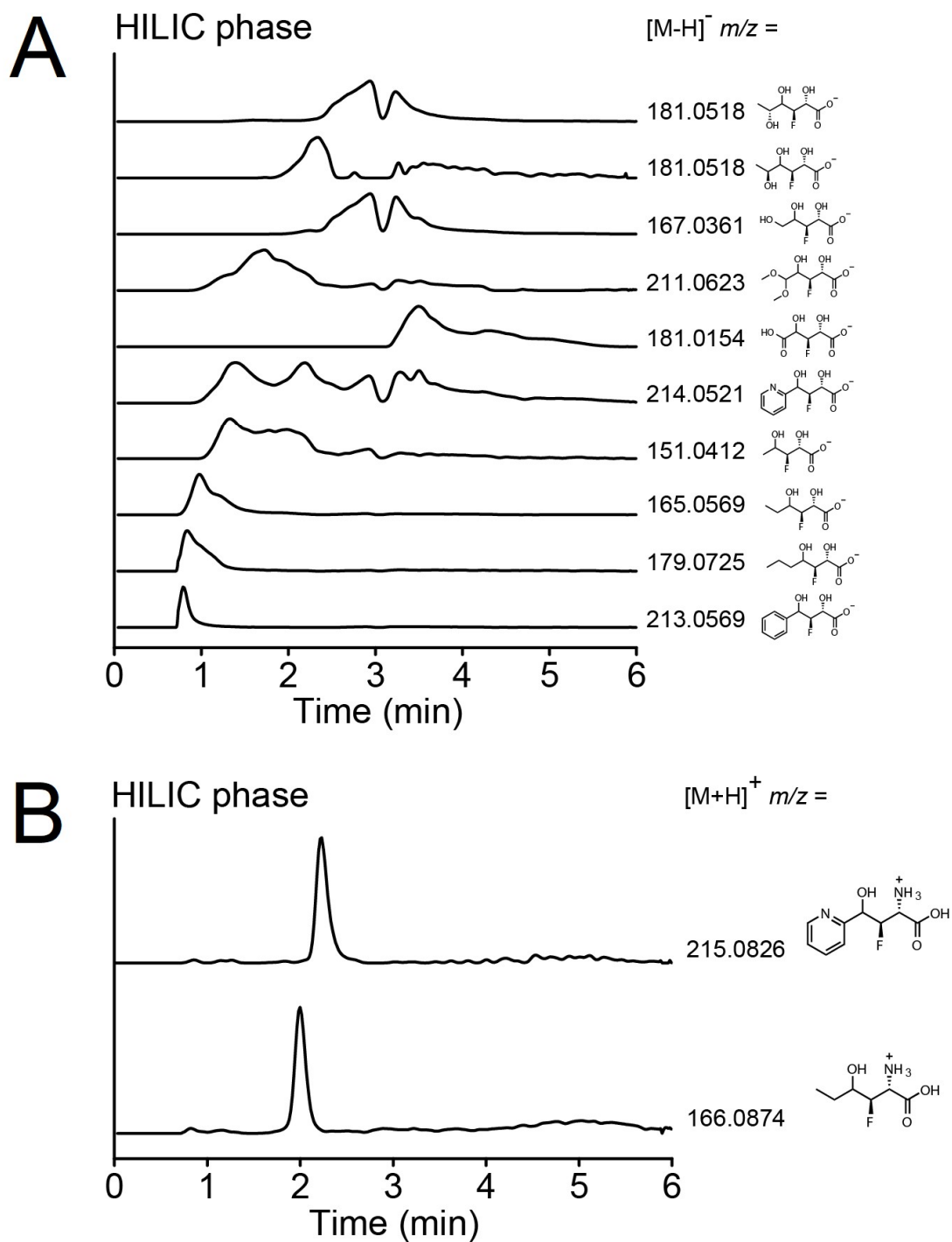


Figure 2.30. High resolution LC-MS of products from downstream reactions. The corresponding masses were not detected in blank injection controls. The intensity of each chromatogram is scaled to the largest in each set. A) 3-Fluoro-2-hydroxyacids prepared by LDH reduction. Poor peak shape for most products may be attributable to complex hydrogen-bonding interactions with the stationary phase. B) 3-Fluoro-2-aminoacids prepared with the transaminase VtAT from *Vibrio fluvialis* and (S)- α -methylbenzylamine as amine donor.

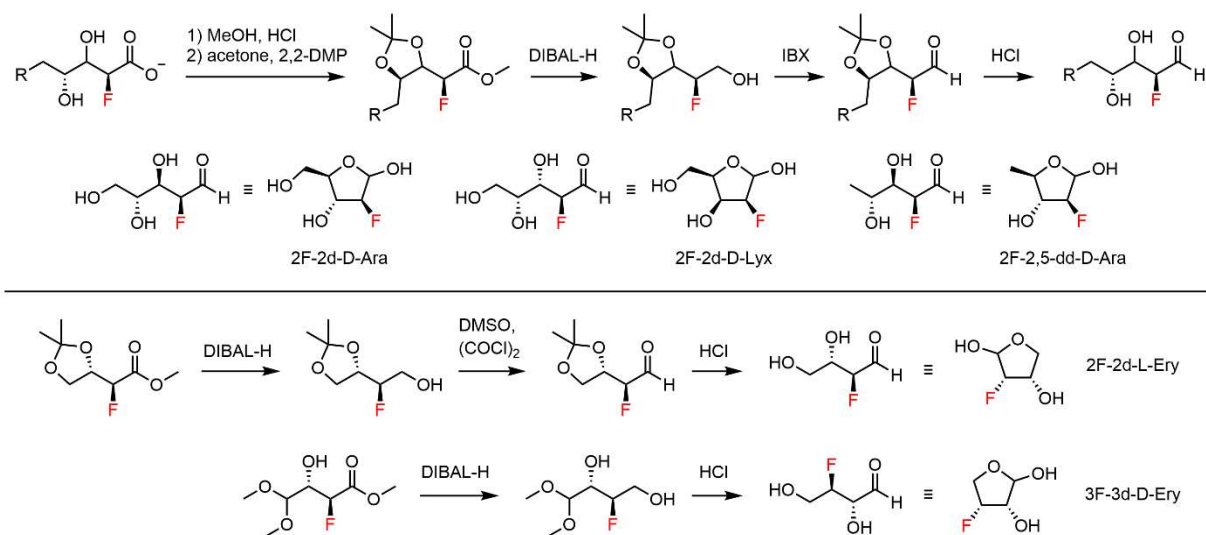


Figure 2.31. Chemical reactions performed to convert enzymatic products into fluorinated sugars. The crude product of each step was used directly for the subsequent step; no intermediates or products were purified. Note that for the trihydroxy compounds ($R = OH$), the actual regioselectivity of acetonide protection was unknown.

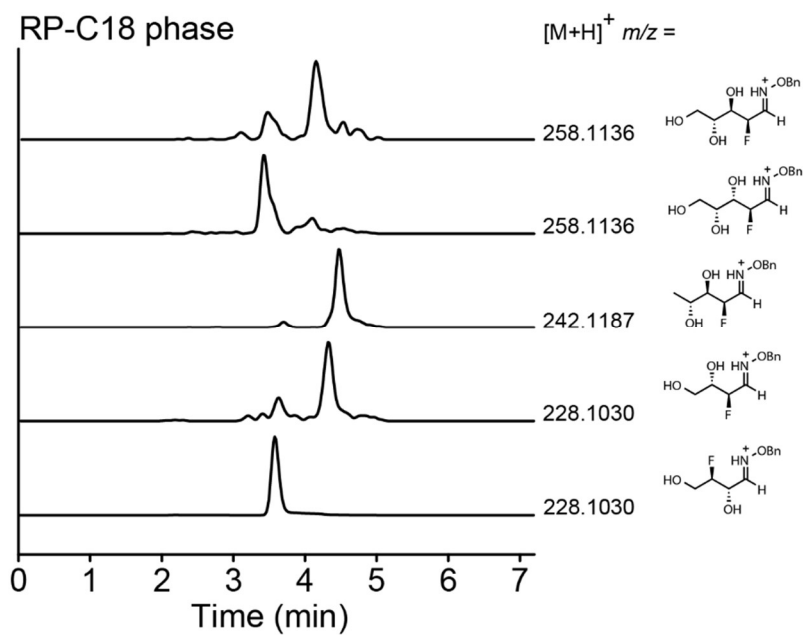


Figure 2.32. High resolution LC-MS of fluorosugars after oxime derivatization with *O*-benzylhydroxylamine (OBHA).

2.4. Conclusion

We have demonstrated that the type II HpcH family of pyruvate-dependent aldolases is capable of efficient fluoropyruvate aldol additions to a variety of aldehyde partners. In accordance with the highly polar native substrates of HpcH aldolases, polar and hydroxylated aldehydes were determined to be superior reaction partners for fluoropyruvate. Our kinetic characterization of EcHpcH comparing pyruvate and fluoropyruvate revealed that the slower reaction rate induced by a fluorine substitution was primarily due to a 500-fold decrease in turnover number. Nevertheless, due to the remarkable innate turnover capability of HpcH aldolases, even the use of non-native substrates in both positions—such as fluoropyruvate with less polar aldehydes—could still be run under practical enzyme loading. Our quantum mechanical modelling of the enzyme active site to interrogate the origin of fluorine stereoselectivity revealed the thermodynamic factors that yield a large driving force for formation of the (*Z*)-fluoroenolate and its corresponding (*3S*)-fluoro product. Additionally, our preliminary engineering work shows potential for augmenting or reversing hydroxyl group stereoselectivity in aldol products. The utility of the HpcH platform was shown concretely with preparative syntheses of organofluorine building blocks and downstream transformations towards fluorinated analogs of amino acids, sugars, and other valuable compounds. Overall, this work paves the way towards a new method to prepare a range of organofluorine molecules, with implications for broadening the role of biocatalysis in pharmaceutical synthesis.

2.5. References

1. Dean, S. M., Greenberg, W. A., Wong, C.-H. Recent advances in aldolase-catalyzed asymmetric synthesis. *Adv Synth Catal* **2007**, *349* (8-9) 1308-1320.
2. Brovetto, M., Gamenara, D., Méndez, P. S., Seoane, G. A. C-C bond-forming lyases in organic synthesis. *Chem Rev* **2011**, *111* (7) 4346-4403.
3. Fesko, K., Gruber-Khadjawi, M. Biocatalytic methods for C-C bond formation. *ChemCatChem* **2013**, *5* (6) 1248-1272.
4. Trost, B. M., Brindle, C. S. The direct catalytic asymmetric aldol reaction. *Chem Soc Rev* **2010**, *39* (5) 1600-1632.
5. Yamashita, Y., Yasukawa, T., Yoo, W.-J., Kitanosono, T., Kobayashi, S. Catalytic enantioselective aldol reactions. *Chem Soc Rev* **2018**, *47* (12) 4388-4480.
6. Schümperli, M., Pellaux, R., Panke, S. Chemical and enzymatic routes to dihydroxyacetone phosphate. *Appl Microbiol Biotechnol* **2007**, *75* (1) 33-45.
7. Li, A., Cai, L., Chen, Z., Wang, M., Wang, N., Nakanishi, H., Gao, X.-D., Li, Z. Recent advances in the synthesis of rare sugars using DHAP-dependent aldolases. *Carbohydr Res* **2017**, *452* 108-115.
8. Welch, J. T., Plummer, J. S. The stereoselective aldol condensation of α -fluoroacetates. *Synth Commun* **1989**, *19* (5-6) 1081-1090.
9. Oda, S., Yamamoto, H. Synthesis of β -hydroxy- α -haloesters through super silyl ester directed syn-selective aldol reaction. *Org Lett* **2013**, *15* (23) 6030-6033.

10. Saadi, J., Wennemers, H. Enantioselective aldol reactions with masked fluoroacetates. *Nat Chem* **2016**, *8* (3) 276-280.
11. Pankiewicz, K. W. Fluorinated nucleosides. *Carbohydr Res* **2000**, *327* (1-2) 87-105.
12. Linclau, B., Ardá, A., Reichardt, N.-C., Sollogoub, M., Unione, L., Vincent, S. P., Jiménez-Barbero, J. Fluorinated carbohydrates as chemical probes for molecular recognition studies. Current status and perspectives. *Chem Soc Rev* **2020**, *49* (12) 3863-3888.
13. Flournoy, D. S., Frey, P. A. Inactivation of the pyruvate dehydrogenase complex of *Escherichia coli* by fluoropyruvate. *Biochemistry* **1989**, *28* (25) 9594-9602.
14. Shelton, M. C., Cotterill, I. C., Novak, S. T. A., Poonawala, R. M., Sudarshan, S., Toone, E. J. 2-Keto-3-deoxy-6-phosphogluconate aldolases as catalysts for stereocontrolled carbon-carbon bond formation. *J Am Chem Soc* **1996**, *118* (9) 2117-2125.
15. Cotterill, I. C., Henderson, D. P., Shelton, M. C., Toone, E. J. The synthetic utility of KDPGal aldolase. *J Mol Catal B Enzym* **1998**, *5* (1-4) 103-111.
16. Stockwell, J., Daniels, A. D., Windle, C. L., Harman, T. A., Woodhall, T., Lebl, T., Trinh, C. H., Mulholland, K., Pearson, A. R., Berry, A., Nelson, A. Evaluation of fluoropyruvate as nucleophile in reactions catalyzed by N-acetyl neuraminic acid lyase variants: scope, limitations and stereoselectivity. *Org Biomol Chem* **2016**, *14* (1) 105-112.
17. Howard, J. K., Müller, M., Berry, A., Nelson, A. An enantio- and diastereoselective chemoenzymatic synthesis of α -fluoro β -hydroxy carboxylic esters. *Angew Chem Int Ed* **2016**, *55* (23) 6767-6770.
18. Windle, C. L., Berry, A., Nelson, A. Aldolase-catalyzed stereoselective synthesis of fluorinated small molecules. *Curr Opin Chem Biol* **2017**, *37* 33-38.
19. Dick, M., Hartmann, R., Weiergräber, O. H., Bisterfeld, C., Classen, T., Schwarten, M., Neudecker, P., Willbold, D., Pietruszka, J. Mechanism-based inhibition of an aldolase at high concentrations of its natural substrate acetaldehyde: structural insights and protective strategies. *Chem Sci* **2016**, *7* (7) 4492-4502.
20. Manjasetty, B. A., Powlowski, J., Vrielink, A. Crystal structure of a bifunctional aldolase-dehydrogenase: sequestering a reactive and volatile intermediate. *Proc Natl Acad Sci U S A* **2003**, *100* (12) 6992-6997.
21. Fang, J., Hait, D., Head-Gordon, M., Chang, M. C. Y. Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angew Chem Int Ed* **2019**, *58* (34) 11841-11845.
22. Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., Vidal, M. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol* **2000**, *328* 575-592.
23. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A., Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009**, *6* (5) 343-345.

24. Wang, W., Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions, and insertions using QuikChange site-directed mutagenesis. *Biotechniques* **1999**, *26* (4) 680-682.
25. 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.
26. Shao, Y., et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol Phys* **2014**, *113* 184-215.
27. Mardirossian, N., Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J Chem Phys* **2016**, *144* 214110.
28. Marenich, A. V., Cramer, C. J., Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B* **2009**, *113* (18) 6378-6396.
29. Weigend, F., Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* **2005**, *7* (18) 3297-3305.
30. Li, Y.-P., Gomes, J., Sharada, S. M., Bell, A. T., Head-Gordon, M. Improved force-field parameters for QM/MM simulations of the energies of adsorption for molecules in zeolites and a free rotor correction to the rigid rotor harmonic oscillator model for adsorption enthalpies. *J Phys Chem C* **2015**, *119* (4) 1840-1850.
31. Hubbard, B. K., Koch, M., Palmer, D. R., Babbitt, P. C., Gerlt, J. A. Evolution of enzymatic activities in the enolase superfamily: characterization of the (D)-glucarate/galactarate catabolic pathway in *Escherichia coli*. *Biochemistry* **1998**, *37* (41) 14369-14375.
32. Izard, T., Blackwell, N. C. Crystal structures of the metal-dependent 2-dehydro-3-deoxygalactarate aldolase suggest a novel reaction mechanism. *EMBO J* **2000**, *19* (15) 3849-3856.
33. Watanabe, S., Saimura, M., Makino, K. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. *J Biol Chem* **2008**, *283* (29) 20372-20382.
34. Rea, D., Hovington, R., Rakus, J. F., Gerlt, J. A., Fülöp, V., Bugg, T. D. H., Roper, D. I. Crystal structure and functional assignment of YfaU, a metal ion dependent class II aldolase from *Escherichia coli* K12. *Biochemistry* **2008**, *47* (38) 9955-9965.
35. Hernández, K., Gómez, A., Joglar, J., Bujons, J., Parella, T., Clapés, P. 2-Keto-3-deoxy-L-rhamnonate aldolase (YfaU) as catalyst in aldol additions of pyruvate to amino aldehyde derivatives. *Adv Synth Catal* **2017**, *359* (12) 2090-2100.
36. Hernández, K., Bujons, J., Joglar, J., Charnock, S. J., de María, P. D., Fessner, W.-D., Clapés, P. Combining aldolases and transaminases for the synthesis of 2-amino-4-hydroxybutanoic acid. *ACS Catal* **2017**, *7* (3) 1707-1711.
37. Wang, W., Seah, S. Y. K. Purification and biochemical characterization of a pyruvate-specific class II aldolase, HpaI. *Biochemistry* **2005**, *44* (27) 9447-9455.

38. Rea, D., Fülöp, V., Bugg, T. D. H., Roper, D. I. Structure and mechanism of HpcH: a metal ion dependent class II aldolase from the homoprotocatechuate degradation pathway of *Escherichia coli*. *J Mol Biol* **2007**, *373* (4) 866-876.
39. Wang, W., Baker, P., Seah, S. Y. K. Comparison of two metal-dependent pyruvate aldolases related by convergent evolution: substrate specificity, kinetic mechanism, and substrate channeling. *Biochemistry* **2010**, *49* (17) 3774-3782.
40. Coincon, M., Wang, W., Sygusch, J., Seah, S. Y. K. Crystal structure of reaction intermediates in pyruvate class II aldolase: substrate cleavage, enolate stabilization, and substrate specificity. *J Biol Chem* **2012**, *287* (43) 36208-36221.
41. De Berardinis, V., Guérard-Hélaine, C., Darii, E., Bastard, K., Hélaine, V., Mariage, A., Petit, J.-L., Poupard, N., Sánchez-Moreno, I., Stam, M., Gefflaut, T., Salanoubat, M., Lemaire, M. Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chem* **2017**, *19* 519-526.
42. Marsden, S. R., Mestrom, L., Bento, I., Hagedoorn, P.-L., McMillan, D. G. G., Hanefeld, U. CH- π interactions promote the conversion of hydroxypyruvate in a class II pyruvate aldolase. *Adv Synth Catal* **2019**, *361* (11) 2649-2658.
43. Marsden, S. R., Mestrom, L., McMillan, D. G. G., Hanefeld, U. Thermodynamically and kinetically controlled reactions in bioatalysis – from concepts to perspectives. *ChemCatChem* **2019**, *12* (2) 426-437.
44. Burmakov, A. I., Motnyak, L. A., Kunshenko, B. V., Alexeeva, L. A., Yagupolskii, L. M. Treatment of dimethyl (+)-L-tartrate with sulfur tetrafluoride. *J Fluor Chem* **1981**, *19* (2) 151-161.
45. Hernández, K., Joglar, J., Bujons, J., Parella, T., Clapés, P. Nucleophile promiscuity of engineered class II pyruvate aldolase YfaU from *E. coli*. *Angew Chem Int Ed* **2018**, *57* (14) 3583-3587.
46. Marín-Valls, R., Hernández, K., Bolte, M., Joglar, J., Bujons, J., Clapés, P. Chemoenzymatic hydroxymethylation of carboxylic acids by tandem stereodivergent biocatalytic aldol reaction and chemical decarboxylation. *ACS Catal* **2019**, *9* (8) 7568-7577.
47. Bea, H.-S., Lee, S.-H., Yun, H. Asymmetric synthesis of (R)-3-fluoroalanine from 3-fluoropyruvate using omega-transaminase. *Biotechnol Bioprocess Eng* **2011**, *16* (2) 291-296.
48. Fateev, I. V., Antonov, K. V., Konstantinova, I. D., Muravyova, T. I., Seela, F., Esipov, R. S., Miroshnikov, A. I., Mikhailopulo, I. A. The chemoenzymatic synthesis of clofarabine and related 2'-deoxyfluoroarabinosyl nucleosides: the electronic and stereochemical factors determining substrate recognition by *E. coli* nucleoside phosphorylases. *Beilstein J Org Chem* **2014**, *10* 1657-1669.

Chapter 3: *Expanding the substrate scope of HpCH aldolase for construction of tertiary fluorides*

Portions of this work were published in the following scientific journal:

Fang, J., Turner, L. E., Chang, M. C. Y. Biocatalytic asymmetric construction of secondary and tertiary fluorides from β -fluoro- α -ketoacids. *Angew Chem Int Ed* **2022**, e202201602.

Portions of this work were performed in collaboration with the following persons:

Laura E. Turner performed cloning and purification of HpCH aldolase variants and conducted preliminary aldol addition assays. Diptarka Hait performed DFT calculations on the carbon acidity of fluorinated donors.

3.1. Introduction

In the previous chapter, the fundamentals of a Type II pyruvate aldolase platform for the synthesis of chiral organofluorines was established. Although marking a significant advance in biocatalysis for fluorinated molecules, the initial study was limited to fluoropyruvate as a novel donor substrate, and thus all products contained a secondary fluoride stereocenter. On the contrary, fully-substituted tertiary fluoride stereocenters are more difficult to construct, and addressing this challenge has been a research topic of major interest. This imbalance is exemplified by the fact that only three approved drugs contain a tertiary fluoride stereocenter—the anti-allergy and asthma medication fluticasone, the antibiotic solithromycin, and the hepatitis C therapy sofosbuvir [1]. (Note: the chiral organofluorines here designated as secondary and tertiary may in other literature be referred to as tertiary and quaternary, respectively. Inconsistencies in terminology arise from whether carbon or halogen descriptors are used, and whether one’s definition of a quaternary stereocenter allows heteroatoms.)

The blockbuster success of fluticasone, solithromycin, and sofosbuvir shows that the underuse of tertiary fluoride stereocenters in medicinal chemistry is not due to a lack of medicinal relevance of these motifs, but due to general difficulties in constructing them (*Figure 3.1A*). In all three of these drugs, tertiary fluoride stereocenters are installed by the action of simple achiral fluorinating agents upon a suitably chiral precursor. The tertiary fluoride of fluticasone is installed by epoxide opening with HF, that of solithromycin by electrophilic fluorination with Selectfluor, and that of sofosbuvir by deoxofluorination with DAST [1-3]. The resulting configuration of fluoride stereocenters is induced by the chiral substrate’s structure and interactions, and so this fluorination approach is difficult to generalize. Rather than rely on chiral induction, modern methods towards stereogenic tertiary fluorides seek to generalize the asymmetric fluorination of achiral substrates, most commonly by combining electrophilic N-F reagents with chiral catalysts [4-7]. A parallel approach involves the use of novel N-F reagents that are chiral themselves [8-10]. The use of chiral catalysts in C-C bond forming reactions to transform achiral organofluorines into chiral ones is also highly suitable [11-15]. Nucleophilic fluorination is the least common mode of construction for stereogenic tertiary fluorides (direct S_N2-type substitution is sterically precluded and S_N1-type substitution leads to poor enantioselectivity) but has been demonstrated with indirect fluorocyclization methods that combine nucleophilic fluoride sources with hypervalent iodine catalysts [16].

As various chemical methods towards tertiary fluoride stereocenters proliferate, it becomes prudent to investigate biocatalytic approaches as well for efficiency and sustainability advantages. However, no enzymes catalyze electrophilic fluorination, and only one enzyme is known to catalyze nucleophilic fluorination on a highly specific substrate [17]. In the space of fluorinated natural products, the only one to feature a tertiary fluoride is nucleocidin, and its biosynthetic mechanism of fluorination is not yet known [18]. As discussed in Chapter 2, biocatalytic asymmetric synthesis of secondary fluorides has been accomplished by our group and others by exploiting various families of aldolases (*Figure 3.1B*). However, biocatalytic asymmetric synthesis of tertiary fluorides has been accomplished only by employing hydrolytic enzymes in the kinetic resolution of racemic tertiary fluorides or desymmetrization of achiral tertiary fluorides [19-21]. These methods use precursors with the same connective complexity as the product, and do not take advantage of the flexibility of C-C bond formation (*Figure 3.1C*).

We envisioned that expanding the fluorinated donor scope of a Type II HpcH family pyruvate aldolase from fluoropyruvate to a broader set of β -fluoro- α -ketoacids would expand the product

space to include stereogenic tertiary fluorides. This goal raised the question of to what extent non-native donors could be tolerated by HpcH aldolases, as fluoropyruvate is sterically similar to pyruvate and no other donors were known at the time of our undertaking investigations with fluoropyruvate in 2016. However, increased interest in the HpcH aldolases in the last five years has revealed a donor scope more promiscuous than expected, including hydroxypyruvate, linear α -ketoacids of up to eight carbons, and some branched and cycloalkyl-substituted α -ketoacids [22-24]. Bulky donors that were unreactive with wild-type aldolases could be accommodated with active site engineering. These studies conceptually inspired our work in this chapter, in which we synthesized and investigated β -fluoro- α -ketoacids larger than fluoropyruvate as substrates of HpcH and its engineered variants (*Figure 3.1D*). These larger fluoro-donors were highly challenging due to being non-native in both steric and electronic aspects, and our endeavor was met with partial success and useful lessons learned. Nevertheless, as a result of pushing the boundaries of aldolase substrate scope, we were able to report the first biocatalytic construction of tertiary fluoride stereocenters by C-C bond formation [25].

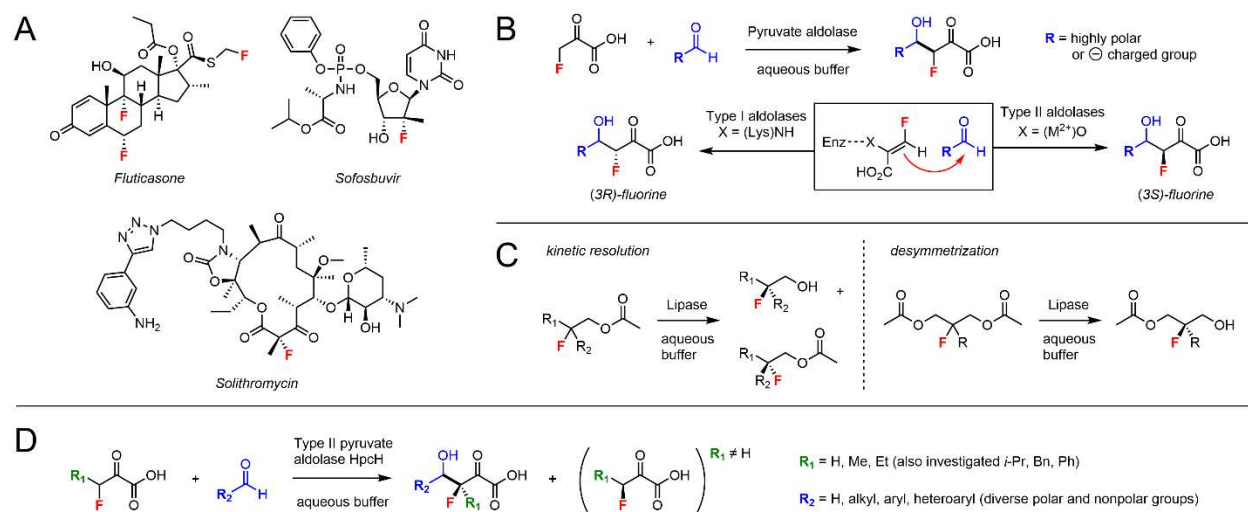


Figure 3.1. Motivation and methodology for enzymatic construction of C-F stereocenters. A) Approved drugs with C-F stereocenters. B) Secondary fluorides from enzymatic aldol addition of fluoropyruvate. C) Lipase-catalyzed formation of tertiary fluorides. D) This chapter: secondary and tertiary fluorides from enzymatic aldol addition of β -fluoro- α -ketoacids to diverse aldehydes, accompanied by kinetic resolution of racemic substrates.

3.2. Materials and methods

Commercial materials. Acetaldehyde, ammonium formate (LC-MS grade), ammonium persulfate, benzaldehyde, *n*-butyllithium (2.5 M in hexanes), catalase from bovine liver, CHES, chlorotrimethylsilane (TMS-Cl), chromium trioxide, cobalt chloride hexahydrate, copper(II) chloride dihydrate, cyclohexanecarboxaldehyde, DIBAL-H (1 M in toluene), diisopropylamine, dimethoxyacetaldehyde (60 wt.% in H₂O), 2,2-dimethoxypropane, 4-dimethylaminopyridine (DMAP), ethyl bromofluoroacetate, 2-ethylbutyraldehyde, ethyl 2-oxo-4-phenylbutyrate, β -fluoropyruvic acid sodium salt monohydrate, 13.4 M formaldehyde, furfural, D-glyceraldehyde, L-glyceraldehyde, glycolaldehyde dimer, glyoxylic acid monohydrate, HEPPS (EPPS), HF-pyridine, iron(II) chloride, isobutyraldehyde, isovaleraldehyde, L-lactate dehydrogenase from rabbit muscle (LDH), lithium hydroxide monohydrate, lysozyme, 2-mercaptoethanol, 3 M methanolic HCl, (R)-(+)- α -methoxy- α -trifluoromethylphenylacetic acid ((+)-MTPA), methyl chloroacetate, methyl 2-

oxobutanoate, 4-methyl-2-oxovaleric acid, nickel chloride hexahydrate, β -nicotinamide adenine dinucleotide hydrate (NAD⁺), β -nicotinamide adenine dinucleotide reduced disodium salt hydrate (NADH), oxalyl chloride (2 M in CH₂Cl₂), 2-oxovaleric acid, phenylmethanesulfonyl fluoride (PMSF), polyethyleneimine (PEI), propionaldehyde, 2-pyridinecarboxaldehyde, 3-pyridinecarboxaldehyde, 4-pyridinecarboxaldehyde, Selectfluor, sodium methoxide, sodium phosphate dibasic heptahydrate, sodium pyruvate, sodium dodecyl sulfate, TAPS, TEMED, 2-thiophenecarboxaldehyde, zinc chloride, and zinc dust were purchased from Sigma-Aldrich (St. Louis, MO). Agarose, Bicine, bromophenol blue, calcium chloride dihydrate, CAPS, carbenicillin disodium salt, deoxynucleotides (dNTPs), dithiothreitol (DTT), EDTA disodium salt dihydrate, formic acid (LC-MS), glycerol, glycine, HEPES, 30 wt.% H₂O₂, imidazole, kanamycin sulfate, manganese chloride tetrahydrate, O'GeneRuler 1 kb Plus DNA Ladder, PageRuler Plus Prestained Protein Ladder, PEG 3350, PEG 8000, potassium chloride, sodium phosphate monobasic monohydrate, TCEP, trichloroacetic acid, Tricine, Tris base, and urea were purchased from Thermo Fisher Scientific (Waltham, MA). LB Miller agar, LB Miller broth, magnesium chloride hexahydrate, Terrific broth (TB), and triethylamine were from EMD-Millipore (Burlington, MA). Ethyl 2-fluoropropionate was from TCI (Tokyo, Japan). InstantBlue Protein Stain was from Expedeon (San Diego, CA). Isopropyl β -D-1-thiogalactopyranoside (IPTG) was from Santa Cruz Biotechnology (Dallas, TX). Acrylamide-bis 30% solution, Bradford reagent, and ethidium bromide were from Bio-Rad (Hercules, CA). Chloroform-*d* and deuterium oxide were from Cambridge Isotope Laboratories (Tewksbury, MA). Phusion polymerase, restriction enzymes, Taq ligase, and T5 exonuclease were from New England Biolabs (Ipswich, MA). Ni-NTA agarose resin and DNA purification kits were from Qiagen (Redwood City, CA). Oligonucleotides and gBlocks gene fragments were synthesized by Integrated DNA Technologies (Coralville, IA). All chemicals were used as purchased without further purification.

Bacterial strains. *Escherichia coli* DH10B-T1^R was used for plasmid construction and *E. coli* BL21(DE3)-T1^R was used for protein production. The strains were made chemically competent by the method of [26].

Construction of expression plasmids. Plasmids were constructed by Gibson isothermal assembly and verified by sequencing (Quintara Biosciences; South San Francisco, CA and Genewiz; South Plainfield, NJ). *In silico* sequence alignments were performed on Benchling (Benchling, San Francisco, CA). All DNA sequences used in this study are listed in *Appendix 3*. To construct pET16hp-HpcH, the parent plasmid pET16hp-IMDH was digested with NdeI/BamHI. To construct pSV272.1-HpcH, the parent plasmid pSV272.1 was digested with SfoI/HindIII. The vector fragments were purified by extraction from 1% agarose gel. The gene encoding HpcH from the *E. coli* C strain was obtained as a codon-optimized gBlock with overhangs homologous to pET16hp, and was resuspended in water to 50 ng/ μ L. To subclone HpcH into pSV272.1, the gene was amplified from pET16hp-HpcH using primers with overhangs homologous to pSV272.1. Primers were reconstituted to 100 μ M in water. The 50 μ L PCR mixture contained 5 \times Phusion HF buffer, 0.2 mM dNTPs, 10 ng template plasmid, 2 units Phusion HF polymerase, and 0.5 μ M each primer. Thermal cycling was as follows: 3 min at 98 $^{\circ}$ C, followed by 35 cycles of 30 s at 98 $^{\circ}$ C, 30 s at (T_m-5) $^{\circ}$ C, and 1 min at 72 $^{\circ}$ C, followed by a final extension of 5 min at 72 $^{\circ}$ C. A 5 μ L sample was visualized on 1% agarose gel and the remaining material was purified by spin column.

Purified vectors and inserts were combined (9:1 v/v) in 5 μ L volume, to which 15 μ L of Gibson master mix, prepared according to [27], was added. The mixture was incubated in a thermocycler

at 50 °C for 1 h, then added to chemically competent *E. coli* DH10B-T1^R along with 20 µL of 5× KCM solution (0.5 M KCl, 0.15 M CaCl₂, 0.25 M MgCl₂) and water to 200 µL total. After incubation on ice for 20 min, the cells were heat-shocked at 42 °C for 90 s and returned to ice for 2 min. Cells were diluted with 1 mL of LB, recovered with shaking at 37 °C for 45 min, and pelleted by centrifugation (7,000 × g, 1 min). The supernatant was partially removed (0.8 mL), then the cells were resuspended and plated on LB agar with 50 µg/mL antibiotic. Plates were incubated at 37 °C overnight. Liquid cultures (5 mL LB with 50 µg/mL antibiotic) were inoculated with single colonies and grown at 37 °C overnight, following which the plasmids were isolated and confirmed by sequencing from the T7 promoter and T7 terminator. Plasmid pSV272.1-HpcH was also sequenced with the primer HpcH-seqL83-R to confirm MBP fusion.

Point mutations in the plasmid pET16hp-HpcH were introduced by the two-step modification of the Quikchange site-directed mutagenesis method [28]. Mutagenic primers were reconstituted to 100 µM in water. Single primer extension reactions (50 µL) were prepared containing 5X Phusion HF buffer, 0.2 mM dNTPs, 20 ng template, 1 unit Phusion HF polymerase, and 0.5 µM forward or reverse primer. Thermal cycling was as follows: 3 min at 98 °C, followed by 10 cycles of 30 s at 98 °C, 30 s at 65 °C, and 3 min at 72 °C, followed by a final extension of 5 min at 72 °C. Then, 25 µL of each primer extension reaction was combined and supplemented with 1 unit Phusion HF polymerase. Thermal cycling was continued for 25 cycles. DpnI (2 units) was added to digest the template for 1 h at 37 °C. *E. coli* DH10B-T1^R was transformed with 5 µL of the mixture (KCM method), then plated on LB agar with 50 µg/mL carbenicillin and incubated at 37 °C overnight. Liquid cultures were inoculated with single colonies and grown at 37 °C overnight. The plasmids were isolated and mutations confirmed by sequencing.

Expression and purification of His-tagged proteins. Expression plasmids (10 ng) were transformed into *E. coli* BL21(DE3)-T1^R (KCM method) and a single colony was used to inoculate 25 mL of TB media with 50 µg/mL of the appropriate antibiotic. The seed culture was grown overnight with shaking at 37 °C (200 RPM). Then, 1 L of TB in an Ultra Yield baffled flask (Thomson Instrument Company; Oceanside, CA) was inoculated with the seed culture and growth was continued to OD₆₀₀ of 0.8-1.2. The culture was chilled in an ice bath for 15 min and protein expression was induced with IPTG (1 mM). Expression proceeded at 16 °C overnight, after which the cells were harvested (7,000 × g, 5 min, 4 °C) and either stored at -80 °C or subjected to protein purification immediately.

Harvested cells were resuspended in 5 mL/(g wet cell wt.) of lysis buffer (20 mM Tris pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% v/v glycerol). To the cells was added PMSF (1 mM) and lysozyme (1 mg/mL). After incubation at r.t. for 30 min, the sample was sonicated (QSonica Q700) with the following program: 10 s on, 20 s off, 1 min total process time, amplitude 50. Cell debris was removed by centrifugation (15,000 × g, 20 min, 4 °C). To the supernatant was added PEI (0.05% w/v) and the precipitated nucleic acids were removed by centrifugation (15,000 × g, 20 min, 4 °C). Ni-NTA agarose resin (50% suspension in 20% EtOH, 0.2 mL/(g wet cell wt.)) was added to the samples, which were shaken gently on an orbital shaker at r.t. for 30 min and then poured into a glass column. The resin was washed with >20 column volumes (CV) of wash buffer (same as lysis buffer with 20 mM imidazole) until the effluent tested negative by the Bradford dye-binding assay. Protein was eluted with >5 CV of elution buffer (same as lysis buffer with 250 mM imidazole) until the effluent tested negative again. The elution fraction was supplemented with β-ME (0.1% v/v) and MgCl₂ (1 mM). Fractions and purified proteins were analyzed by SDS-PAGE.

The proteins were concentrated in Amicon Ultra-15 centrifugal filter units (EMD Millipore; Burlington, MA) with molecular weight cutoff of 10 kDa for HpcH aldolase and its mutants, and molecular weight cutoff of 30 kDa for MBP-HpcH. Note that HpcH aldolase and its mutants have poor solubility at 4 °C if not fused to MBP, and concentration past 2 mg/mL should be performed at room temperature. Buffer exchange was conducted with PD-10 columns (GE Healthcare; Chicago, IL) into HEPES storage buffer (50 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% v/v glycerol, 1 mM TCEP). Desalted samples were supplemented with MgCl₂ (1 mM) and further concentrated to 150 μM. Protein concentrations were determined by the absorbance at 280 nm, using extinction coefficients (wt-HpcH and all mutants except W19A: 32430 M⁻¹cm⁻¹, HpcH-W19A: 26930 M⁻¹cm⁻¹, MBP-HpcH: 100270 M⁻¹cm⁻¹) predicted with ExPASy ProtParam (Swiss Institute of Bioinformatics).

Safety note on handling of organofluorine compounds. Fluoroacetate is an extremely toxic metabolic poison. The decarboxylative treatment of fluoropyruvate aldol reactions, used throughout this chapter, generates fluoroacetate in solution. Such samples should be handled with great care. Other organofluorines including fluoropyruvate, larger β-fluoro-α-ketoacids and their decarboxylated products, and novel synthesized compounds may have uncharacterized toxicity and warrant general handling precautions.

General remarks for 1D-NMR spectroscopy. NMR was performed at the UC Berkeley College of Chemistry NMR Facility. Characterization of isolated compounds (¹H, ¹³C, and ¹⁹F NMR) and ¹⁹F NMR aldolase assays were performed on a Bruker AV-600 (600 MHz) equipped with a Z-gradient broadband cryoprobe (supported by NIH S10OD024998). ¹H NMR assays for native activity of aldolases was performed on a Bruker AV-500 (500 MHz) equipped with a Z-gradient broadband probe. All experiments were conducted at 298 K. Chemical shifts were referenced to the solvent and expressed relative to tetramethylsilane (¹H and ¹³C) or trichlorofluoromethane (¹⁹F). The spectral center of ¹⁹F NMR was -150 ppm and the spectral width was 300 ppm. In general, the number of ¹H or ¹⁹F scans was 8, and the number of ¹³C scans was >256. For ¹⁹F NMR aldolase assays, the number of ¹⁹F scans was 1. Data was processed in MestreNova (MestreLab Research). Baseline distortion in the -150 to -200 ppm region of ¹⁹F NMR due to fluoropolymer in the probes was corrected with backward linear prediction (Toeplitz) of the first 128 data points, followed by manual phase correction and baseline correction (Whittaker smoother).

General procedures for synthesis of 3-fluoro-2-oxoacids (2a-d).

Silylation procedure for synthesis of 2-oxoester silyl enol ethers. To a solution of 2-oxoester (20 mmol) in CH₃CN (40 mL) under N₂ was added TMS-Cl (2.39 g, 2.79 mL, 22 mmol) followed by Et₃N (2.43 g, 3.35 mL, 24 mmol). The mixture was stirred vigorously at r.t. for 4 h, then diluted with H₂O (60 mL) and extracted with hexanes (3 × 60 mL). The organic layer was washed with brine (60 mL), dried over Na₂SO₄ and concentrated. The silyl enol ether was purified by silica gel chromatography using the solvents indicated below.

Fluorination procedure for synthesis of 3-fluoro-2-oxoesters. To a solution of silyl enol ether (20 mmol) in CH₃CN (80 mL) was added finely crushed Selectfluor (7.79 g, 22 mmol). The suspension was stirred at r.t. overnight, then the solvent was evaporated. The product was extracted from the white solids by trituration with Et₂O (2 × 100 mL) which was then filtered and concentrated. The 3-fluoro-2-oxoester was purified by silica gel chromatography using the

solvents indicated below. For mixtures of keto and hydrate forms, reported yields are based on the keto molecular weight.

Saponification procedure for synthesis of 3-fluoro-2-oxoacids. To a solution of 3-fluoro-2-oxoester (5 mmol, keto wt.) in CH₃CN (20 mL) was added LiOH-H₂O (231 mg, 5.5 mmol). The reaction was stirred at r.t. for 1 h, then the solvent was evaporated. The residue was dissolved in a minimal amount of water. After washing with EtOAc (2 × 30 mL) the aqueous layer was acidified to pH 0 with 6 M HCl and the product was extracted with EtOAc (3 × 30 mL). The organic layer was washed with brine (30 mL), dried over Na₂SO₄ and concentrated. The free acid was weighed, then neutralized according to the following procedure. A solution of the acid in CH₃CN (20 mL) was treated with quantitatively-prepared 3M NaOH (1.0 eq. hydrate wt.) and stirred vigorously for 15 min. The solvent was evaporated and the sodium salt dried under vacuum.

3-Fluoro-2-oxobutanoic acid (FKB, 2a).

TMS-enol ether of methyl 2-oxobutanoate. Methyl 2-oxobutanoate (2.32 g, 20 mmol) was silylated by the general procedure and chromatographed (29:1 hexanes/Et₂O) to give the silyl enol ether as a clear liquid (1.83 g, 49%). ¹H NMR (600 MHz, CDCl₃) δ 6.15 (q, *J* = 7.1 Hz, 1H), 3.76 (s, 3H), 1.72 (d, *J* = 7.1 Hz, 3H), 0.24 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 165.2, 141.4, 118.4, 51.8, 11.3, 0.4.

Methyl 3-fluoro-2-oxobutanoate. The silyl enol ether of methyl 2-oxobutanoate (3.77 g, 20 mmol) was fluorinated by the general procedure and chromatographed (1:1 hexanes/Et₂O) to give the fluoroester as a yellow oil (2.24 g, 84%, keto/hydrate 31:69). ¹H NMR (600 MHz, CDCl₃) Keto δ 5.53 (dq, *J* = 48.0, 7.0 Hz, 1H), 3.93 (s, 3H), 1.64 (dd, *J* = 23.7, 7.0 Hz, 3H); Hydrate δ 4.87 (dq, *J* = 46.8, 6.4 Hz, 1H), 4.31 (s, 1H), 3.91 (s, 3H), 3.72 (s, 1H), 1.42 (dd, *J* = 24.7, 6.4 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) Keto δ 190.9 (d, *J* = 23.8 Hz), 160.8 (d, *J* = 1.8 Hz), 89.7 (d, *J* = 181.0 Hz), 53.2, 17.0 (d, *J* = 22.2 Hz); Hydrate δ 171.1 (d, *J* = 1.2 Hz), 93.4 (d, *J* = 25.0 Hz), 90.5 (d, *J* = 174.2 Hz), 53.7, 13.9 (d, *J* = 22.1 Hz). ¹⁹F NMR (565 MHz, CDCl₃) Keto δ -189.5 (dq, *J* = 47.9, 23.9 Hz); Hydrate δ -188.8 (dq, *J* = 48.9, 24.7 Hz).

3-Fluoro-2-oxobutanoic acid (FKB, 2a). Methyl 3-fluoro-2-oxobutanoate (671 mg, 5 mmol) was saponified by the general procedure to give the fluoroacid as a cloudy syrup (273 mg, 40%) which was converted to a sodium salt prior to NMR analysis and experimental use. In aqueous solution the salt becomes ~91% hydrated, and spectral data is given only for the hydrate. ¹H NMR (600 MHz, D₂O) δ 4.79 (dq, *J* = 47.3, 6.5 Hz, 1H), 1.26 (dd, *J* = 25.5, 6.4 Hz, 3H). ¹³C NMR (151 MHz, D₂O) δ 175.7 (d, *J* = 2.8 Hz), 94.4 (d, *J* = 22.2 Hz), 92.3 (d, *J* = 170.9 Hz), 14.1 (d, *J* = 21.8 Hz). ¹⁹F NMR (565 MHz, D₂O) δ -188.6 (dq, *J* = 50.5, 25.8 Hz).

3-Fluoro-2-oxopentanoic acid (FKV, 2b).

Methyl 2-oxopentanoate. 2-Oxovaleric acid (2.32 g, 20 mmol) was dissolved in 2,2-dimethoxypropane (40 mL) and MeOH (10 mL). TMS-Cl (217 mg, 254 μL, 2 mmol) was added and the mixture was stirred at r.t. overnight. The solvent was evaporated and the residue purified by silica gel chromatography (4:1 hexanes/Et₂O) to give the methyl ester as a yellow liquid (2.34 g, 90%). ¹H NMR (600 MHz, CDCl₃) δ 3.88 (s, 3H), 2.84 (t, *J* = 7.2 Hz, 2H), 1.69 (sext, *J* = 7.3 Hz, 2H), 0.98 (t, *J* = 7.4 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 194.2, 161.6, 52.8, 41.2, 16.5, 13.5.

TMS-enol ether of methyl 2-oxopentanoate. Methyl 2-oxopentanoate (2.60 g, 20 mmol) was silylated by the general procedure and chromatographed (29:1 hexanes/Et₂O) to give the silyl enol

ether as a clear liquid (2.79 g, 69%). ^1H NMR (600 MHz, CDCl_3) δ 6.07 (t, $J = 7.5$ Hz, 1H), 3.77 (s, 3H), 2.19 (quint, $J = 7.5$ Hz, 2H), 1.04 (t, $J = 7.6$ Hz, 3H), 0.23 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 165.4, 139.9, 125.4, 51.8, 19.2, 13.2, 0.4.

Methyl 3-fluoro-2-oxopentanoate. The silyl enol ether of methyl 2-oxopentanoate (4.05 g, 20 mmol) was fluorinated by the general procedure and chromatographed (4:1 hexanes/ Et_2O) to give the fluoroester as a yellow liquid (2.23 g, 75%, keto/hydrate 53:47.). ^1H NMR (600 MHz, CDCl_3) Keto δ 5.37 (ddd, $J = 48.7, 7.6, 4.2$ Hz, 1H), 3.93 (s, 3H), 2.14-1.88 (m, 2H), 1.09 (t, $J = 3.8$ Hz, 3H); Hydrate δ 4.60 (ddd, $J = 47.4, 9.8, 2.9$ Hz, 1H), 4.17 (s, 1H), 3.91 (s, 3H), 3.45 (s, 1H), 1.87-1.66 (m, 2H), 1.07 (t, $J = 3.8$ Hz, 3H). ^{13}C NMR (151 MHz, CDCl_3) Keto δ 190.9 (d, $J = 23.9$ Hz), 161.0 (d, $J = 2.1$ Hz), 94.0 (d, $J = 184.1$ Hz), 53.1, 24.7 (d, $J = 21.4$ Hz), 8.7 (d, $J = 4.0$ Hz); Hydrate δ 171.0 (d, $J = 1.5$ Hz), 95.4 (d, $J = 176.8$ Hz), 93.3 (d, $J = 25.3$ Hz), 53.7, 21.5 (d, $J = 21.1$ Hz), 9.7 (d, $J = 3.9$ Hz). ^{19}F NMR (565 MHz, CDCl_3) Keto δ -198.8 (ddd, $J = 49.9, 27.8, 23.2$ Hz); Hydrate δ -197.8 (ddd, $J = 47.2, 39.5, 15.7$ Hz).

3-Fluoro-2-oxopentanoic acid (FKV, 2b). Methyl 3-fluoro-2-oxopentanoate (741 mg, 5 mmol) was saponified by the general procedure to give the fluoroacid as a cloudy syrup (403 mg, 53%) which was converted to a sodium salt prior to NMR analysis and experimental use. In aqueous solution the salt becomes ~68% hydrated. ^1H NMR (600 MHz, D_2O) Keto δ 5.46 (ddd, $J = 49.0, 6.9, 4.4$ Hz, 1H), 2.06-1.77 (m, 2H), 0.93 (t, $J = 7.4$ Hz, 3H); Hydrate δ 4.53 (ddd, $J = 47.9, 8.9, 3.7$ Hz, 1H), 1.66-1.47 (m, 2H), 0.95 (t, $J = 7.4$ Hz, 3H); ^{13}C NMR (151 MHz, D_2O) Keto δ 202.2 (d, $J = 20.6$ Hz), 168.6 (d, $J = 8.3$ Hz), 95.3 (d, $J = 180.0$ Hz), 24.2 (d, $J = 20.9$ Hz), 7.9 (d, $J = 3.9$ Hz). Hydrate δ 175.7 (d, $J = 2.9$ Hz), 97.4 (d, $J = 174.0$ Hz), 94.3 (d, $J = 22.1$ Hz), 21.9 (d, $J = 20.9$ Hz), 9.2 (d, $J = 3.8$ Hz). ^{19}F NMR (565 MHz, D_2O) Keto δ -198.8 (dt, $J = 52.2, 27.2$ Hz); Hydrate δ -198.4 (m).

3-Fluoro-4-methyl-2-oxopentanoic acid (FKMV, 2c).

Methyl 4-methyl-2-oxopentanoate. 4-Methyl-2-oxovaleric acid (2.60 g, 20 mmol) was dissolved in 2,2-dimethoxypropane (40 mL) and methanol (10 mL). Chlorotrimethylsilane (217 mg, 254 μL , 2 mmol) was added and the mixture was stirred at r.t. overnight. The solvent was evaporated and the residue purified by silica gel chromatography (9:1 hexanes/ Et_2O) to give the methyl ester as a yellow liquid (2.63 g, 91%). ^1H NMR (600 MHz, CDCl_3) δ 3.88 (s, 3H), 2.73 (d, $J = 6.9$ Hz, 2H), 2.21 (non, $J = 6.7$ Hz, 1H), 0.98 (d, $J = 6.7$ Hz, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 194.0, 161.7, 52.9, 47.9, 24.2, 22.4.

TMS-enol ether of methyl 4-methyl-2-oxopentanoate. Methyl 4-methyl-2-oxopentanoate (2.88 g, 20 mmol) was silylated by the general procedure and chromatographed (29:1 hexanes/ Et_2O) to give the silyl enol ether as a clear liquid (2.95 g, 68%). ^1H NMR (600 MHz, CDCl_3) δ 5.92 (d, $J = 9.6$ Hz, 1H), 3.77 (s, 3H), 2.77 (dsept, $J = 9.7, 6.7$ Hz, 1H), 1.03 (d, $J = 6.7$ Hz, 6H), 0.24 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 165.6, 138.6, 130.5, 51.8, 25.4, 22.1, 0.4.

Methyl 3-fluoro-4-methyl-2-oxopentanoate. The silyl enol ether of methyl 4-methyl-2-oxopentanoate (4.33 g, 20 mmol) was fluorinated by the general procedure and chromatographed (9:1 hexanes/ Et_2O) to give the fluoroester as a yellow liquid (2.59 g, 80%, keto/hydrate 88:12). The spectral data is given only for the keto form. ^1H NMR (600 MHz, CDCl_3) δ 5.22 (dd, $J = 48.6, 4.1$ Hz, 1H), 3.93 (s, 3H), 2.39 (m, 1H), 1.14 (d, $J = 6.9$ Hz, 3H), 0.99 (d, $J = 6.9$ Hz, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 191.2 (d, $J = 24.5$ Hz), 161.4 (d, $J = 2.2$ Hz), 96.8 (d, $J = 186.5$ Hz), 53.1, 30.4 (d, $J = 20.5$ Hz), 18.5 (d, $J = 3.9$ Hz), 15.9 (d, $J = 5.3$ Hz). ^{19}F NMR (565 MHz, CDCl_3) δ -207.0 (dd, $J = 48.6, 26.6$ Hz).

3-Fluoro-4-methyl-2-oxopentanoic acid (FKMV, 2c). Methyl 3-fluoro-4-methyl-2-oxopentanoate (811 mg, 5 mmol) was saponified by the general procedure to give the fluoroacid as a cloudy syrup (636 mg, 77%) which was converted to a sodium salt prior to NMR analysis and experimental use. In aqueous solution the salt becomes ~25% hydrated. ¹H NMR (600 MHz, D₂O) Keto δ 5.38 (dd, *J* = 48.9, 2.6 Hz, 1H), 2.36-2.23 (m, 1H), 1.04 (d, *J* = 7.0 Hz, 3H), 0.84 (d, *J* = 7.0 Hz, 3H); Hydrate δ 4.38 (dd, *J* = 46.5, 6.4 Hz, 1H), 1.99-1.88 (m, 1H), 0.94 (d, *J* = 6.9 Hz, 3H), 0.91 (d, *J* = 6.9 Hz, 3H). ¹³C NMR (151 MHz, D₂O) Keto δ 202.3 (d, *J* = 19.9 Hz), 168.8 (d, *J* = 2.3 Hz), 98.1 (d, *J* = 183.5 Hz), 29.8 (d, *J* = 20.1 Hz), 17.9 (d, *J* = 3.3 Hz), 14.9 (d, *J* = 5.6 Hz); Hydrate δ 175.8 (d, *J* = 3.3 Hz), 99.6 (d, *J* = 176.3 Hz), 94.5 (d, *J* = 22.2 Hz), 28.2 (d, *J* = 20.5 Hz), 19.0 (d, *J* = 6.5 Hz), 17.1 (d, *J* = 7.5 Hz). ¹⁹F NMR (565 MHz, D₂O) Keto δ -208.6 (dd, *J* = 48.9, 30.6 Hz); Hydrate δ -200.7 (dd, *J* = 46.2, 19.0 Hz).

3-Fluoro-2-oxo-4-phenylbutanoic acid (FKPB, 2d).

TMS-enol ether of ethyl 2-oxo-4-phenylbutanoate. Ethyl 2-oxo-4-phenylbutanoate (4.12 g, 20 mmol) was silylated by the general procedure with the modification that the product was extracted with 1:1 hexanes/toluene, then chromatographed (20:9:1 hexanes/toluene/Et₂O) to give the silyl enol ether as a clear liquid (3.57 g, 64%). ¹H NMR (600 MHz, CDCl₃) δ 7.35-7.16 (m, 5H), 6.25 (t, *J* = 7.6 Hz, 1H), 4.24 (q, *J* = 7.1 Hz, 2H), 3.54 (d, *J* = 7.6 Hz, 2H), 1.33 (t, *J* = 7.1 Hz, 3H), 0.28 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 164.8, 141.0, 139.7, 128.53, 128.47, 126.2, 121.3, 61.0, 32.0, 14.2, 0.6.

Ethyl 3-fluoro-2-oxo-4-phenylbutanoate. The silyl enol ether of ethyl 2-oxo-4-phenylbutanoate (5.57 g, 20 mmol) was fluorinated by the general procedure and chromatographed (1:1 hexanes/Et₂O) to give the fluoroester as bright yellow syrup (3.75 g, 84%, keto/hydrate 19:81). ¹H NMR (600 MHz, CDCl₃) δ 7.38-7.21 (m, 5H), 5.64 (ddd, *J* = 48.4, 8.3, 4.0 Hz, 1H keto), 4.92 (ddd, *J* = 46.9, 8.6, 4.1 Hz, 1H hydrate), 4.40-4.24 (m, 2H), 4.33 (s, 1H hydrate), 3.49 (s, 1H hydrate), 3.39-2.96 (m, 2H), 1.39 (t, *J* = 7.2 Hz, 3H keto), 1.35 (t, *J* = 7.2 Hz, 3H hydrate). ¹³C NMR (151 MHz, CDCl₃) Keto δ 190.3 (d, *J* = 23.3 Hz), 160.4, 134.6 (d, *J* = 1.6 Hz), 129.4, 128.8, 127.4, 93.2 (d, *J* = 186.4 Hz), 62.9, 37.8 (d, *J* = 21.0 Hz), 13.9; Hydrate δ 170.3, 136.5 (d, *J* = 2.7 Hz), 129.4, 128.6, 126.8, 94.3 (d, *J* = 179.4 Hz), 93.1 (d, *J* = 25.4 Hz), 63.3, 34.8 (d, *J* = 21.2 Hz), 14.0. ¹⁹F NMR (565 MHz, CDCl₃) Keto δ -193.9 (ddd, *J* = 48.4, 30.8, 22.1 Hz); Hydrate δ -193.7 (ddd, *J* = 46.8, 35.0, 19.3 Hz).

3-Fluoro-2-oxo-4-phenylbutanoic acid (FKPB, 2d). Ethyl 3-fluoro-2-oxo-4-phenylbutanoate (1.12 g, 5 mmol) was saponified by the general procedure to give the fluoroacid as a yellow paste-like solid (825 mg, 77%) which was converted to a sodium salt prior to NMR analysis and experimental use. In aqueous solution the salt becomes ~90% hydrated, and spectral data is given only for the hydrate. ¹H NMR (600 MHz, D₂O) δ 7.39-7.23 (m, 5H), 4.89-4.75 (m, 1H), 3.03-2.85 (m, 2H). ¹³C NMR (151 MHz, D₂O) δ 175.6, 137.6, 129.3, 128.8, 126.8, 96.3 (d, *J* = 176.5 Hz), 94.2 (d, *J* = 22.6 Hz), 34.9 (d, *J* = 20.9 Hz). ¹⁹F NMR (565 MHz, D₂O) δ -196.2 (m).

3-Fluoro-3-phenylpyruvic acid (FPP, 2e).

Methyl phenylglycidate. To a mixture of benzaldehyde (4.24 g, 4.07 mL, 40 mmol) and methyl chloroacetate (8.68 g, 7.01 mL, 80 mmol) was slowly added a suspension of sodium methoxide (6.48 g, 120 mmol) in methanol (60 mL). The thick mixture was stirred vigorously for 4 h, then diluted with water (150 mL). The product was extracted with EtOAc (3 × 70 mL), washed with brine (70 mL), dried over Na₂SO₄, and concentrated. Purification by silica gel chromatography (9:1 hexanes/EtOAc) gave the product as a clear liquid (1.69 g, 24%). ¹H NMR (600 MHz, CDCl₃)

δ 7.41-7.36 (m, 3H), 7.34-7.30 (m, 2H), 4.13 (d, J = 1.8 Hz, 1H), 3.86 (s, 3H), 3.54 (d, J = 1.8 Hz, 1H). ^{13}C NMR (151 MHz, CDCl_3) δ 168.6, 134.9, 129.0, 128.7, 125.8, 58.0, 56.6, 52.6.

Methyl 3-fluoro-3-phenyllactate. In a polypropylene tube, methyl phenylglycidate (1.78 g, 10 mmol) was dissolved in dichloromethane (20 mL), then 70% HF/pyridine (0.91 mL, \sim 35 mmol HF) was added dropwise. After stirring 1 h the mixture was diluted with EtOAc (60 mL). The organic layer was washed with sat. NaHCO_3 (2×40 mL), water (40 mL), and brine (40 mL), then dried over Na_2SO_4 and concentrated. Purification by silica gel chromatography (4:1 hexanes/EtOAc) gave the product (3:1 syn/anti mixture of diastereomers) as a yellow oil (1.42 g, 72%). ^1H NMR (600 MHz, CDCl_3) δ 7.46-7.33 (m, 5H), 5.81 (dd, J = 45.1, 2.3 Hz, 1H syn), 5.75 (dd, J = 44.9, 3.8 Hz, 1H anti), 4.69 (ddd, J = 15.0, 7.3, 3.8 Hz, 1H anti), 4.47 (ddd, J = 27.0, 6.8, 2.3 Hz, 1H syn), 3.90 (s, 3H syn), 3.79 (s, 3H anti), 3.03 (d, J = 6.9 Hz, 1H syn), 2.87 (d, J = 7.5 Hz, 1H anti). ^{13}C NMR (151 MHz, CDCl_3) Syn δ 171.8 (d, J = 2.8 Hz), 135.6 (d, J = 20.9 Hz), 128.8 (d, J = 1.2 Hz), 128.4, 125.9 (d, J = 7.6 Hz), 93.2 (d, J = 180.2 Hz), 73.6 (d, J = 23.7 Hz), 53.1; Anti δ 171.3 (d, J = 9.3 Hz), 134.7 (d, J = 21.0 Hz), 129.0 (d, J = 1.2 Hz), 128.4, 126.1 (d, J = 7.2 Hz), 93.7 (d, J = 180.7 Hz), 73.7 (d, J = 25.9 Hz), 52.8. ^{19}F NMR (565 MHz, CDCl_3) Syn δ -194.5 (dd, J = 45.1, 27.0 Hz); Anti δ -188.0 (dd, J = 45.3, 15.3 Hz).

Methyl 3-fluoro-3-phenylpyruvate. The 3-fluoro-3-phenyllactate (1.39 g, 7 mmol) was dissolved in acetone (10 mL) followed by addition of Jones reagent (10 mL, 2.5 M CrO_3 in 25% aq. H_2SO_4). The mixture was stirred overnight, quenched with isopropanol (2 mL), and diluted with water (60 mL). The product was extracted with EtOAc (3×40 mL), washed with sat. NaHCO_3 (40 mL) and brine (40 mL), then dried over Na_2SO_4 and concentrated. Purification by silica gel chromatography (4:1 hexanes/EtOAc) gave the product as a yellow oil (437 mg, 32%, keto/hydrate 66:34). ^1H NMR (600 MHz, CDCl_3) δ 7.52-7.40 (m, 5H), 6.48 (d, J = 47.1 Hz, 1H keto), 5.69 (d, J = 44.6 Hz, 1H hydrate), 3.94 (s, 3H hydrate), 3.84 (s, 3H keto). ^{13}C NMR (151 MHz, CDCl_3) Keto δ 187.0 (d, J = 24.9 Hz), 160.1 (d, J = 3.2 Hz), 131.2 (d, J = 20.4 Hz), 130.3 (d, J = 2.5 Hz), 129.2, 128.0 (d, J = 5.1 Hz), 93.2 (d, J = 187.0 Hz), 53.2; Hydrate δ 170.6, 132.9 (d, J = 20.6 Hz), 129.5 (d, J = 1.2 Hz), 128.2, 127.6 (d, J = 7.0 Hz), 93.3 (d, J = 181.2 Hz), 93.2 (d, J = 28.7 Hz), 53.8. ^{19}F NMR (565 MHz, CDCl_3) Keto δ -183.6 (d, J = 47.0 Hz); Hydrate δ -189.2 (d, J = 44.4 Hz).

3-Fluoro-3-phenylpyruvic acid (FPP, 2e). Methyl 3-fluorophenylpyruvate (588 mg, 3 mmol) was saponified by the general procedure with lithium hydroxide monohydrate (138 mg, 3.3 mmol) to give the fluoroacid as a cloudy yellow syrup (508 mg, 85%), which was converted to a sodium salt prior to NMR analysis and experimental use. In aqueous solution the salt becomes \sim 80% hydrated, and spectral data is given only for the hydrate. ^1H NMR (600 MHz, D_2O) δ 7.48-7.35 (m, 5H), 5.63 (d, J = 44.8 Hz). ^{13}C NMR (151 MHz, D_2O) δ 175.31 (d, J = 2.2 Hz), 134.5 (d, J = 20.2 Hz), 129.1 (d, J = 1.7 Hz), 128.1, 127.4 (d, J = 7.3 Hz), 95.1 (d, J = 176.3 Hz), 94.5 (d, J = 26.1 Hz). ^{19}F NMR (565 MHz, D_2O) δ -189.8 (d, J = 44.3 Hz).

Native activity assay of pyruvate aldolases. Reactions of 0.6 mL total volume were prepared containing 50 mM sodium pyruvate (from 1 M in H_2O , 30 μL), 100 mM glycolaldehyde (from 2 M in H_2O , 30 μL), 0.1 mol% aldolase (from 150 μM solution in storage buffer, 200 μL), and 20 mM HEPES, 1 mM MgCl_2 , pH 7.5 (340 μL). Reactions proceeded at r.t. for 1 h and were quenched by decarboxylation with 60 μL of 30 wt.% H_2O_2 and incubation for 30 min. After removal of precipitated proteins by centrifugation (15,000 $\times g$, 10 min), the samples were frozen at -80°C and lyophilized in a vacuum concentrator, and remaining solids were resuspended in D_2O (700 μL) for ^1H NMR analysis. The extent of conversion was obtained from integration of acetate

methyl protons (2.81 ppm) and 3,4-dihydroxybutanoate methylene protons (2.23 ppm). The native activity of HpcH aldolase, all studied HpcH mutants, and MBP-HpcH was confirmed by observation of >90% conversion. In the absence of enzyme (volume replaced with storage buffer), 0% conversion was observed.

Aldol addition assay of fluorinated donor substrates. Reactions of 0.6 mL total volume were prepared containing 50 mM fluoro-donor substrate, 100 mM acceptor substrate (30 μ L of a 2 M stock solution in most cases), 0.1 mol% aldolase (from 150 μ M solution in storage buffer, 200 μ L), and 20 mM HEPES, 1 mM MgCl₂, pH 7.5 (370 μ L). The desired amount of fluoro-donor was weighed and dissolved in reaction buffer, and this solution was aliquoted before adding acceptors or enzymes. Formaldehyde, acetaldehyde, propionaldehyde, D-glyceraldehyde, glycolaldehyde, dimethoxyacetaldehyde, glyoxylic acid, 2-, 3-, and 4-pyridinecarboxaldehyde were added from 2 M solutions in H₂O. The solution of glyoxylic acid was neutralized with 3 M NaOH prior to use. The expensive acceptor L-glyceraldehyde was applied at only 50 mM, and was added from a 1 M solution in H₂O. The acceptors 2-ethylbutyraldehyde, cyclohexanecarboxaldehyde, benzaldehyde, and 2-thiophenecarboxaldehyde were added from 2 M solutions in DMSO, and thus the reactions contained 5% DMSO v/v as co-solvent. Isobutyraldehyde, isovaleraldehyde, and furfural were added as pure liquids (60 μ mol) followed by supplementary H₂O (to 30 μ L). Reactions proceeded at r.t. for 1 h, 4 h, 16 h, 48 h, or 7 d, and were quenched by decarboxylation with 60 μ L of 30 wt.% H₂O₂ and incubation for 30 min. After removal of precipitated proteins by centrifugation (15,000 \times g, 10 min), D₂O (100 μ L) was added for ¹⁹F NMR analysis. The extent of conversion was obtained from integration of the resulting 2-fluoroacids. The chemical shift of the decarboxylated fluoro-donor was referenced to its average value across all samples (fluoroacetate: -216.979 ppm, 2-fluoropropionate: -173.288 ppm, 2-fluorobutyrate: -183.244 ppm). See *Appendix 2* for full conversion data.

In control reactions during HpcH mutant screening, the volume of enzyme was replaced with storage buffer. In the secondary screen of HpcH mutants at 0.3 mol%, enzyme was added from stocks that had been concentrated three-fold (450 μ M solution in storage buffer, 200 μ L). In reactions of fluoropyruvate accelerated by basic pH, the reaction buffer was 20 mM CHES, 1 mM MgCl₂, pH 9.3. The combination of this buffer with the enzyme solution gives a final pH of 8.5. Various other buffers (HEPPS pH 8.5, Bicine pH 8.5, Tricine pH 8.5, TAPS pH 8.5, CAPS pH 10.5, sodium carbonate pH 10.5, sodium phosphate pH 12.5) and other divalent metals (Mn, Fe, Co, Ni, Cu, Zn) were also tested for their effect on promoting background aldol addition. These non-enzymatic reactions contained only the fluoro-donor and formaldehyde dissolved in the buffer of interest (570 μ L) plus metal dichloride (1 mM).

High resolution LC-MS analysis. Verification of all products observed in ¹⁹F NMR analysis was performed by electrospray ionization (ESI) high resolution LC-MS. A small portion (10 μ L) of each NMR sample was mixed with water (80 μ L) and 100% w/v trichloroacetic acid (10 μ L), and incubated at 4 $^{\circ}$ C for 30 min to precipitate residual proteins. The samples were clarified by centrifugation (15,000 \times g, 20 min, 4 $^{\circ}$ C) and diluted 100-fold into either 0.1% formic acid (for reverse-phase chromatography) or 90% acetonitrile (for HILIC-phase chromatography). An Agilent 1290 Infinity LC system was equipped with an Agilent Poroshell 120 EC-C18 column (2.7 μ m, 2.1 \times 50 mm) or an Ascentis Express HILIC column (2 μ m, 2.1 \times 50 mm) and coupled to an Agilent 6530 Q-TOF MS operated in negative full-scan mode (capillary voltage 3500 V, nozzle voltage 500 V, fragmentor voltage 175 V, scan range 100-1700 *m/z*). Reverse-phase chromatography conditions were as follows. Solvent A: 0.1% v/v formic acid; solvent B:

acetonitrile; flow rate 0.6 mL/min; maximum pressure 600 bar; timetable (%B): 0% at 0 min, 0% at 1 min, 100% at 6 min, 100% at 6.5 min, 0% at 7.5 min. HILIC-phase chromatography conditions were as follows. Solvent A: 10% acetonitrile, 10 mM ammonium formate; solvent B: 90% acetonitrile, 10 mM ammonium formate; flow rate 0.4 mL/min; maximum pressure 600 bar; timetable (%A): 0% at 0 min, 0% at 1 min, 75% at 6 min, 75% at 6.5 min, 0% at 7.5 min.

Characterization of kinetic resolution by enzymatic reduction. Enzymatic aldol reactions of 3-fluoro-2-oxobutanoate (0.3 mL total volume) were prepared by the method above, using formaldehyde, propionaldehyde, glycolaldehyde, or 2-pyridinecarboxaldehyde as acceptors. A sample with no aldolase and no aldehyde served as control. After 48 h of incubation, to the samples was added 100 mM EDTA pH 9.0 (100 μ L) to inactivate the aldolase, followed 10 min later by NADH (200 μ L of a 100 mM stock in 10 mM NaOH) and LDH (100 μ L of 2 mg/mL stock in 50 mM sodium phosphate pH 7.5). The enzymatic reduction reactions were incubated at r.t. for 16 h followed by addition of D₂O (100 μ L) for ¹⁹F NMR analysis.

DFT calculation of carbon acidity of 3-fluoro-2-oxoacids. The carbon acidity of fluorinated donors was computed from the proton exchange reaction (pyruvate dianion) + (donor monoanion) \rightarrow (pyruvate monoanion) + (donor dianion). The Hartree energy of each species was calculated with the Q-Chem 5.2 package [29] using the ω B97M-V range-separated hybrid density functional [30] and the SMD solvent model [31]. After optimization of geometries with the def2-SVPD basis set, single point energies were recomputed with the larger def2-TZVPD basis set [32]. Only the donor 3-fluoro-3-phenylpyruvate adopted the *trans*-fluoroenolate geometry, while all other donors adopted the *cis*-fluoroenolate geometry. The enthalpy and entropy of nuclear contributions to the free energy were estimated using a modified quasi rigid rotor harmonic oscillator model with ω B97M-V/def2-SVPD geometries and frequencies [33]. The change in each free energy component (ΔE , ΔH , ΔS) was obtained as (sum of product energies) - (sum of reactant energies), and the overall change in Gibbs free energy was obtained as $\Delta G = \Delta E + \Delta H - T\Delta S$, with $T = 298.15$ K. Using the ideal gas constant and temperature of $T = 298.15$ K, the values of ΔG were converted to a pK_a difference relative to the carbon acidity of pyruvate, which is assigned the value of $pK_a = 16.58$ according to [34].

General procedure for synthesis of 2-fluoroesters (3a-c) from formaldehyde. In a plastic tube, the starting 3-fluoro-2-oxoacid sodium salt (0.6 mmol) was dissolved in the specified enzymatic reaction buffer (7.4 mL). Formaldehyde solution (2 M in H₂O, 0.6 mL, 1.2 mmol) was added, followed by the specified preparation of HpcH aldolase or mutant thereof (4 mL in HEPES storage buffer containing 1 mM MgCl₂, pH 7.5). The reaction mixture (12 mL) was incubated without stirring at r.t. for the specified duration, then was quenched with H₂O₂ (30% w/w, 1.2 mL) and stirred for 1 h for decarboxylation. Excess H₂O₂ was decomposed with catalase (1 mg) and stirring for 30 min. The mixture was concentrated to near dryness by co-evaporation with CH₃CN (50 mL), then the residue was dissolved in methanolic HCl (1 M, 30 mL) and stirred at r.t. overnight for esterification. The solvent was evaporated and replaced with saturated aqueous NaHCO₃ (20 mL), from which the product was extracted with EtOAc (6 \times 20 mL). The organic layer was washed with brine (20 mL), dried over Na₂SO₄, and concentrated. Products were purified by silica gel chromatography with the specified solvents. Racemic standards were prepared by replacing the volume of enzymatic reaction buffer and aldolase solution with the specified racemic reaction buffer containing CoCl₂ (11.4 mL). The mixture was incubated without stirring at r.t. for the specified duration. All other aspects of the reaction and product isolation remain identical.

Methyl (S)- or (RS)-2-fluoro-3-hydroxypropanoate (3a or rac-3a). Sodium fluoropyruvate (88 mg, 0.6 mmol) was treated by the general procedure above, with the following details. Enzymatic reaction buffer: 20 mM CHES-NaOH, 1 mM MgCl₂, pH 9.3. Enzyme preparation: wt-HpcH, 4.65 mg/mL (150 μM), 0.1 mol% catalyst load. Enzymatic reaction time: 4 h. Note the final mixture has pH 8.5. A racemic standard (*RS*) was prepared with the following details. Racemic reaction buffer: 20 mM HEPPS-NaOH, 1 mM CoCl₂, pH 8.5. Racemic reaction time: 48 h. Chromatography with 1:1 hexanes/EtOAc provided the stereopure product (35 mg, 48%) or the racemic product (16 mg, 22%) as a colorless oil. ¹H NMR (600 MHz, CDCl₃) δ 5.03 (ddd, *J* = 48.2, 4.5, 2.9 Hz, 1H), 4.11-3.98 (m, 2H), 3.86 (s, 3H), 2.11 (s, 1H). ¹³C NMR (151 MHz, CDCl₃) δ 168.2 (d, *J* = 24.2 Hz), 89.5 (d, *J* = 185.4 Hz), 62.8 (d, *J* = 21.3 Hz), 52.6. ¹⁹F NMR (565 MHz, CDCl₃) δ -199.8 (ddd, *J* = 48.5, 28.5, 20.8 Hz).

Methyl (S)- or (RS)-2-fluoro-3-hydroxy-2-methylpropanoate (3b or rac-3b). Sodium 3-fluoro-2-oxobutanoate (96 mg, 0.6 mmol) was treated by the general procedure above, with the following details. Enzymatic reaction buffer: 20 mM HEPES-NaOH, 1 mM MgCl₂, pH 7.5. Enzyme preparation: wt-HpcH, 4.65 mg/mL (150 μM), 0.1 mol% catalyst load. Enzymatic reaction time: 48 h. A racemic standard (*RS*) was prepared with the following details. Racemic reaction buffer: 20 mM Na₂HPO₄-NaOH, 1 mM CoCl₂, pH 12.5. Racemic reaction time: 48 h. Chromatography with 2:1 hexanes/EtOAc provided the stereopure product (15 mg, 37%, theoretical yield 0.3 mmol) or the racemic product (10 mg, 13%) as a colorless oil. ¹H NMR (600 MHz, CDCl₃) δ 3.96-3.81 (m, 2H), 3.85 (s, 3H), 2.00 (s, 1H), 1.57 (d, *J* = 21.4 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 171.1 (d, *J* = 25.5 Hz), 95.5 (d, *J* = 184.7 Hz), 67.0 (d, *J* = 23.7 Hz), 52.8, 19.8 (d, *J* = 23.5 Hz). ¹⁹F NMR (565 MHz, CDCl₃) δ -164.8 (m).

Methyl (S)- or (RS)-2-fluoro-2-(hydroxymethyl)butanoate (3c or rac-3c). Sodium 3-fluoro-2-oxopentanoate (104 mg, 0.6 mmol) was treated by the general procedure above, with the following details. Enzymatic reaction buffer: 20 mM HEPES-NaOH, 1 mM MgCl₂, pH 7.5. Enzyme preparation: HpcH F170V, 13.95 mg/mL (450 μM), 0.3 mol% catalyst load. Enzymatic reaction time: 1 week. A racemic standard (*RS*) was prepared with the following details. Racemic reaction buffer: 20 mM Na₂HPO₄-NaOH, 1 mM CoCl₂, pH 12.5. Racemic reaction time: 48 h. Chromatography with 2:1 hexanes/EtOAc provided the stereopure product (17 mg, 38%, theoretical yield 0.3 mmol) or the racemic product (22 mg, 24%) as a colorless oil. ¹H NMR (600 MHz, CDCl₃) δ 3.96-3.83 (m, 2H), 3.85 (s, 3H), 2.02 (s, 1H), 1.97-1.82 (m, 2H), 0.98 (t, *J* = 7.5 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.6 (d, *J* = 26.5 Hz), 98.8 (d, *J* = 187.0 Hz), 66.2 (d, *J* = 23.6 Hz), 52.6, 26.6 (d, *J* = 22.2 Hz), 7.2 (d, *J* = 4.4 Hz). ¹⁹F NMR (565 MHz, CDCl₃) δ -175.2 (tt, *J* = 28.6, 14.8 Hz).

Ethyl (2S,3RS)-2-fluoro-3-hydroxy-3-phenylpropanoate (16a). In a plastic tube, sodium fluoropyruvate (88 mg, 0.6 mmol) was dissolved in 20 mM CHES-NaOH, 1 mM MgCl₂, pH 9.3 (7.4 mL). A solution of benzaldehyde (2 M in DMSO, 0.6 mL, 1.2 mmol) was added, followed by the enzyme preparation: MBP-HpcH, 10.89 mg/mL (150 μM), 0.1 mol% catalyst load, in 4 mL of HEPES storage buffer containing 1 mM MgCl₂, pH 7.5. Note the final mixture has pH 8.5. The reaction mixture (12 mL) was stirred very slowly (~60 rpm) at r.t. for 16 h, then was quenched with H₂O₂ (30% w/w, 1.2 mL) and stirred for 1 h for decarboxylation. Excess H₂O₂ was decomposed with catalase (1 mg) and stirring for 30 min. The mixture was diluted with water (40 mL), acidified to pH 0 with 6 M HCl, and then extracted with methyl *tert*-butyl ether (3 × 30 mL). The organic layers were washed with brine (30 mL) and concentrated, then the residue was dissolved in ethanol (30 mL) and thionyl chloride (1 mL) added slowly. The esterification reaction

was stirred at r.t. overnight, then the solvent was evaporated and replaced with saturated aqueous NaHCO₃ (30 mL), from which the product was extracted with EtOAc (3 × 30 mL). The organic layer was washed with brine (30 mL), dried over Na₂SO₄, and concentrated. Purification by silica gel chromatography (9:1 then 4:1 hexanes/EtOAc) afforded the product (61:39 syn/anti mixture of diastereomers) as a colorless oil (55 mg, 43%). ¹H NMR (600 MHz, CDCl₃) δ 7.44-7.34 (m, 5H), 5.16 (dd, *J* = 22.1, 3.7 Hz, 1H syn), 5.15 (dd, *J* = 12.1, 4.3 Hz, 1H anti), 5.07 (dd, *J* = 48.2, 5.0 Hz, 1H anti), 5.04 (dd, *J* = 48.1, 3.8 Hz, 1H syn), 4.28-4.20 (m, 2H), 1.25 (t, *J* = 7.2 Hz, 3H syn), 1.22 (t, *J* = 7.2 Hz, 3H anti). ¹³C NMR (151 MHz, CDCl₃) Syn δ 167.68 (d, *J* = 24.3 Hz), 137.9 (d, *J* = 2.8 Hz, syn), 128.62, 128.5, 126.5, 91.5 (d, *J* = 191.9 Hz), 74.0 (d, *J* = 20.0 Hz), 61.9, 13.98; Anti δ 167.69 (d, *J* = 23.2 Hz), 137.6 (d, *J* = 2.8 Hz) 128.65, 128.62, 126.8, 90.9 (d, *J* = 191.1 Hz), 73.7 (d, *J* = 22.0 Hz), 61.8, 13.96. ¹⁹F NMR (565 MHz, CDCl₃) Syn δ -202.7 (dd, *J* = 48.5, 22.3 Hz); Anti δ -197.7 (dd, *J* = 48.2, 16.2 Hz).

Ethyl (2*RS*,3*RS*)-2-fluoro-3-hydroxy-3-phenylpropanoate (rac-16a). A racemic standard of the previous product was prepared by Reformatsky reaction of benzaldehyde (531 mg, 508 μL, 5 mmol) and ethyl bromofluoroacetate (1.29 g, 827 μL, 7 mmol). To a suspension of zinc dust (588 mg, 9 mmol) in THF (10 mL) under nitrogen gas was added DIBAL-H (1 M in toluene, 0.1 mL, 2 mol%) and 5% of the alkyl bromide. Zinc was activated by vigorous stirring for 15 min, then the rest of the alkyl bromide and the benzaldehyde were added successively. The reaction was stirred at r.t. for 16 h, then diluted with EtOAc (30 mL) and filtered through Celite. After thorough shaking with 1 M NH₄Cl (30 mL), the layers were separated and the aqueous layer was extracted with EtOAc (3 × 30 mL). The organic layers were washed with brine (30 mL), dried over Na₂SO₄, and concentrated. Purification by silica gel chromatography (9:1 then 4:1 hexanes/EtOAc) afforded the product (55:45 syn/anti mixture of diastereomers) as a colorless oil (603 mg, 57%). Spectral data was identical to the previous section aside from the diastereomeric ratio.

Ethyl (2*S*,3*RS*)-2-fluoro-3-hydroxy-2-methyl-3-(2-pyridyl)propanoate (17b). In a plastic tube, sodium 3-fluoro-2-oxobutanoate (96 mg, 0.6 mmol) was dissolved in 20 mM HEPES-NaOH, 1 mM MgCl₂, pH 7.5 (7.4 mL). A solution of 2-pyridinecarboxaldehyde (2 M in H₂O, 0.6 mL, 1.2 mmol) was added, followed by the enzyme preparation: wt-HpcH, 4.65 mg/mL (150 μM), 0.1 mol% catalyst load, in 4 mL of HEPES storage buffer containing 1 mM MgCl₂, pH 7.5. The reaction mixture (12 mL) was incubated without stirring at r.t. for 48 h, then was quenched with H₂O₂ (30% w/w, 1.2 mL) and stirred for 1 h for decarboxylation. Excess H₂O₂ was decomposed with catalase (1 mg) and stirring for 30 min. The mixture was concentrated to near dryness by co-evaporation with CH₃CN (50 mL), then the residue was dissolved in ethanol (30 mL) and thionyl chloride (1 mL) added slowly. The esterification reaction was stirred at r.t. overnight, then the solvent was evaporated and replaced with saturated aqueous NaHCO₃ (30 mL), from which the product was extracted with EtOAc (3 × 30 mL). The organic layer was washed with brine (30 mL), dried over Na₂SO₄, and concentrated. The residue was passed through a short column of silica gel (eluted with 1:1 hexanes/EtOAc), then concentrated and suspended in 1 M NaHSO₃ (10 mL) to remove residual 2-pyridinecarboxaldehyde. Extraction with CH₂Cl₂ (3 × 20 mL) followed by washing with brine (20 mL), drying over Na₂SO₄ and concentration provided a yellow oil, containing the product (21:79 syn/anti mixture of diastereomers) and the diethyl acetal of 2-pyridinecarboxaldehyde. Due to difficulty in thoroughly removing the latter impurity, relative integrals from ¹H NMR were used to subtract its weight, giving a corrected yield of product (31 mg, 45%, theoretical yield 0.3 mmol). ¹H NMR (600 MHz, CDCl₃) δ 8.63-8.61 (m, 1H anti), 8.61-8.60 (m, 1H syn), 7.81-7.78 (m, 1H syn), 7.78-7.75 (m, 1H anti), 7.47-7.44 (m, 1H syn), 7.36-7.32 (m, 1H), 5.14 (d, *J* = 15.3 Hz, 1H anti), 5.01 (d, *J* = 15.5 Hz, 1H syn), 4.33-4.27 (m, 2H anti), 4.22-

4.18 (m, 2H syn), 1.67 (d, $J = 22.0$ Hz, 3H syn), 1.53 (d, $J = 22.4$ Hz, 3H anti), 1.32 (t, $J = 7.1$ Hz, 3H anti), 1.23 (t, $J = 7.1$ Hz, 3H syn). ^{13}C NMR (151 MHz, CDCl_3) Syn δ 169.9 (d, $J = 25.3$ Hz), 155.6, 147.60, 137.2, 123.65, 123.2 (d, $J = 3.0$ Hz), 96.2 (d, $J = 190.2$ Hz), 75.5 (d, $J = 24.8$ Hz), 61.7, 20.7 (d, $J = 23.0$ Hz), 14.0; Anti δ 170.3 (d, $J = 23.4$ Hz), 155.4, 147.57, 137.3, 123.59, 122.5 (d, $J = 2.8$ Hz), 96.3 (d, $J = 189.7$ Hz), 73.9 (d, $J = 25.2$ Hz), 61.9, 18.5 (d, $J = 23.2$ Hz), 14.1. ^{19}F NMR (565 MHz, CDCl_3) Syn δ -165.11 (quint, $J = 19.9, 19.2$ Hz); Anti δ -163.75 (quint, $J = 21.6$ Hz).

Ethyl (2RS,3RS)-2-fluoro-3-hydroxy-3-(2-pyridyl)propanoate (rac-17b). A racemic standard of the previous product was prepared by LDA-promoted aldol reaction of 2-pyridinecarboxaldehyde (643 mg, 570 μL , 6 mmol) and ethyl 2-fluoropropionate (601 mg, 5 mmol). To a solution of diisopropylamine (607 mg, 847 μL , 6 mmol) in THF (10 mL) under nitrogen gas at -78 $^\circ\text{C}$ was added n-BuLi (2.5 M in hexanes, 2.2 mL, 5.5 mmol). After stirring 1 min the ethyl 2-fluoropropionate was added and the solution stirred for 15 min. Then, the 2-pyridinecarboxaldehyde was added and stirring was continued at -78 $^\circ\text{C}$ for 30 min. The reaction was quenched at -78 $^\circ\text{C}$ with 1 M NH_4Cl (5 mL), then diluted with sat. NaHCO_3 (40 mL) and warmed to room temperature. The product was extracted with EtOAc (3×30 mL), washed with brine (30 mL), dried over Na_2SO_4 and concentrated. The residue was passed through a short column of silica gel (eluted with 1:1 hexanes/EtOAc), then concentrated and suspended in 1 M NaHSO_3 (10 mL) to remove residual 2-pyridinecarboxaldehyde. Extraction with CH_2Cl_2 (3×20 mL) followed by washing with brine (20 mL), drying over Na_2SO_4 and concentration provided the product (35:65 syn/anti mixture of diastereomers) as a yellow oil (277 mg, 24%). Spectral data was identical to the previous section aside from the diastereomeric ratio.

Chiral derivatization with Mosher's acid. To a solution of (+)-MTPA (0.1 mmol, 23.4 mg) in dry hexanes (2 mL) was added catalytic amount of DMF (~ 1 μL) followed by oxalyl chloride (2 M in CH_2Cl_2 , 0.5 mmol, 0.25 mL). After stirring at r.t. for 1 h, the solvent was evaporated. The acid chloride was dissolved in dry CH_2Cl_2 (2 mL), to which was added the alcohol to be derivatized (0.08 mmol), triethylamine (0.3 mmol, 42 μL), and catalytic amount of DMAP (~ 1 mg). The mixture was stirred overnight, then diluted with ether (20 mL) and washed with water (2×10 mL) and brine (10 mL). The organic layer was dried over Na_2SO_4 and concentrated. The residue was dissolved in CDCl_3 (0.7 mL) for analysis by ^{19}F NMR spectroscopy. The derivatives of alcohol **3c** had poor ^{19}F NMR separation, and were additionally analyzed by TLC. The samples were spotted 2 cm from the bottom of 10 cm silica plates. The plates were eluted for 7 min or 15 min in 9:1 hexanes/EtOAc, and the Mosher's esters visualized by UV lamp.

Ammonium (2S,3S,4R)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate (12b). In a plastic tube, sodium 3-fluoro-2-oxopentanoate (26 mg, 0.15 mmol) was dissolved in 20 mM HEPES-NaOH, 1 mM MgCl_2 , pH 7.5 (1.85 mL). A solution of D-glyceraldehyde (2 M in H_2O , 0.15 mL, 0.3 mmol) was added, followed by the enzyme preparation: wt-HpcH, 4.65 mg/mL (150 μM), 0.1 mol% catalyst load, in 1 mL of HEPES storage buffer containing 1 mM MgCl_2 , pH 7.5. The reaction mixture (4 mL) was incubated without stirring at r.t. for 48 h, then was quenched with H_2O_2 (30% w/w, 0.4 mL) and stirred for 1 h for decarboxylation. Excess H_2O_2 was decomposed with catalase (1 mg) and stirring for 30 min. After removing the precipitate by centrifugation ($13,000 \times g$, 5 min), the sample was lyophilized in a vacuum concentrator to reduce the volume to 2 mL and then filtered through a 0.22 μm syringe filter. The product was purified semi-preparative HPLC on an Agilent 1200 Series HPLC System equipped with a SeQuant ZIC-HILIC column (5 μm , 200 Å , 21.2×150 mm). Chromatography conditions were as follows. Solvent A: 10%

acetonitrile, 10 mM ammonium formate; solvent B: 90% acetonitrile, 10 mM ammonium formate; flow rate 5 mL/min; maximum pressure 400 bar; timetable (%A): 0% at 0 min, 0% at 10 min, 75% at 70 min, 100% at 80 min, 100% at 90 min, 0% at 100 min, 0% at 120 min; fraction collection: 384 fractions in deep 2 mL 96-well plates. Without dilution, the fractions were directly screened by electrospray ionization (ESI) LC-MS. Agilent 1290 Infinity LC system was equipped with an Ascentis Express HILIC column, and chromatography conditions were identical to those described previously in the section “*High resolution LC-MS analysis*”. Agilent 6460 Triple-Quad MS was operated in negative mode with single-ion monitoring of the 181.0518 *m/z* ion. Fractions containing the product were combined and concentrated by rotary evaporation to ~10 mL, then lyophilized in a vacuum concentrator to reveal white solids, consisting of ammonium formate and the product (5:95 syn/anti mixture of diastereomers). The amount of product was estimated by addition of DMSO (37.5 μ mol) as internal standard to the NMR sample (2.4 mg, 16%, theoretical yield 75 μ mol). Spectral data is given for the *anti* diastereomer only. ^1H NMR (600 MHz, D_2O) δ 3.86-3.80 (m, 1H), 3.77 (dd, $J = 10.9, 5.9$ Hz, 1H), 3.67-3.55 (m, 2H), 1.35 (d, $J = 22.8$ Hz, 3H). ^{13}C NMR (151 MHz, D_2O) δ 94.0 (d, $J = 178.2$ Hz), 71.2 (d, $J = 19.2$ Hz), 64.6 (d, $J = 2.2$ Hz), 60.3, 14.8 (d, $J = 24.2$ Hz). ^{19}F NMR (565 MHz, D_2O) δ -172.1 (quint, $J = 23.5$ Hz). The carboxylate peak in ^{13}C NMR was not assigned due to presence of a large formate peak.

Ammonium (2S,3RS,4S)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate (13b). The previous procedure was repeated using L-glyceraldehyde (1 M in H_2O , 0.15 mL, 0.15 mmol) as acceptor. All other aspects of the reaction and product isolation remained identical, providing a sample of white solids, consisting of ammonium formate and the product (33:67 syn/anti mixture of diastereomers). The amount of product was estimated by addition of DMSO (37.5 μ mol) as internal standard to the NMR sample (1.9 mg, 13%, theoretical yield 75 μ mol). ^1H NMR (600 MHz, D_2O) δ 4.22-4.09 (m, 1H), 3.99-3.94 (m, 1H), 3.84-3.79 (m, 1H), 3.70-3.61 (m, 1H), 1.45 (d, $J = 23.2$ Hz, 3H anti), 1.32 (d, $J = 23.5$ Hz, 3H syn). ^{13}C NMR (151 MHz, D_2O) Syn δ 104.0 (d, $J = 200.9$ Hz), 80.97, 72.7 (d, $J = 17.2$ Hz), 60.8, 15.8 (d, $J = 25.5$ Hz); Anti δ 101.5 (d, $J = 179.0$ Hz), 80.95, 72.9 (d, $J = 17.4$ Hz), 61.5, 15.4 (d, $J = 24.4$ Hz). ^{19}F NMR (565 MHz, D_2O) Syn δ -173.5 (quint, $J = 23.3$ Hz); Anti δ -172.3 (quint, $J = 23.8$ Hz). The carboxylate peak in ^{13}C NMR was not assigned due to presence of a large formate peak.

3.3. Results and discussion

Among the five homologs of HpcH aldolase that were discussed in the previous chapter, we began by selecting one aldolase that would be a good candidate to explore the use of substituted pyruvate donors, which would generate tertiary fluoride stereocenters. Although all candidates had high fluorine stereoselectivity with fluoropyruvate substrate as well as similarly broad acceptor scopes, we reasoned that HpcH from *Escherichia coli* C strain (EcHpcH) would serve as the best scaffold for protein engineering efforts, due to its high solubility and protein expression yield (up to 80 mg/L culture). All work on biocatalysis with HpcH family aldolases up to this point had been performed with RhmA from *E. coli* (also known as YfaU), whose poor solubility always necessitated the use of a maltose-binding protein (MBP) fusion of the aldolase [23,24,35,36]. On the other hand, EcHpcH had not been used for biocatalysis, but numerous studies on its biochemical characterization showed that it is structurally similar and mechanistically identical to RhmA [37-40].

We next decided upon a panel of β -fluoro- α -ketoacids bearing different C-3 substitutions to investigate as new fluoro-donors for EcHpcH. Although sodium fluoropyruvate (FP, **1**, R = H),

the donor that would give rise to secondary fluoride stereocenters, is commercially available, no other β -fluoro- α -ketoacids are commercially available, and thus the donors that would provide tertiary fluoride stereocenters needed to be synthesized (Figure 3.2). In course, β -fluoro- α -ketobutyrate (FKB, **2a**, R = Me), β -fluoro- α -ketovalerate (FKV, **2b**, R = Et), β -fluoro- α -keto- γ -methylvalerate (FKMV, **2c**, R = *i*-Pr), and β -fluoro- α -keto- γ -phenylbutyrate (FKPB, **2d**, R = Bn) were prepared as racemic compounds from the corresponding α -ketoesters in a three step sequence adapted from the literature [41]. In cases where only the α -ketoacid was available as starting material, it was esterified with methanol, 2,2-dimethoxypropane, and catalytic TMS-Cl. Each α -ketoester was converted to a silyl enol ether with TMS-Cl and Et₃N, then treated with Selectfluor reagent, providing a β -fluoro- α -ketoester that was saponified with LiOH. The products were purified as free acids by acid-base extraction prior to neutralization to sodium salts for enzymatic use. The use of acetonitrile as reaction solvent for all three steps from the α -ketoesters was a critical optimization that allowed multi-gram scale synthesis of these fluoro-donors. When silylation was performed in acetonitrile instead of THF, the Et₃NHCl by-product produced a less viscous slurry. Dilution with just 1.5 volumes of water completely solubilized Et₃NHCl, and the water/acetonitrile layer was efficiently extracted with hexanes. Selectfluor was more soluble in acetonitrile than THF, and the fluorinated products could be extracted with diethyl ether from the solid by-product of Selectfluor after evaporation. Saponification of β -fluoro- α -ketoester in acetonitrile was found to prevent significant fluoride elimination that occurred in aqueous solvents. Due to the relatively high water solubility of FKB and FKV free acids, it was important to minimize the volume of water used during their purification by acid-base extraction. The final desired substrate of fluoro(phenyl)pyruvate (FPP, **2e**, R = Ph) could not be accessed with the Selectfluor route due to

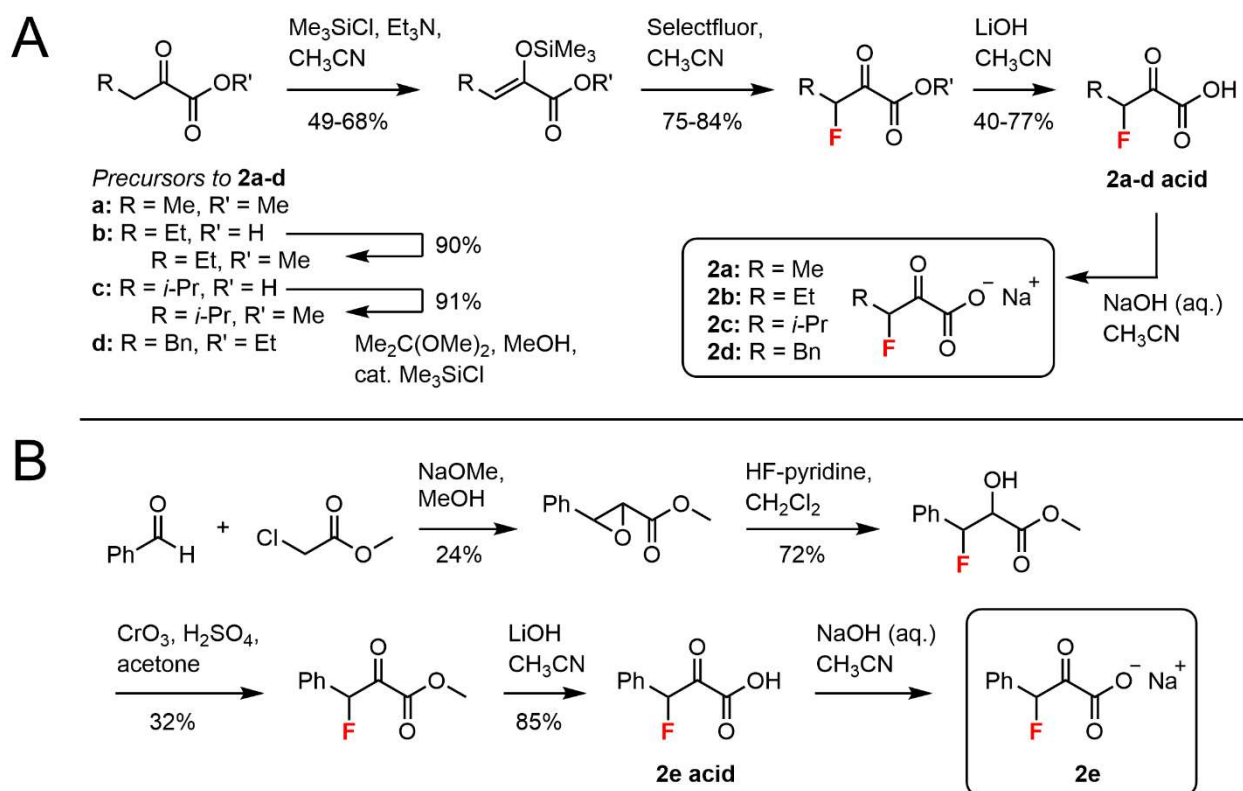


Figure 3.2. A) Synthesis of C3-alkyl substituted fluoropyruvate analogs by silylation, fluorination, and saponification. B) Synthesis of C3-aryl substituted fluoropyruvate analog by Darzens condensation, epoxide opening with fluoride, Jones oxidation, and saponification.

aromatic conjugation. Therefore, the methyl ester of FPP was prepared by a known route of Darzens condensation of benzaldehyde with methyl chloroacetate, epoxide opening with HF-pyridine, and Jones oxidation [42]. This product was also saponified with LiOH and converted to a sodium salt in the same manner as the other four synthesized fluoro-donors.

In order to test the competence of these new fluoro-donors in enzymatic aldol reactions, we decided to initially use formaldehyde as a model acceptor to simplify the products to those with one nascent stereocenter. As the simplest possible acceptor, formaldehyde would essentially alkylate the fluoro-donor with a hydroxymethyl group. After inspecting the active site of EcHpcH for bulky non-catalytic residues, a collection of EcHpcH mutants with size reduction of varying degrees at the Trp-19, Phe-170, and Leu-212 residues were prepared in hopes of steric alleviation towards bulky fluoro-donors (Figure 3.3). All enzyme variants were expressed recombinantly in *E. coli* with N-terminal His₁₀ tags, purified by metal affinity chromatography, and confirmed by SDS-PAGE analysis (Figure 3.4). All enzyme variants were confirmed to possess native activity, by incubation of pyruvate (50 mM) and glycolaldehyde (0.1 M) with 0.1 mol% enzyme for 1 h,

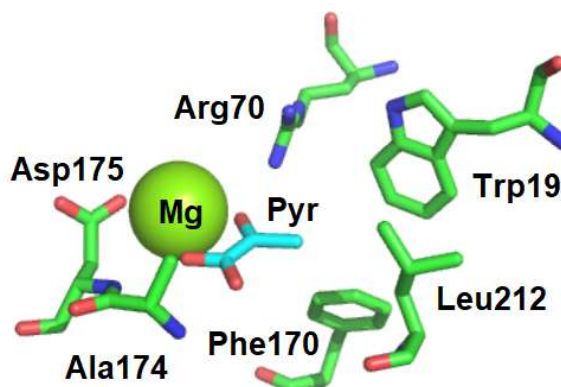


Figure 3.3. Active site organization of EcHpcH, showing pyruvate chelated to a divalent metal, and the bulky residues Trp-19, Phe-170, and Leu-212 that were targeted for mutations. Crystal structure is from RCSB PDB 4B5U [Ref. 40].

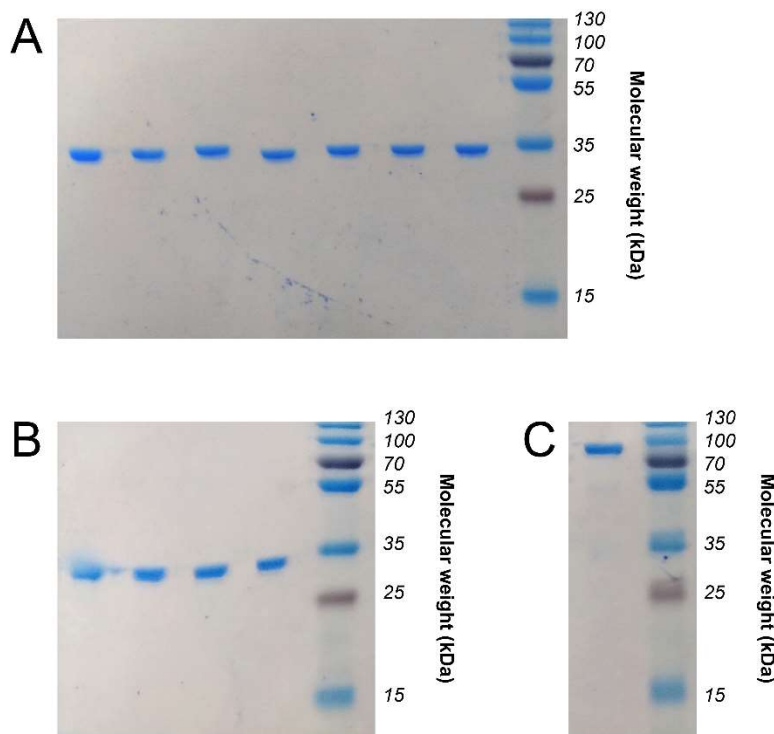
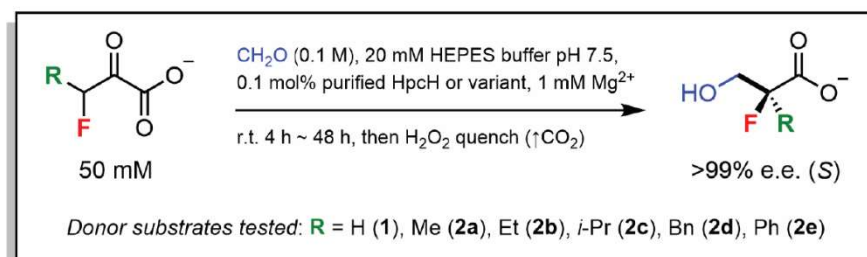


Figure 3.4. SDS-PAGE analysis of EcHpcH variants. A) From left to right: WT, W19A, F170A, L212A, F170A/L212A, F170V, F170V/L212A. B) From left to right: F170L, F170I, L212I, L212V. C) MBP-tagged EcHpcH.

decarboxylation with H₂O₂, and observation of 90% or greater conversion by ¹H NMR. An initial screen for fluoro-aldol addition was conducted by incubating fluoro-donors (50 mM) and formaldehyde (0.1 M) with 0.1 mol% enzyme, followed by H₂O₂ decarboxylation. The percent conversion was determined by integration of ¹⁹F NMR peaks corresponding to α-fluoroacids arising from reactants and products (*Figure 3.5*). We first observed that FP and FKB reacted well with formaldehyde. The best conversions of greater than 50% were achieved with wild-type enzyme (wt-HpcH), suggesting that the mutation of large residues to smaller ones decreased the intrinsic enzyme activity and were detrimental with respect to sterically non-challenging donors. However, wt-HpcH was ineffective on FKV, producing less than 2% of the expected product. Attempts to turn over FKV with mutants at Phe-170 or Leu-212 gave increased conversion at 6-12%. This result suggests that steric relief only becomes important upon introducing a C-3 substituent the size of an ethyl group or larger. The reaction time necessary to collect reliable screening data accordingly increased depending on the difficulty of the fluoro-donor: 4 h was used for FP, 16 h for FKB, and 48 h for FKV.

The bulkiest fluoro-donors FKMV and FKPB did not react with formaldehyde after 48 h using any of the EcHpcH variants. In an effort to produce any trace amount of aldol products, the enzyme loading of mutants F170L, F170V, F170A, and L212A was tripled to 0.03 mol%, and the acceptor



Variant of HpcH aldolase	Donor substrate	1	2a	2b	2c	2d	% Conversion by NMR
	(time)	(4 h)	(16 h)	(48 h)	(48 h)	(48 h)	
	No enzyme	tr.	n.d.	n.d.	n.d.	n.d.	
	Wild-type	69	50	2	n.d.	n.d.	
	W19A	9	tr.	tr.	n.d.	n.d.	
	F170A	5	2	6	n.d.	n.d.	
	L212A	7	16	10	n.d.	n.d.	
	F170A/L212A	2	tr.	2	n.d.	n.d.	
	F170V	6	tr.	10	n.d.	n.d.	
	F170V/L212A	4	tr.	2	n.d.	n.d.	
	F170L	30	24	13	n.d.	n.d.	
	F170I	7	2	13	n.d.	n.d.	
	L212I	57	tr.	tr.	n.d.	n.d.	
	L212V	33	n.d.	tr.	n.d.	n.d.	
	Optimized condition	98 ^a	82 ^b	40 ^c	-	-	

Figure 3.5. Screening fluoro-donors for C-F stereocenter formation. No product detected (n.d.) or trace product of <1% conversion (tr.) where indicated. Due to kinetic resolution of C-3 substituted fluoro-donors, 100% conversion equals 50% donor consumption. Optimized conditions had the following modifications. ^aUsed wt-HpcH, pH 8.5 (2:1 CHES/HEPES). ^bUsed wt-HpcH, 48 h. ^cUsed 0.3 mol% HpcH F170V, 1 week.

was switched from formaldehyde to glycolaldehyde, which is known to be an extremely good acceptor. Unfortunately, no reaction of FKMV or FKPB was observed under these forcing conditions. It is likely that the steric abnormality of a C-3 isopropyl or benzyl group cannot be compensated by simple point mutations, as the effect of creating space in the active site is counterbalanced by the intrinsically lower activity of mutants. In the future, a more extensively engineered aldolase that includes mutations outside the active site may be able to utilize such large fluoro-donors. Lastly, the fluoro-donor FPP which bears a C-3 aryl group was observed to react non-enzymatically with formaldehyde, with approximately 30% conversion to a product regardless of the presence or absence of aldolase. As this reaction would be non-stereoselective, FPP was not further investigated.

Additionally, a kinetic resolution phenomenon was observed for FKB and FKV, the two C-3 substituted fluoro-donors that were competent for aldol reaction (*Figure 3.6*). We saw that the total consumption of FKB and FKV never exceeded 50%, and proved that racemic FKB undergoes kinetic resolution by analysis of samples subjected to stereoselective reduction at C-2 by *L*-lactate dehydrogenase (LDH). Using this approach, the enantiomers of FKB were converted to

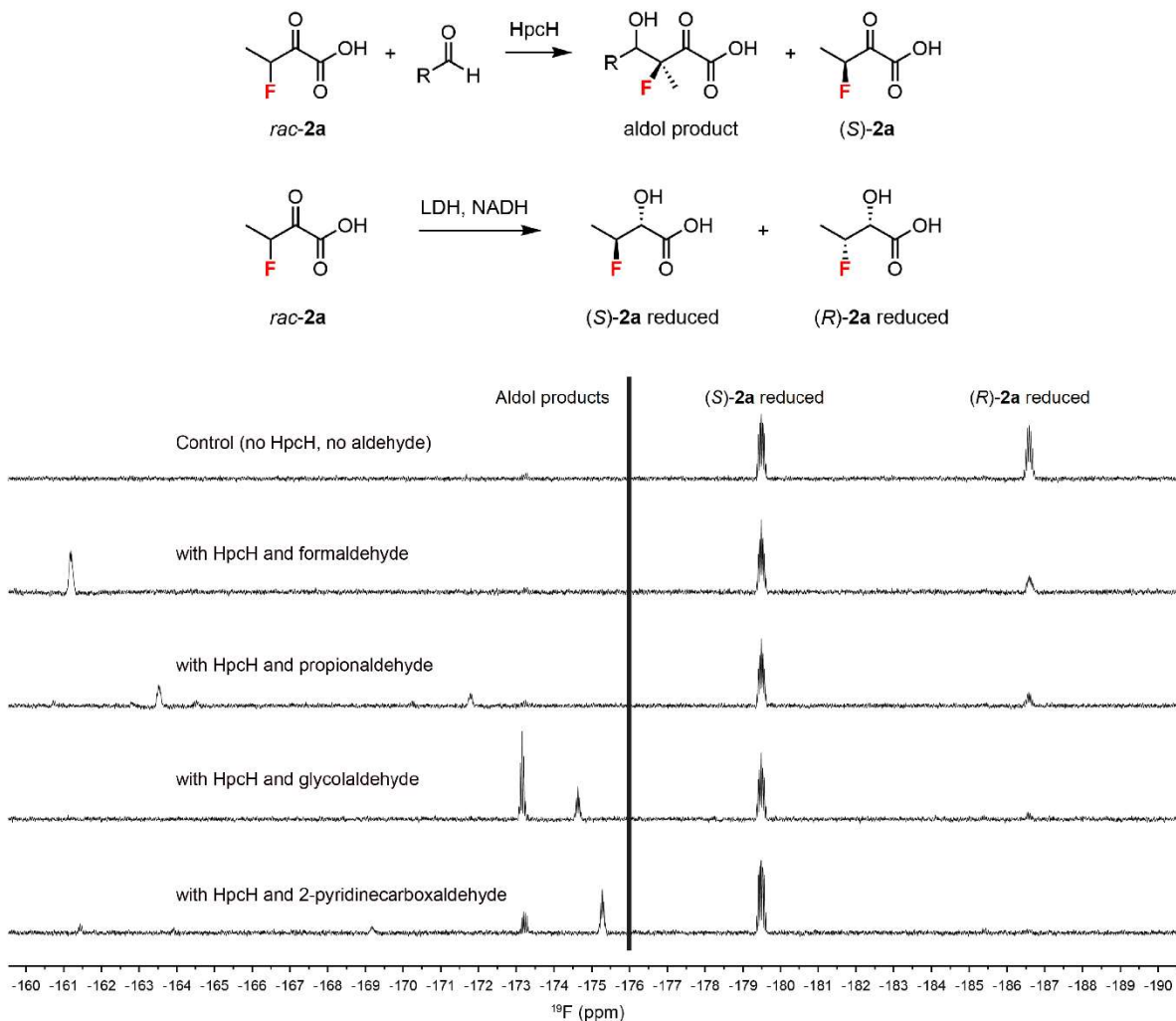


Figure 3.6. Kinetic resolution of a racemic fluoro-donor as shown by stereoselective reduction of the fluoro-donor enantiomers with lactate dehydrogenase, and ^{19}F NMR analysis of the resulting products.

diastereomers of 3-fluoro-2-hydroxybutanoate that could be distinguished by NMR, indicating that only (*R*)-FKB is depleted by the aldolase and (*S*)-FKB remains after reaction completion. Although we have not directly proved that kinetic resolution also occurs for FKV because it not a substrate of LDH, similar stalling of reactions at 50% donor consumption indicates that FKV is also kinetically resolved. As such, all conversion values for reactions of FKB and FKV (in this chapter text, in figures, and in *Appendix 2*) are expressed such that 100% conversion is equal to 50% consumption of the racemic donor. This method of kinetic resolution could be potentially useful for the preparation of enantiopure β -fluoro- α -ketoacids.

The origin of the exquisite stereoselectivity of HpcH in generating (*S*)-fluoro stereocenters is thought to arise from two main factors: (1) the electronic requirement for *cis*-geometry of the metal-bound fluoroenolate intermediate, and (2) the enzyme-determined *Si*-facial consistency (relative to C-3 of the fluoroenolate) for proton abstraction and attack onto an aldehyde [43]. This provides the framework to rationalize the kinetic resolution phenomenon and the beneficial nature of Phe-170 and Leu-212 mutations observed in the mutational study. For the racemic fluoro-donors, only one enantiomer has an acidic proton orientable such that H-abstraction leads to a *cis*-fluoroenolate. Thus, the alkyl group is forced into the *trans* position of the fluoroenolate, becoming directed towards the space occupied by Phe-170 and Leu-212 (*Figure 3.7*). During the course of this study, it was reported that mutation of Trp-19 to a small residue (W19V or W19A) was the key change that unlocked activity with bulky non-fluorinated α -ketoacids [23]. The resulting product stereochemistry with these donors indicated that the C-3 bulk had occupied the *cis* position of the enolate, which is directed towards Trp-19. In contrast, the W19A mutation was totally ineffective with our fluoro-donors, corroborating our theory that fluorine defeats an alkyl group in the competition of both groups for the *cis* position.

To rule out any electronic as opposed to steric explanations of reactivity differences between our tested fluoro-donors, we note that in the hypothetical absence of steric factors, the inductive effect of electron-releasing alkyl substituents at C-3 would increase the fluoroenolate nucleophilicity, and thus accelerate the rate-determining step of C-C bond formation [44]. DFT

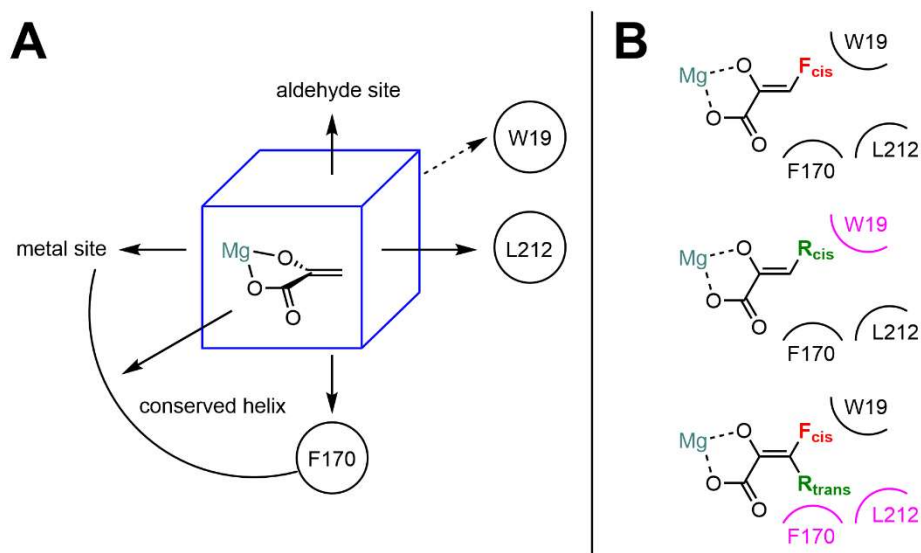


Figure 3.7. Stereochemical outcome of the EChpcH reaction with fluoro-donors. A) Schematic of a metal-bound enolate at the enzyme active site. Non-catalytic residues W19, F170, and L212 form a pocket around the C-3 position. B) Geometric preference of C-3 fluorine or alkyl substituents on the enolate, in different cases resulting in steric clash of the alkyl group with W19 or with F170/L212.

calculations were performed to obtain the pK_a of the acidic carbon in each fluoro-donor. These values were derived from the free energy change associated with transfer of a proton from the fluoro-donor monoanion to the pyruvate dianion, as a carbon acidity of $pK_a = 16.58$ for pyruvate is known [34]. For fluoro-donors, the following pK_a values were found: FP 13.90, FKB 16.16, FKV 16.19, FKMV 16.51, FKPB 15.72, FPP 12.93 (all ionizing to the *cis*-fluoroenolate except for FPP which preferred the *trans*-fluoroenolate). Using acidity values as a proxy for nucleophilicity, it can be seen that aliphatic alkyl substituents at C-3 cancel out nearly the entire effect of fluorine, so that FKB, FKV, FKMV, and FKPB would have similar electronic properties to pyruvate. But far from reacting at the speed of native pyruvate (500-fold faster than FP as discussed in Chapter 2), these extended β -fluoro- α -ketoacids either reacted slower than FP or not at all. We also ruled out any inhibition of the aldolase by the extended β -fluoro- α -ketoacids, as their co-presence had no effect on the conversion rate of FP. These results reinforce our finding that the fluoro-donor scope of EcHpcH is limited by steric and not electronic factors, and that these steric factors (the required geometry and positioning of the bulky group in the active site) are different from those of non-fluorinated donors.

We proceeded to optimize the individual reactions of FP, FKB, and FKV with formaldehyde, based on time and buffer pH. We were surprised to find that reactions of FP were markedly accelerated when increasing the pH to 8.5, allowing shorter reaction times than described in the initial studies in Chapter 2. However, the higher pH also accelerated the background rate of aldol addition, which would be non-stereoselective and undesirable. Screening of several amine-based buffers at pH 8.5 (Bicine, HEPPS, TAPS, Tricine) for the background aldol addition of FP to formaldehyde revealed that tertiary amine buffers promoted this non-enzymatic reaction but secondary amine buffers suppressed it. The reason for this is unclear, but may be due to condensation of secondary amines with the donor or acceptor carbonyl. For the final condition of FP, we settled on using a 2:1 mixture of CHES pH 9.3 and HEPES pH 7.5 to achieve an overall pH of 8.5 and the inclusion of a secondary amine (a mixture was used because the aldolases were stored in HEPES pH 7.5 at fixed concentrations that consistently occupied one-third of the reaction volume). This condition gave 98% conversion of FP within just 4 h. Unlike for FP, increased pH was not beneficial for FKB and FKV, and pH 7.5 was maintained for these. For the final condition of FKB, the reaction time was simply increased from 16 to 48 h, thus increasing FKB conversion to 82%. For the difficult fluoro-donor FKV, the mutants F170L, F170V, F170A, and L212A were re-screened at the tripled enzyme loading of 0.3 mol%, and the reaction time was increased to 1 week. F170V provided the highest FKV conversion of 40%, despite F170L appearing slightly better in the first screen (F170I was not included in the re-screening due to issues of low protein solubility).

Next, the optimized conditions were used to expand the electrophile scope of FP, FKB, or FKV to a panel of 17 aldehydes with a diverse spectrum of polarity (*Figure 3.8*). All products were characterized by ^{19}F NMR and high resolution LC-MS (*Figure 3.9* and *Figure 3.10*). FKB and FKV showed good conversions with a number of aldehydes, with aliphatic and polar heteroaromatic aldehydes especially tolerated. Gratifyingly, near quantitative formation of several products was observed even with the difficult fluoro-donor FKV (products **4c**, **5c**, **18c**, **19c**). However, there remained difficult acceptors that reacted only with FP (products **6a**, **8a**, **9a**, **10a**, **11a**, **16a**). These bulky or branched aldehydes had been rendered viable by our optimization of increasing the pH to 8.5. A maltose-binding protein fusion (MBP-HpcH) was used to stabilize the enzyme in the presence of the most hydrophobic aldehydes, which were solubilized by inclusion of DMSO co-solvent at 5% of the reaction volume. Although the fluorine chiral center is expected

to be consistently (*S*)-fluoro in all products, the diastereomeric ratio of *syn* and *anti* products which can be observed by ^{19}F NMR is still highly variable. This is because the aldolase's lack of aldehyde-interacting residues leads to attack at both aldehyde faces, giving access to both configurations of the hydroxy chiral center.

Having expanded our platform for chiral organofluorine synthesis to the possibility of products with tertiary fluoride stereocenters, we concluded the project by demonstrating the synthetic utility

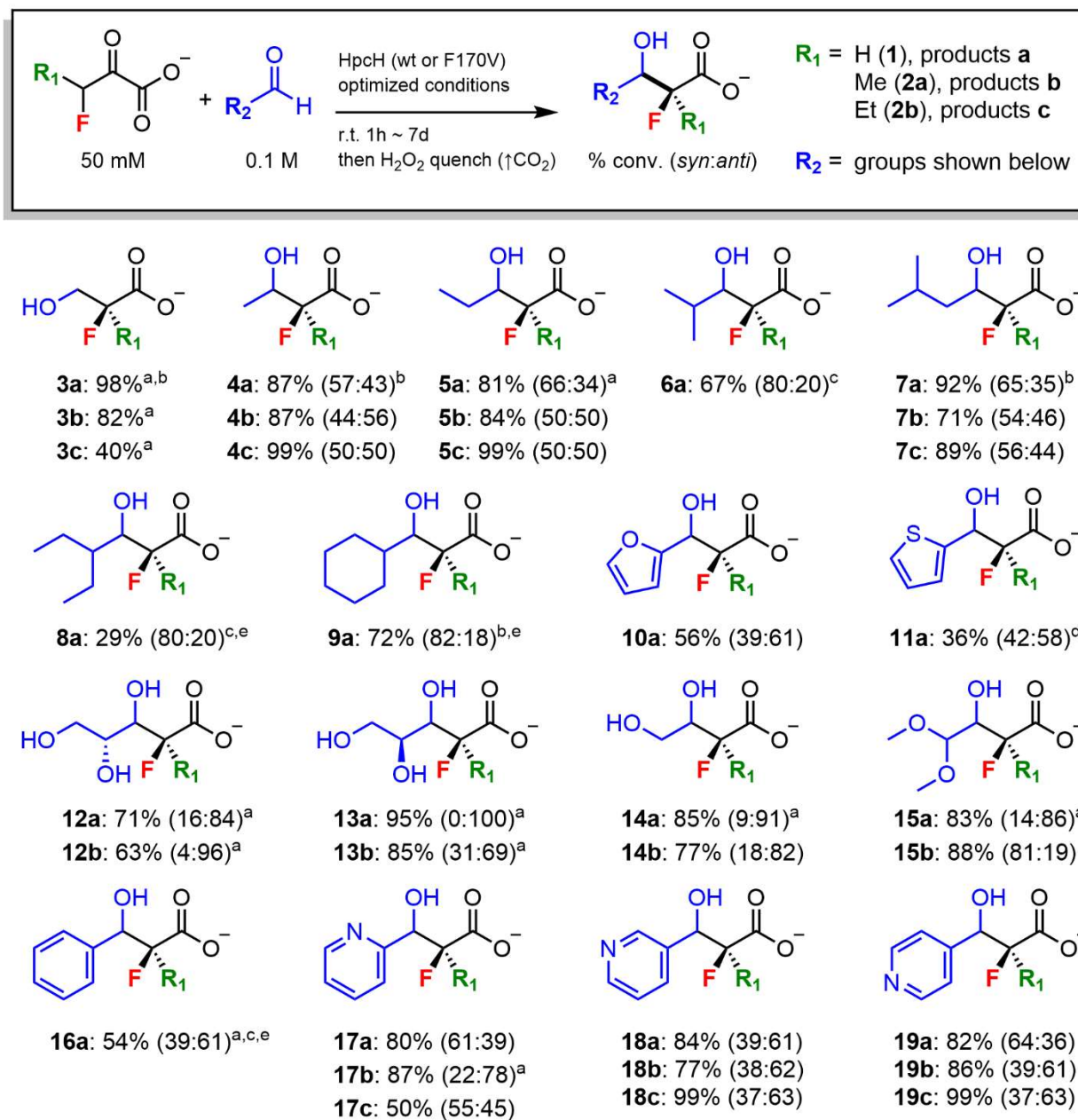


Figure 3.8. Substrate scope in the aldehyde for aldol additions of β -fluoro- α -ketoacids. Due to kinetic resolution of C-3 substituted fluoro-donors, 100% conversion equals 50% donor consumption. Optimized conditions are the same as in Figure 3.5 except where indicated. For combinations of donor and aldehyde omitted here, the reactions were attempted but produced little to no product. ^aStereopurity at fluorine of >99% (*S*) was either confirmed in this chapter (Figure 3.20) or in the previous chapter. ^bReaction time was 4 h. ^cReaction time was 16 h. ^dReaction time was 48 h. ^eMBP-tagged HpcH was used.

	R	R'	<i>anti</i> product		<i>syn</i> product	
			δ (ppm)	J_{F-H} (Hz)	δ (ppm)	J_{F-H} (Hz)
3a	H	H	-188.62 (dt)	51.1, 27.1	-	-
4a	H	Me	-196.37 (dd)	49.6, 23.7	-198.91 (dd)	52.2, 27.1
5a	H	Et	-196.72 (dd)	51.4, 27.2	-198.58 (dd)	49.4, 26.3
6a	H	<i>i</i> -Pr	-193.85 (dd)	50.6, 26.4	-200.35 (dd)	48.8, 28.9
7a	H	<i>i</i> -Bu	-197.36 (dd)	52.3, 27.5	-197.67 (dd)	49.6, 25.9
8a	H	CH(Et) ₂	-193.01 (dd)	50.8, 26.5	-199.13 (dd)	49.6, 28.4
9a	H	<i>c</i> -Hex	-193.50 (dd)	50.1, 26.6	-200.30 (dd)	49.1, 29.5
10a	H	2-Furyl	-194.38 (dd)	49.3, 23.5	-195.88 (dd)	50.7, 27.2
11a	H	Thiophen-2-yl	-194.37 (dd)	49.1, 24.3	-196.27 (dd)	51.4, 28.3
12a	H	(<i>R</i>)-CH(OH)CH ₂ OH	-189.92 (dd)	49.9, 20.1	-202.51 (dd)	48.4, 28.9
13a	H	(<i>S</i>)-CH(OH)CH ₂ OH	-193.78 (dd)	49.8, 25.3	-203.91 (dd)	51.1, 30.7
14a	H	CH ₂ OH	-194.78 (dd)	50.5, 24.9	-200.47 (dd)	49.3, 27.1
15a	H	CH(OMe) ₂	-196.51 (dd)	49.5, 26.3	-200.83 (dd)	49.0, 28.7
16a	H	Ph	-191.99 (dd)	50.6, 23.4	-196.97 (dd)	48.9, 26.5
17a	H	2-Pyridyl	-195.47 (dd)	49.2, 24.5	-196.55 (dd)	51.7, 26.9
18a	H	3-Pyridyl	-195.41 (dd)	51.4, 26.2	-196.80 (dd)	48.9, 25.8
19a	H	4-Pyridyl	-194.98 (dd)	49.4, 24.6	-196.01 (dd)	51.7, 26.9
3b	Me	H	-155.75 (h)	22.3	-	-
4b	Me	Me	-163.29 (p)	20.6	-166.33 (p)	21.7
5b	Me	Et	-160.33 (p)	22.0	-164.59 (p)	22.9
7b	Me	<i>i</i> -Bu	-160.32 (p)	20.0	-164.95 (p)	22.8
12b	Me	(<i>R</i>)-CH(OH)CH ₂ OH	-171.95 (p)	23.3	-	-
13b	Me	(<i>S</i>)-CH(OH)CH ₂ OH	-172.18 (p)	23.8	-173.46 (p)	23.5
14b	Me	CH ₂ OH	-162.76 (p)	22.2	-	-
15b	Me	CH(OMe) ₂	-159.45 (p)	21.0	-165.14 (p)	23.2
17b	Me	2-Pyridyl	-161.43 (p)	21.6	-163.79 (p)	21.3
18b	Me	3-Pyridyl	-163.84 (p)	20.8	-164.63 (p)	22.6
19b	Me	4-Pyridyl	-161.93 (p)	20.7	-163.45 (p)	22.5
3c	Et	H	-166.89 (p)	22.4	-	-
4c	Et	Me	-175.29 (m)	-	-177.71 (m)	-
5c	Et	Et	-172.52 (m)	-	-175.90 (m)	-
7c	Et	<i>i</i> -Bu	-172.82 (m)	-	-176.34 (m)	-
17c	Et	2-Pyridyl	-173.35 (m)	-	-175.28 (m)	-
18c	Et	3-Pyridyl	-175.34 (m)	-	-176.10 (m)	-
19c	Et	4-Pyridyl	-174.24 (m)	-	-175.40 (m)	-

Figure 3.9. ¹⁹F NMR data for all fluorinated products of aldol reactions (post-decarboxylation) displayed in Figure 3.8.

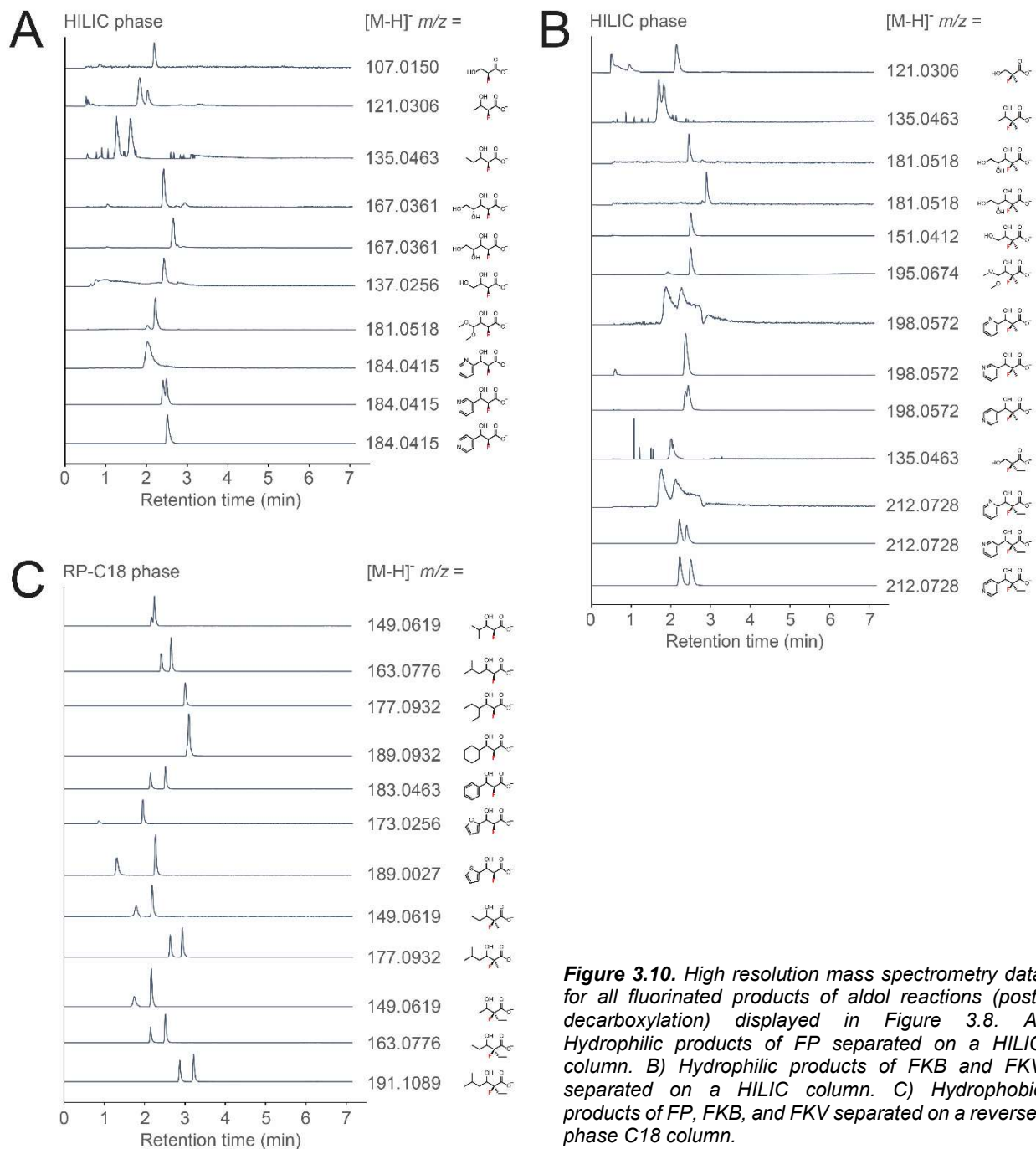


Figure 3.10. High resolution mass spectrometry data for all fluorinated products of aldol reactions (post-decarboxylation) displayed in Figure 3.8. A) Hydrophilic products of FP separated on a HILIC column. B) Hydrophilic products of FKB and FKV separated on a HILIC column. C) Hydrophobic products of FP, FKB, and FKV separated on a reverse-phase C18 column.

of β -fluoro- α -ketoacids in selected reactions relevant to bioactive compounds (Figure 3.11 and Figure 3.12). The compounds **3a**, **3b**, and **3c** correspond to the donors FP, FKB, and FKV undergoing hydroxymethylation with formaldehyde. They are also fluorine-bearing analogs of the Roche ester (methyl 3-hydroxy-2-methylpropionate), a versatile building block that has been employed in landmark total syntheses of complex natural products such as the anticancer polyketide discodermolide [45,46]. The aromatic products **16a** formed from FP and benzaldehyde, and **17b** formed from FKB and 2-pyridinecarboxaldehyde, resemble known arylbutyryl intermediates of the nikkomycins, a family of antifungal natural products that has also been

targeted by total synthesis [47,48]. The above five α -fluoroester products were prepared by enzymatic aldol addition (0.6 mmol), H₂O₂ decarboxylation, removal of water and enzymes, and esterification in MeOH/HCl or EtOH/SOCl₂. Lastly, the α -fluoro- α -methyl branched sugar acids **12b** and **13b** formed from either D-glyceraldehyde or L-glyceraldehyde were enzymatically synthesized (0.15 mmol) and directly purified by semi-preparative HPLC. These products complement the FP-derived unbranched versions **12a** and **13a** which were isolated in high diastereomeric purities in Chapter 2. Similar poly-hydroxylated intermediates are used to synthesize the fluorinated pentose moiety of the blockbuster antiviral sofosbuvir [49]. All isolated products were characterized by ¹H, ¹³C, and ¹⁹F NMR (*Figure 3.13* through *Figure 3.19*).

Although the stereopurity of products **12b** and **13b** was directly verified by ¹⁹F NMR due to the presence of a fixed stereocenter arising from a chiral aldehyde, the enantiopurity of **3a**, **3b**, **3c**, **16a**, and **17b** needed verification by a method of chiral derivatization. Thus, these five α -fluoroesters were derivatized at the alcohol group with Mosher's acid, and compared to similarly-derivatized racemic standards that were independently synthesized [50]. The standards *rac-3a*, *rac-3b*, and *rac-3c* were prepared in aqueous conditions with a workflow parallel to the enzymatic procedure, except that CoCl₂ and a buffer of sufficiently high pH (HEPPS pH 8.5 for FP, phosphate pH 12.5 for FKB and FKV) were used to promote the non-enzymatic aldol addition. Meanwhile, anhydrous synthesis of *rac-16a* and *rac-17b* was performed by Reformatsky reaction of ethyl bromofluoroacetate and LDA-promoted aldol reaction of ethyl 2-fluoropropionate, respectively [51]. In general, the Mosher esters of enzymatic products displayed single sets of peaks, while those of racemic standards displayed doubled peaks (*Figure 3.20*). This result confirmed that our methodology of organofluorine synthesis offers exquisite fluorine stereoselectivity that extends to tertiary fluorides, and may enable the asymmetric synthesis of new fluorinated drugs or fluorinated analogs of known drugs.

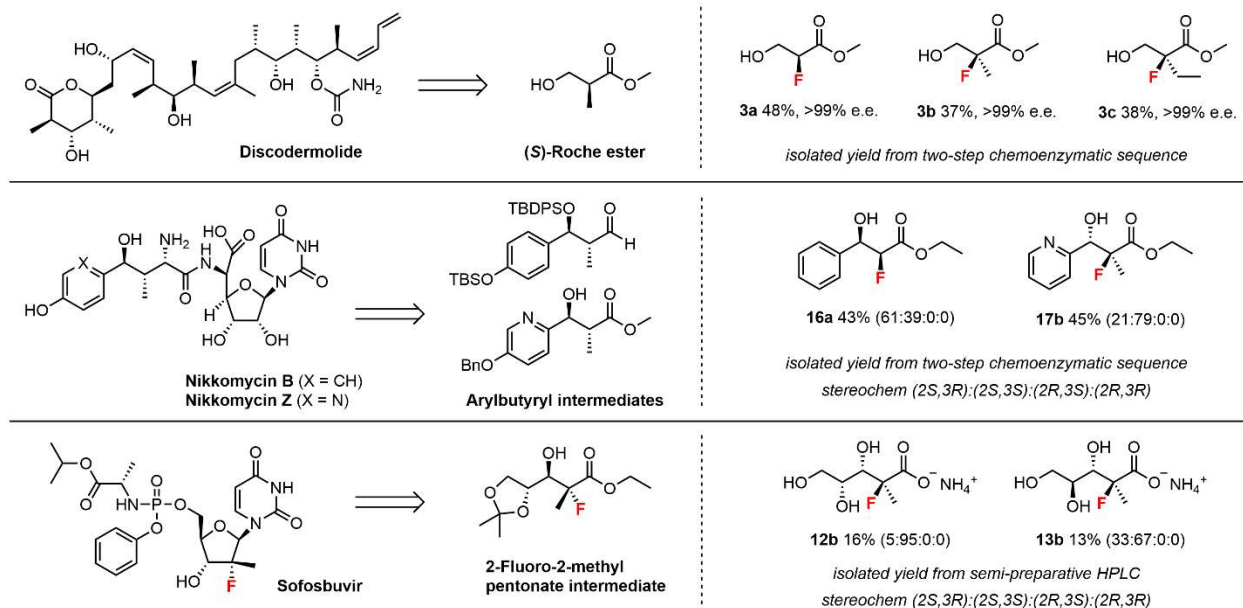


Figure 3.11. Examples of fluorinated compounds with relevance to building blocks of bioactive compounds, isolated in high enantiomeric or diastereomeric purity. The drawn structures correspond to the major stereoisomer. Enantiopurity of fluoroesters was confirmed by Mosher analysis and the minor enantiomers were not detected (Figure S20).

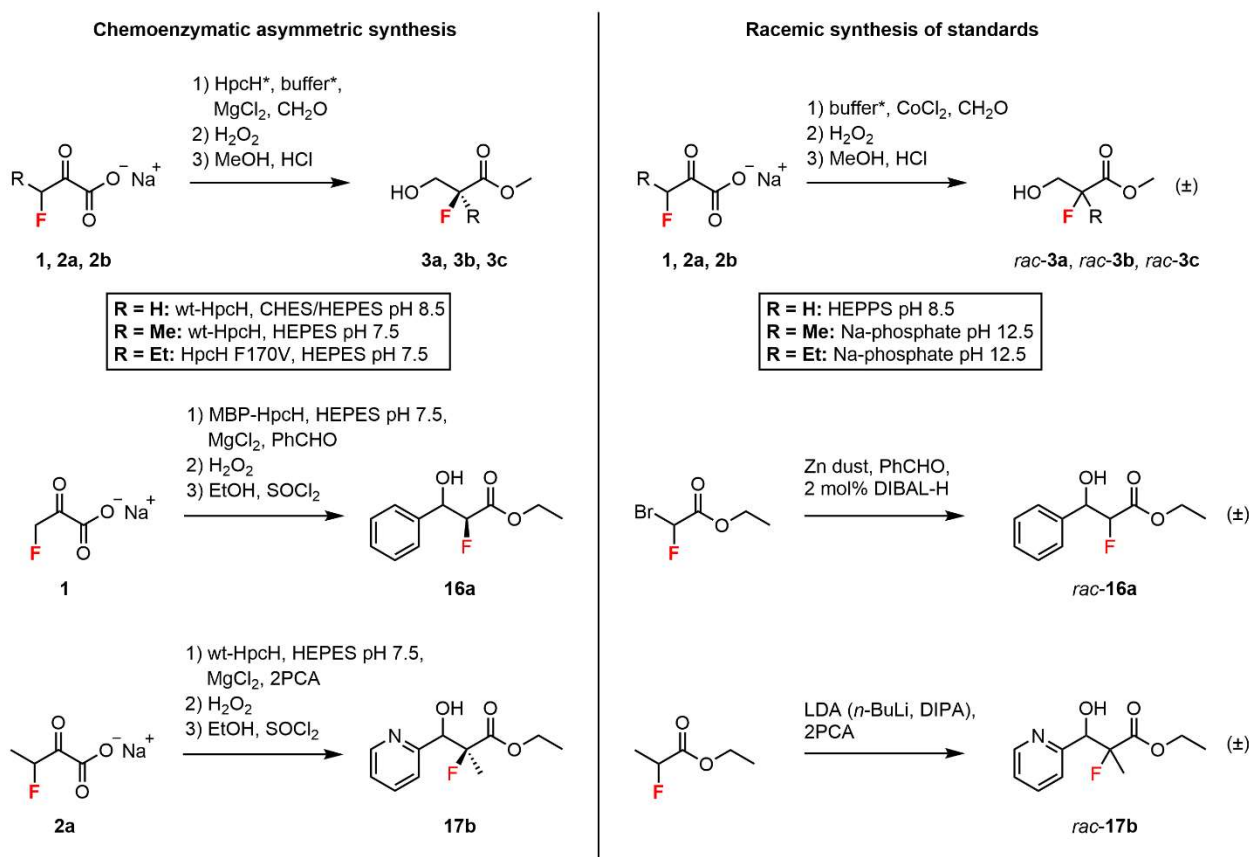


Figure 3.12. Details of compound synthesis of isolated fluoroesters in enantiopure form (from enzymatic reaction) or racemic form (from non-enzymatic reaction).

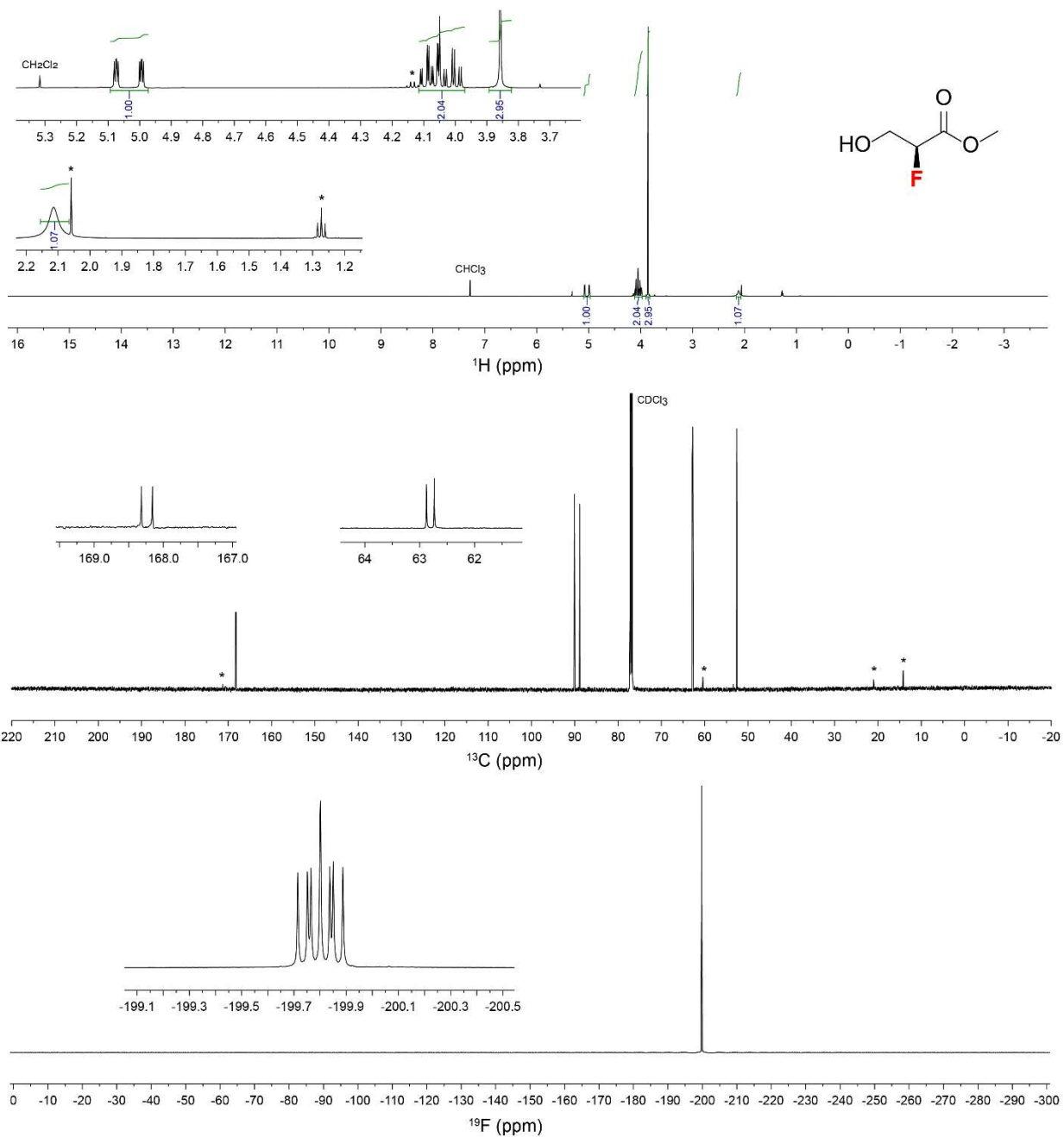


Figure 3.13. NMR spectra of methyl (S)-2-fluoro-3-hydroxypropanoate (**3a**). ^1H , ^{13}C , and ^{19}F spectra. Asterisk indicates peaks due to residual ethyl acetate.

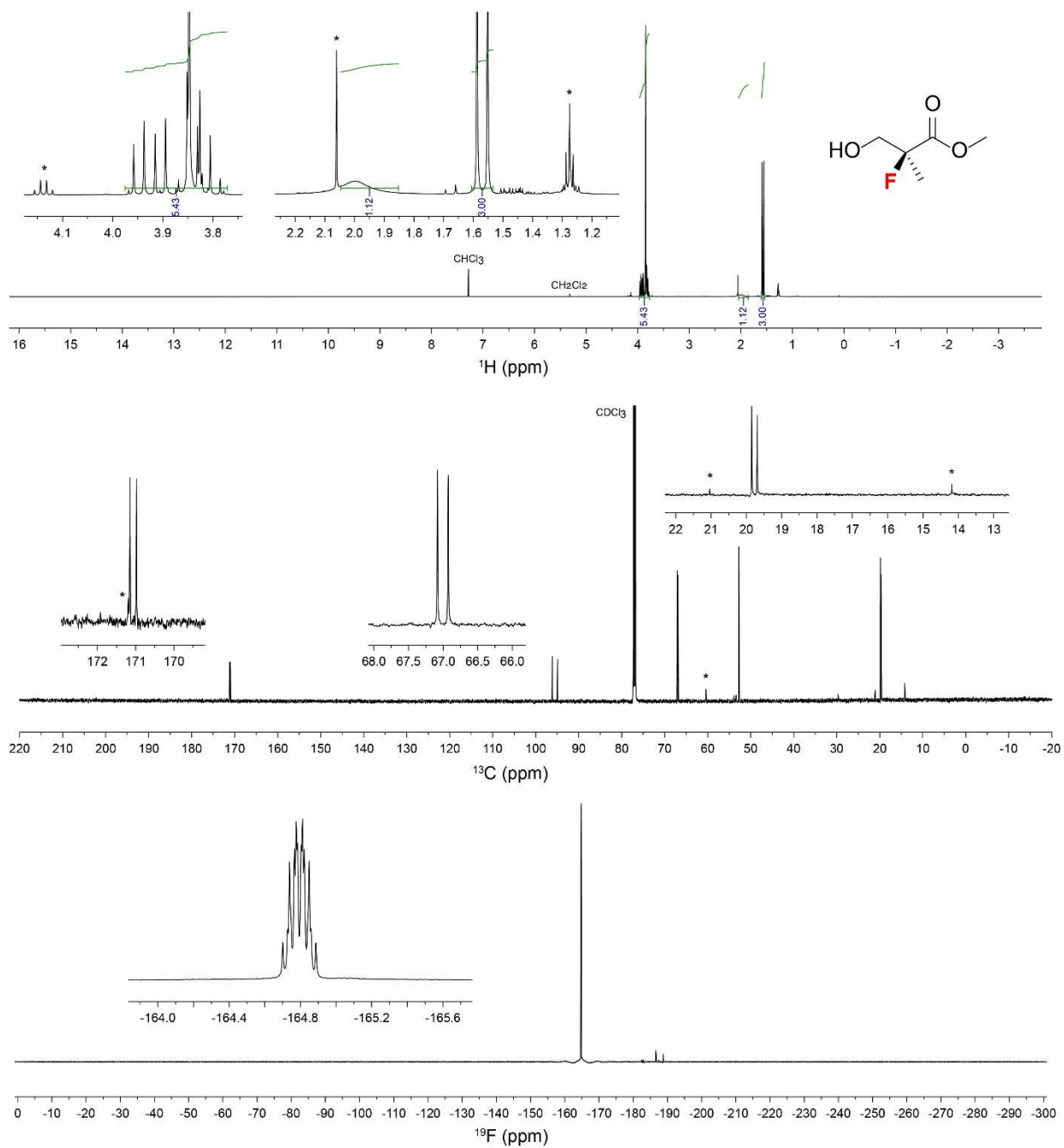


Figure 3.14. NMR spectra of methyl (S)-2-fluoro-3-hydroxy-2-methylpropanoate (**3b**). ¹H, ¹³C, and ¹⁹F spectra. Asterisk indicates peaks due to residual ethyl acetate.

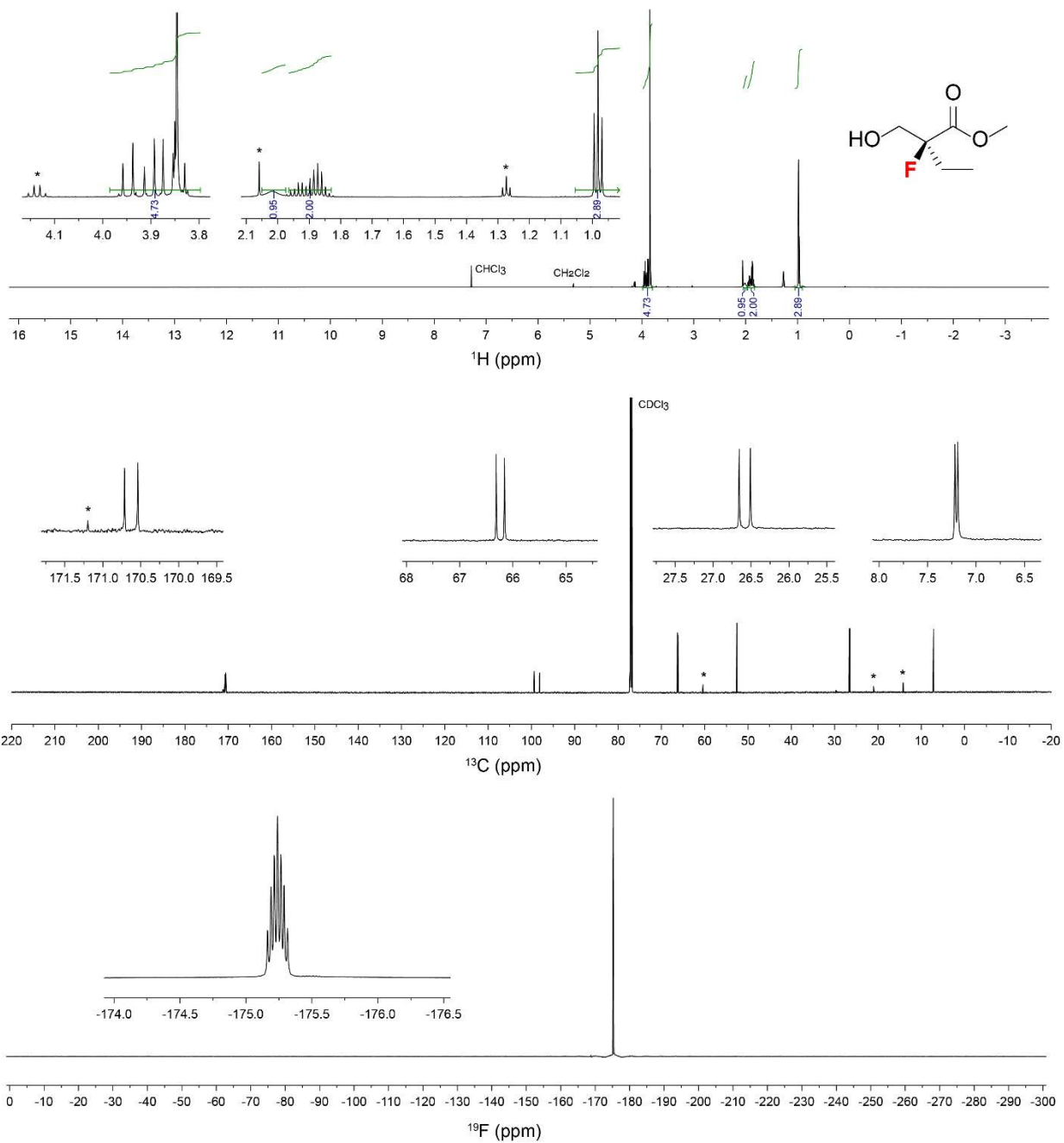


Figure 3.15. NMR spectra of methyl (S)-2-fluoro-2-(hydroxymethyl)butanoate (**3c**). ^1H , ^{13}C , and ^{19}F spectra. Asterisk indicates peaks due to residual ethyl acetate.

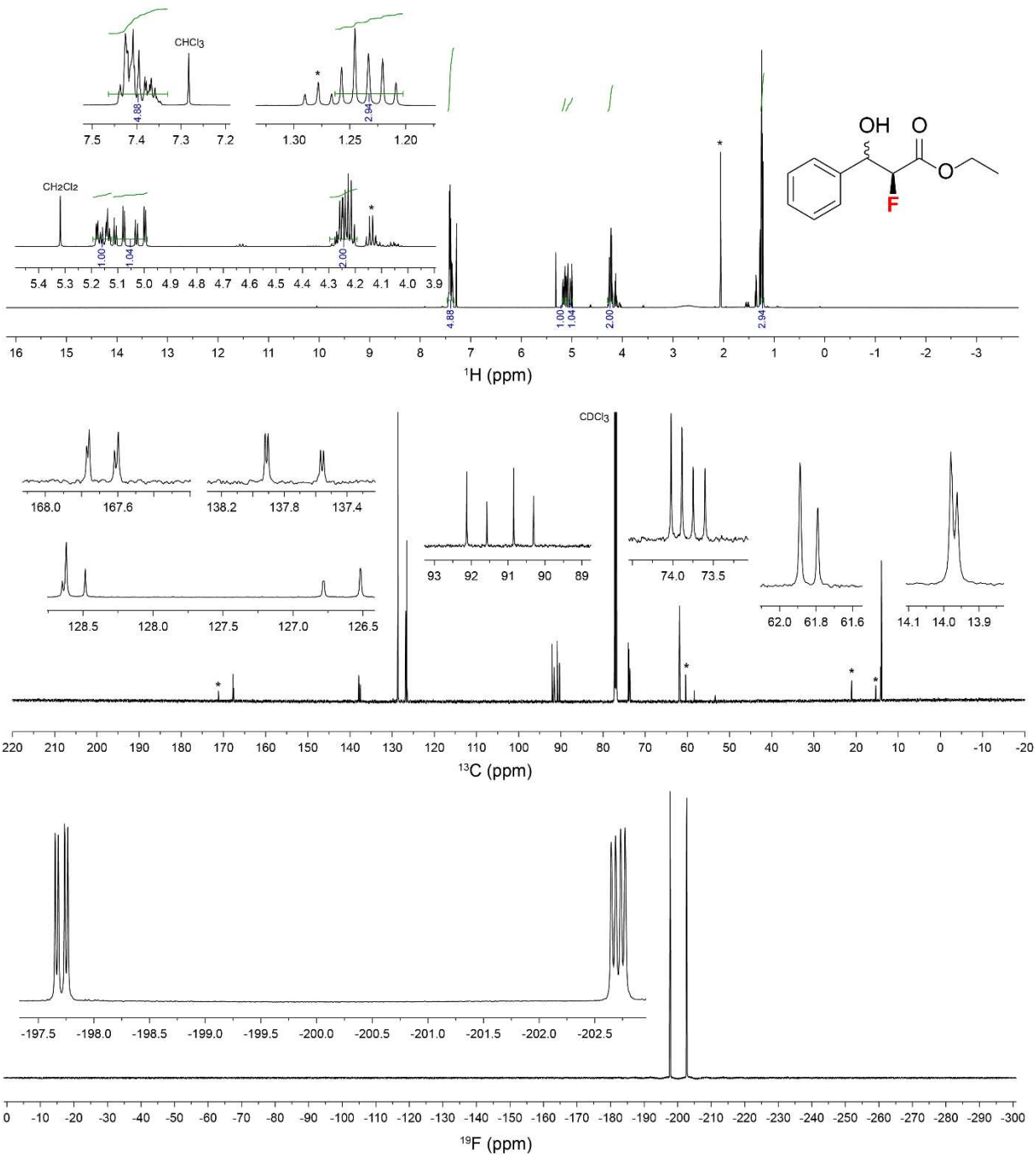


Figure 3.16. NMR spectra of ethyl (2*S*,3*RS*)-2-fluoro-3-hydroxy-3-phenylpropanoate (**16a**). ¹H, ¹³C, and ¹⁹F spectra (61:39 *syn/anti* ratio). Asterisk indicates peaks due to residual ethyl acetate.

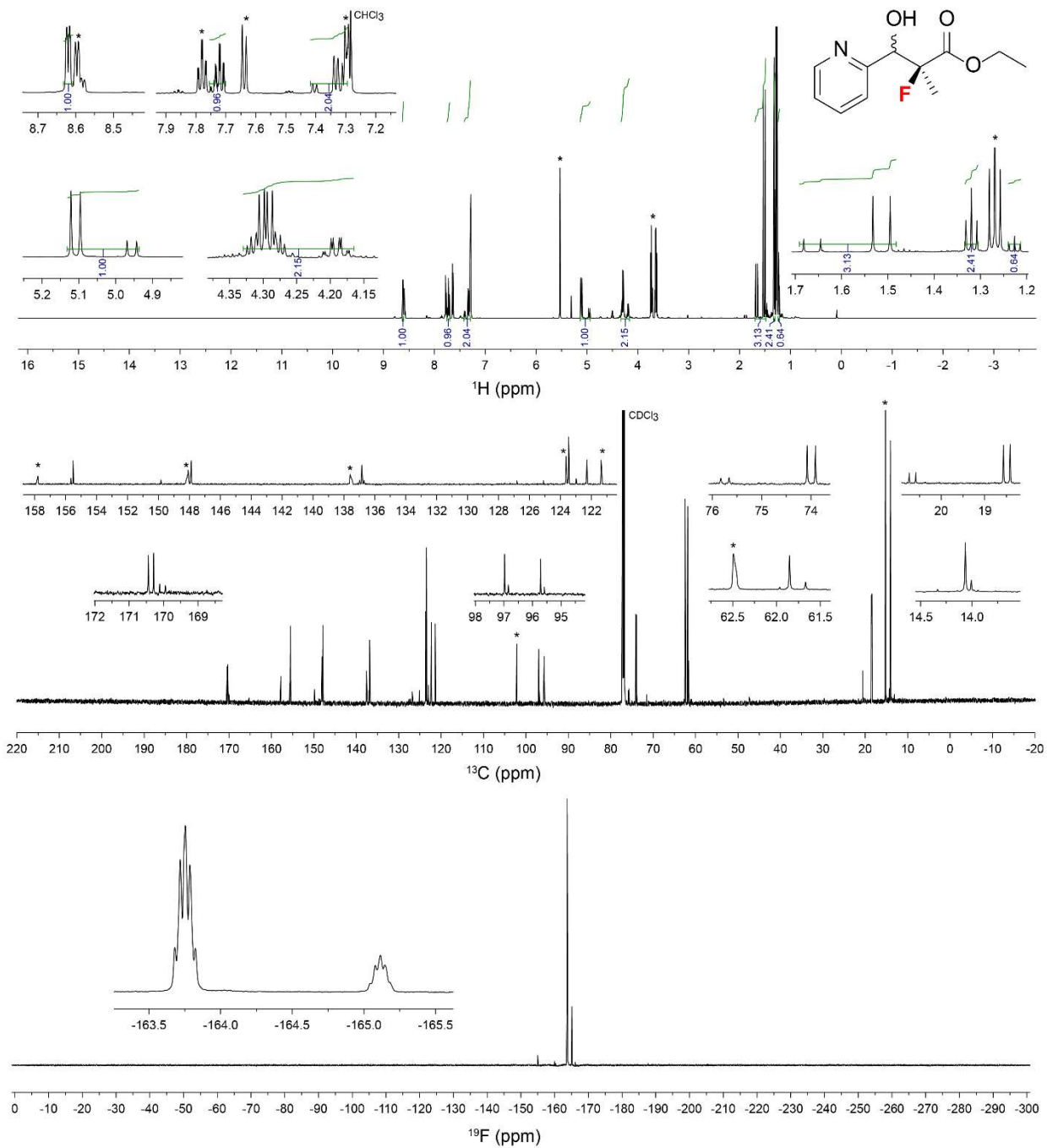


Figure 3.17. NMR spectra of ethyl (2S,3RS)-2-fluoro-3-hydroxy-2-methyl-3-(2-pyridyl)propanoate (**17b**). ¹H, ¹³C, and ¹⁹F spectra (21:79 syn/anti ratio). Asterisk indicates peaks from the diethyl acetal of 2-pyridinecarboxaldehyde.

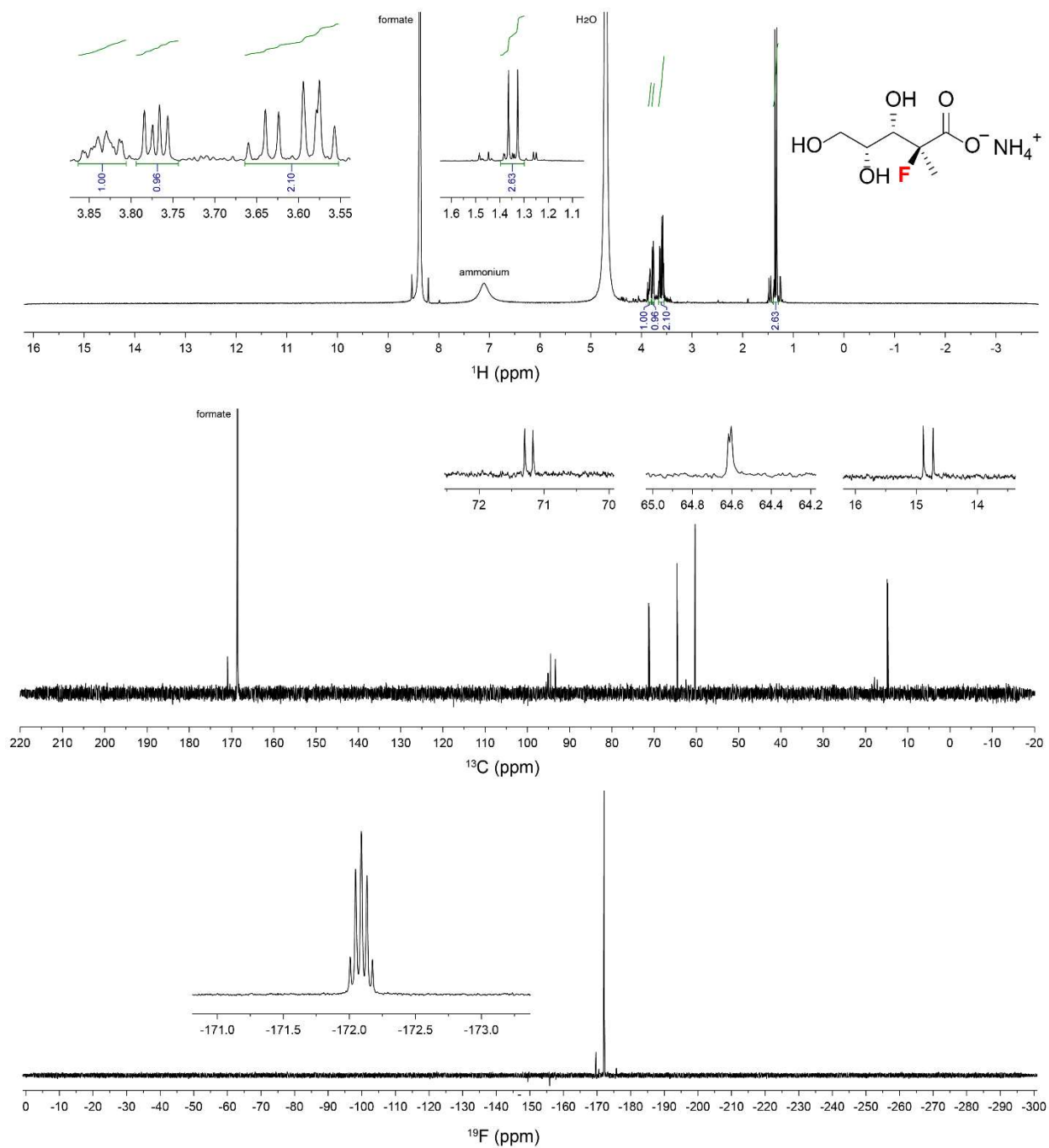


Figure 3.18. NMR spectra of ammonium (2S,3S,4R)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate (**12b**). ^1H , ^{13}C , and ^{19}F spectra (5:95 syn/anti ratio, only the anti product is annotated).

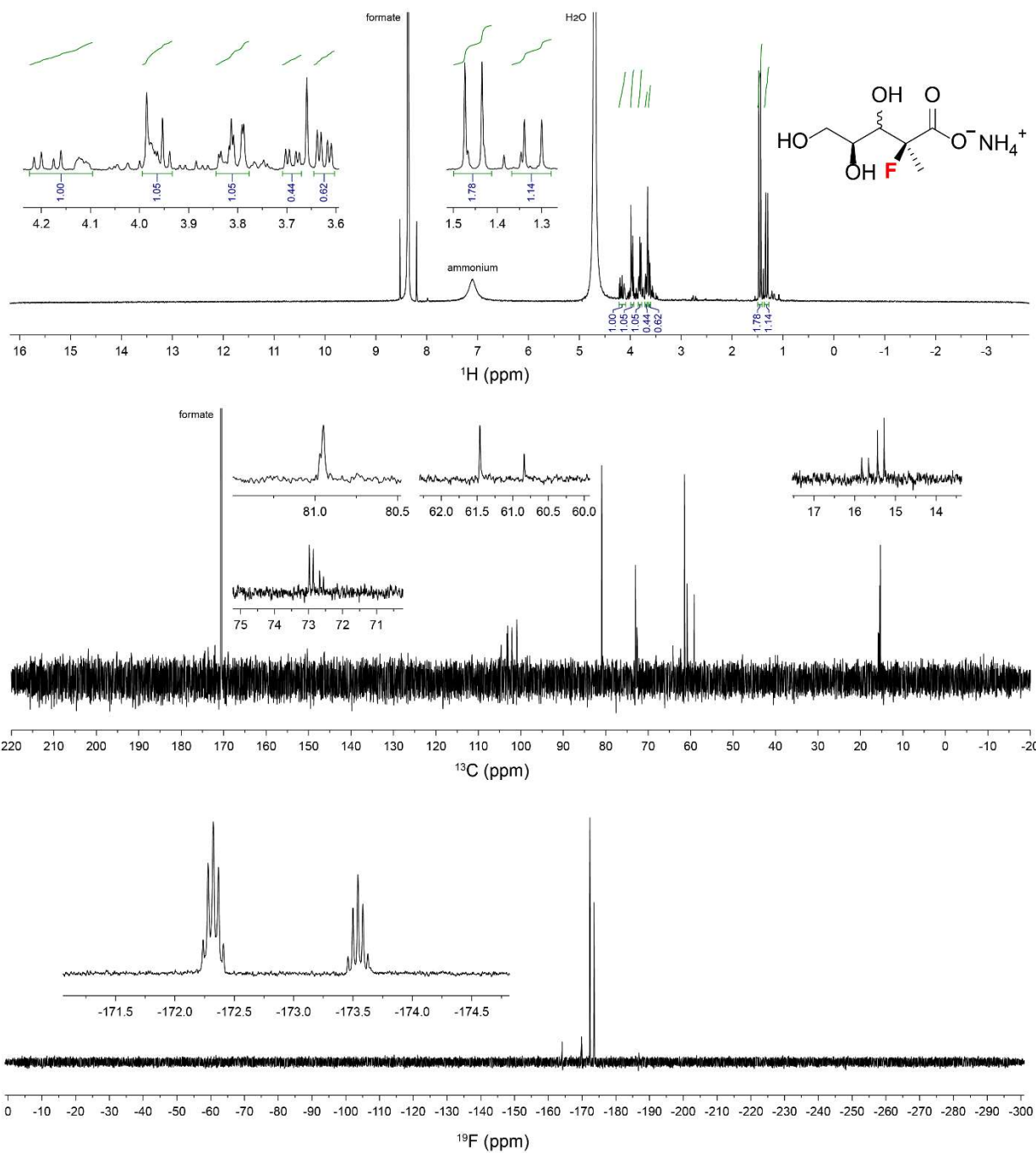


Figure 3.19. NMR spectra of ammonium (2S,3RS,4S)-2-fluoro-3,4,5-trihydroxy-2-methylpentanoate (**13b**). ¹H, ¹³C, and ¹⁹F spectra (33:67 syn/anti ratio).

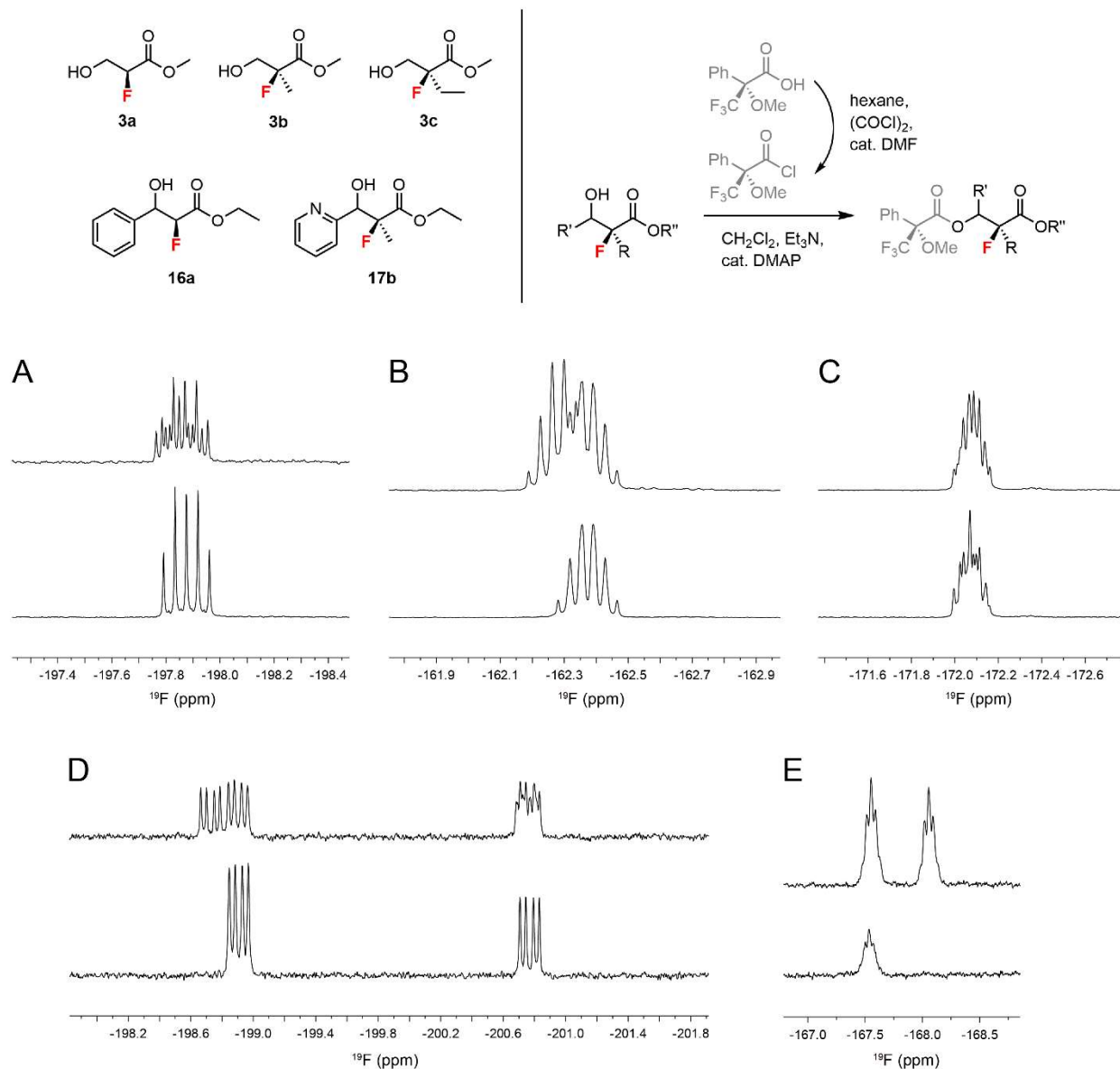


Figure 3.20. Enantiomeric purity of α -fluoroesters determined by chiral derivatization with (+)-Mosher's acid. ^{19}F NMR spectra were compared between derivatized racemic standards (upper spectra, doubled peaks) and the derivatized enzymatic products (lower spectra, single peaks). No minor enantiomers were detectable for enzymatic products, indicating >99% e.e. for formation of the (S)-configured C-F stereocenter at the limit of detection. (A) Spectra for 3a. (B) Spectra for 3b. (C) Spectra for 3c. (D) Spectra for 16a (syn and anti diastereomers). (E) Spectra for 17b shown for the syn diastereomer, as the anti diastereomer was poorly derivatized from strong steric hindrance of the alcohol.

3.4. Conclusion

The work described in this chapter shows that tertiary fluoride stereocenters, which have previously not been synthesized with biocatalytic C-C bond formation, can be prepared with a Type II HpcH family pyruvate aldolase. The racemic β -fluoro- α -ketoacid substrates required for this reaction could be chemically synthesized in a short sequence from commercially available α -ketoacids. Using a panel of HpcH variants with active site residues engineered for potential steric relief, we were able to demonstrate incorporation of β -fluoro- α -ketobutyrate and β -fluoro- α -ketovalerate with concomitant kinetic resolution of the racemic fluoro-donor. Unfortunately, our efforts to incorporate larger fluoro-donors were unsuccessful, but our mechanistic analysis of the enzyme active site revealed that the steric factors leading to rejection of bulky fluoro-donors were indeed different from those of bulky non-fluorinated donors. Our optimization of reaction conditions for fluoropyruvate, β -fluoro- α -ketobutyrate, and β -fluoro- α -ketovalerate established a diverse substrate scope for asymmetric construction of both secondary and tertiary fluoride stereocenters. We furthermore demonstrated the robustness of this system with examples of isolated products relevant to bioactive compounds of interest. Taken together, our work in Chapter 2 to lay a foundation for biocatalytic organofluorine synthesis and our work in Chapter 3 to expand this method to tertiary fluorides can supplement the repertoire of methodology for functional units that are difficult to access chemically, and may provide new sustainable routes towards fluorinated pharmaceuticals.

3.5. References

1. Zhu, Y., Han, J., Wang, J., Shibata, N., Sodeoka, M., Soloshonok, V. A., Coelho, J. A. S., Toste, F. D. Modern approaches for asymmetric construction of carbon-fluorine quaternary stereogenic centers: Synthetic challenges and pharmaceutical needs. *Chem Rev* **2018**, *118* (7) 3887-3964.
2. Fried, J., Sabo, E. F. 9α -Fluoro derivatives of cortisone and hydrocortisone. *J Am Chem Soc* **1954**, *76* (5) 1455-1456.
3. Clark, J. L., Hollecker, L., Mason, J. C., Stuyver, L. J., Tharnish, P. M., Lostia, S., McBrayer, T. R., Schinazi, R. F., Watanabe, K. A., Otto, M. J., Furman, P. A., Stec, W. J., Patterson, S. E., Pankiewicz, K. W. Design, synthesis, and antiviral activity of 2'-deoxy-2'-fluoro-2'-C-methylecytidine, a potent inhibitor of hepatitis C virus replication. *J Med Chem* **2005**, *48* (17) 5504-5508.
4. Althaus, M., Becker, C., Togni, A., Mezzetti, A. Ruthenium-catalyzed asymmetric electrophilic fluorination of 1,3-dicarbonyl compounds. *Organometallics* **2007**, *26* (24) 5902-5911.
5. Rauniyar, V., Lackner, A. D., Hamilton, G. L., Toste, F. D. Asymmetric electrophilic fluorination using an anionic chiral phase-transfer catalyst. *Science* **2011**, *334* (6063) 1681-1684.
6. Meng, W.-T., Zheng, Y., Nie, J., Xiong, H.-Y., Ma, J.-A. Organocatalytic asymmetric one-pot sequential conjugate addition/dearomative fluorination: Synthesis of chiral fluorinated isoxazol-5(4H)-ones. *J Org Chem* **2013**, *78* (2) 559-567.
7. Xu, J., Hu, Y., Huang, D., Wang, K.-H., Xu, C., Niu, T. Thiourea-catalyzed enantioselective fluorination of β -keto esters. *Adv Synth Catal* **2012**, *354* (2-3) 515-526.

8. Shibata, N., Suzuki, E., Takeuchi, Y. A fundamentally new approach to enantioselective fluorination based on cinchona alkaloid derivatives/Selectfluor combination. *J Am Chem Soc* **2000**, *122* (43) 10728-10729.
9. Zhu, C.-L., Maeno, M., Zhang, F.-G., Shigehiro, T., Kagawa, T., Kawada, K., Shibata, N., Ma, J.-A., Cahard, D. Chiral *N*-fluorodibenzenesulfonimide analogues for enantioselective electrophilic fluorination and oxidative fluorination. *Eur J Org Chem* **2013**, *2013* (29) 6501-6505.
10. Wolstenhulme, J. R., Rosenqvist, J., Lozano, O., Ilupeju, J., Wurz, N., Engle, K. M., Pidgeon, G. W., Moore, P. W., Sandford, G., Gouverneur, V. Asymmetric electrophilic fluorocyclization with carbon nucleophiles. *Angew Chem Int Ed* **2013**, *52* (37) 9796-9800.
11. Han, X., Luo, J., Liu, C., Lu, Y. Asymmetric generation of fluorine-containing quaternary carbons adjacent to tertiary stereocenters: Uses of fluorinated methines as nucleophiles. *Chem Commun* **2009**, *15* 2044-2046.
12. Liang, Y., Fu, G. C. Catalytic asymmetric synthesis of tertiary alkyl fluorides: Negishi cross-couplings of racemic α,α -dihaloketones. *J Am Chem Soc* **2014**, *136* (14) 5520-5524.
13. Cuadros, S., Dell'Amico, L., Melchiorre, P. Forging fluorine-containing quaternary stereocenters by a light-driven organocatalytic aldol desymmetrization process. *Angew Chem Int Ed* **2017**, *56* (39) 11875-11879.
14. He, Z.-T., Jiang, X., Hartwig, J. F. Stereodivergent construction of tertiary fluorides in vicinal stereogenic pairs by allylic substitution with iridium and copper catalysts. *J Am Chem Soc* **2019**, *141* (33) 13066-13073.
15. Liu, J., Yuan, Q., Toste, F. D., Sigman, M. S. Enantioselective construction of remote tertiary carbon-fluorine bonds. *Nat Chem* **2019**, *11* (8) 710-715.
16. Wang, Q., Lübcke, M., Biosca, M., Hedberg, M., Eriksson, L., Himo, F., Szabó, K. J. Enantioselective construction of tertiary fluoride stereocenters by organocatalytic fluorocyclization. *J Am Chem Soc* **2020**, *142* (47) 20048-20057.
17. Zhu, X., Robinson, D. A., McEwan, A. R., O'Hagan, D., Naismith, J. H. Mechanism of enzymatic fluorination in *Streptomyces cattleya*. *J Am Chem Soc* **2007**, *129* (47) 14597-14604.
18. O'Hagan, D., Harper, D. B. Fluorine-containing natural products. *J Fluor Chem* **1999**, *100* (1-2) 127-133.
19. Goj, O., Burchardt, A., Haufe, G. A versatile approach to optically active primary 2-fluoro-2-phenylalkanols through lipase-catalyzed transformations. *Tetrahedron: Asymmetry* **1997**, *8* 399-408.
20. Yokoyama, H., Hyodo, R., Nakada, A., Yamaguchi, S., Hirai, Y., Kometani, T., Goto, M., Shibata, N., Takeuchi, Y. Asymmetric synthesis of chiral 2-fluorinated 1,3-propanediols and its application to the preparation of monofluorinated chiral synthon. *Tetrahedron Lett* **1998**, *39* (42) 7741-7744.
21. Guanti, G., Narisano, E., Riva, R. Lipase mediated preparation of differently protected homochiral 2-aryl-2-fluoro-1,3-propanediols. *Tetrahedron: Asymmetry* **1998**, *9* 1859-1862.

22. De Berardinis, V., Guérard-Hélaine, C., Darii, E., Bastard, K., Hélaine, V., Mariage, A., Petit, J.-L., Poupard, N., Sánchez-Moreno, I., Stam, M., Gefflaut, T., Salanoubat, M., Lemaire, M. Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chem* **2017**, *19* 519-526.
23. Hernández, K., Joglar, J., Bujons, J., Parella, T., Clapés, P. Nucleophile promiscuity of engineered class II pyruvate aldolase YfaU from *E. coli*. *Angew Chem Int Ed* **2018**, *57* (14) 3583-3587.
24. Marín-Valls, R., Hernández, K., Bolte, M., Joglar, J., Bujons, J., Clapés, P. Chemoenzymatic hydroxymethylation of carboxylic acids by tandem stereodivergent biocatalytic aldol reaction and chemical decarboxylation. *ACS Catal* **2019**, *9* (8) 7568-7577.
25. Fang, J., Turner, L. E., Chang, M. C. Y. Biocatalytic asymmetric construction of secondary and tertiary fluorides from β -fluoro- α -ketoacids. *Angew Chem Int Ed* **2022**, e202201602.
26. Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., Vidal, M. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol* **2000**, *328* 575-592.
27. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A., Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009**, *6* (5) 343-345.
28. Wang, W., Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions, and insertions using QuikChange site-directed mutagenesis. *Biotechniques* **1999**, *26* (4) 680-682.
29. Shao, Y., et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol Phys* **2014**, *113* 184-215.
30. Mardirossian, N., Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J Chem Phys* **2016**, *144* 214110.
31. Marenich, A. V., Cramer, C. J., Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B* **2009**, *113* (18) 6378-6396.
32. Weigend, F., Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* **2005**, *7* (18) 3297-3305.
33. Li, Y.-P., Gomes, J., Sharada, S. M., Bell, A. T., Head-Gordon, M. Improved force-field parameters for QM/MM simulations of the energies of adsorption for molecules in zeolites and a free rotor correction to the rigid rotor harmonic oscillator model for adsorption enthalpies. *J Phys Chem C* **2015**, *119* (4) 1840-1850.
34. Chiang, Y., Kresge, A. J., Pruszynski, P. Keto-enol equilibria in the pyruvic acid system: determination of the keto-enol equilibrium constants of pyruvic acid and pyruvate anion and the acidity constant of pyruvate enol in aqueous solution. *J Am Chem Soc* **1992**, *114* (8) 3103-3107.

35. Hernández, K., Gómez, A., Joglar, J., Bujons, J., Parella, T., Clapés, P. 2-Keto-3-deoxy-L-rhamnonate aldolase (YfaU) as catalyst in aldol additions of pyruvate to amino aldehyde derivatives. *Adv Synth Catal* **2017**, 359 (12) 2090-2100.
36. Hernández, K., Bujons, J., Joglar, J., Charnock, S. J., de María, P. D., Fessner, W.-D., Clapés, P. Combining aldolases and transaminases for the synthesis of 2-amino-4-hydroxybutanoic acid. *ACS Catal* **2017**, 7 (3) 1707-1711.
37. Wang, W., Seah, S. Y. K. Purification and biochemical characterization of a pyruvate-specific class II aldolase, HpaI. *Biochemistry* **2005**, 44 (27) 9447-9455.
38. Rea, D., Fülöp, V., Bugg, T. D. H., Roper, D. I. Structure and mechanism of HpcH: a metal ion dependent class II aldolase from the homoprotocatechuate degradation pathway of *Escherichia coli*. *J Mol Biol* **2007**, 373 (4) 866-876.
39. Wang, W., Baker, P., Seah, S. Y. K. Comparison of two metal-dependent pyruvate aldolases related by convergent evolution: substrate specificity, kinetic mechanism, and substrate channeling. *Biochemistry* **2010**, 49 (17) 3774-3782.
40. Coincon, M., Wang, W., Sygusch, J., Seah, S. Y. K. Crystal structure of reaction intermediates in pyruvate class II aldolase: substrate cleavage, enolate stabilization, and substrate specificity. *J Biol Chem* **2012**, 287 (43) 36208-36221.
41. Okonya, J. F., Johnson, M. C., Hoffman, R. V. Preparation of β -fluoro- α -ketoesters from α -ketoesters and their conversion to (Z)- β -fluoro- α -aminoacrylate derivatives. *J Org Chem* **1998**, 63 (18) 6409-6413.
42. Ling, T.-S., Shiu, S., Yang, D.-Y. Design and synthesis of 3-fluoro-2-oxo-3-phenylpropionic acid derivatives as potent inhibitors of 4-hydroxyphenylpyruvate dioxygenase from pig liver. *Bioorg Med Chem* **1999**, 7 (7) 1459-1465.
43. Fang, J., Hait, D., Head-Gordon, M., Chang, M. C. Y. Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angew Chem Int Ed* **2019**, 58 (34) 11841-11845.
44. Marsden, S. R., Mestrom, L., Bento, I., Hagedoorn, P.-L., McMillan, D. G. G., Hanefeld, U. CH- π interactions promote the conversion of hydroxypyruvate in a class II pyruvate aldolase. *Adv Synth Catal* **2019**, 361 (11) 2649-2658.
45. Smith, A. B., Beauchamp, T. J., LaMarche, M. J., Kaufman, M. D., Qiu, Y., Arimoto, H., Jones, D. R., Kobayashi, K. Evolution of a gram-scale synthesis of (+)-discodermolide. *J Am Chem Soc* **2000**, 122 (36) 8654-8664.
46. Paterson, I., Florence, G. J., Gerlach, K., Scott, J. P., Sereinig, N. A practical synthesis of (+)-discodermolide and analogues: fragment union by complex aldol reactions. *J Am Chem Soc* **2001**, 123 (39) 9535-9544.
47. Akita, H., Chen, C. Y., Kato, K. Formal total synthesis of nikkomycin B based on a lipase-catalyzed hydrolysis of an acetate possessing two stereogenic centers. *Tetrahedron* **1998**, 54 11011-11026.
48. Akita, H., Takano, Y., Nedu, K., Kato, K. First chiral synthesis of the N-terminal amino acid congener of nikkomycin Z based on lipase-catalyzed enantioselective acetylation of a

- primary alcohol possessing two stereogenic centers. *Tetrahedron: Asymmetry* **2006**, *17* 1705-1714.
49. Wang, P., Chun, B.-K., Rachakonda, S., Du, J., Khan, N., Shi, J., Stec, W., Cleary, D., Ross, B. S., Sofia, M. J. An efficient and diastereoselective synthesis of PSI-6130: A clinically efficacious inhibitor of HCV NS5B polymerase. *J Org Chem* **2009**, *74* (17) 6819-6824.
 50. Ward, D. E., Rhee, C. K. A simple method for the microscale preparation of Mosher's acid chloride. *Tetrahedron Lett* **1991**, *32* (49) 7165-7166.
 51. Ocampo, R., Dolbier, W. R., Abboud, K. A., Zuluaga, F. Catalyzed Reformatsky reactions with ethyl bromofluoroacetate for the synthesis of α -fluoro- β -hydroxy acids. *J Org Chem* **2002**, *67* (1) 72-78.

Chapter 4: *Investigating the aldolase-catalyzed synthesis of fluorinated sugars from fluoroketones*

Portions of this work were performed in collaboration with the following persons:

Matthew Kojima performed cloning and purification of aldolases. Gabby E. Dolgonos performed cloning and purification of glycerol dehydrogenase. Both Matthew Kojima and Gabby E. Dolgonos characterized enzymes and developed enzymatic reactions for alcohol oxidation and aldol addition.

4.1. Introduction

The work described up to this point has focused on the use of Type II pyruvate aldolases to synthesize chiral organofluorines with carboxylic acid motifs. In this chapter, we pivot to describe the synthesis of fluorosugars using two privileged Type I aldolases that can utilize simple uncharged aldehydes and ketones as a donor substrate. This greatly broadens the aldolase product space to value-added aldehydes and ketones, including rare sugars, without the need for additional dephosphorylation or reduction steps in the case of DHAP-dependent and pyruvate-dependent aldolases respectively.

Fluorinated sugars have a plethora of applications in pharmaceutical design, chemical biology, and clinical diagnostics (*Figure 4.1A*). These applications take advantage of the fact that fluorine substitution can significantly modulate carbohydrate properties such as torsional configuration, hydrolytic stability at glycosidic bonds, lipophilicity, and non-covalent interactions protein residues [1,2]. Many nucleoside-mimicking drugs contain fluorination at the sugar moiety, with the antiviral agent sofosbuvir and anticancer agents clofarabine and gemcitabine being prominent examples [3,4]. Fluorinated sugars serve as inhibitors and mechanistic probes of protein targets implicated in oncology and immunology, such as viral neuraminidases, fucosyltransferases, and glycosidases [5-8]. Due to the high sensitivity of ^{19}F in nuclear resonance, fluorinated sugars are also used in chemical biology to study carbohydrate transporters and carbohydrate-binding proteins via NMR techniques [9,10]. The radioisotope ^{18}F with a half-life of 110 minutes, most commonly in the form of ^{18}F -fluorodeoxyglucose, is widely used as a tracer for positron emission tomography (PET) in medical imaging [11]. As such, the diverse areas in which fluorinated sugars are required fuel the development of methods for synthesizing them with good regioselectivity and stereoselectivity.

Aldolase enzymes are inherently well-suited towards the biocatalytic synthesis of rare sugars, as they generally possess native functions in carbohydrate-related pathways. The field of aldolase-catalyzed sugar synthesis has been dominated by DHAP-dependent aldolases, since these enzymes are well-characterized and diverse, providing flexibility in substrate scope and stereoselectivity [12]. However, reaction scalability can be limited by the expensive nature of DHAP and the need to dephosphorylate the product to obtain ketose sugar [13]. To prepare fluorosugars with a new fluorine stereocenter, 1-fluoro-3-hydroxyacetone phosphate (FHAP) has been investigated as a novel donor for DHAP-dependent aldolases [14]. However, the only known synthesis of FHAP is lengthy, and its enzymatic incorporation was met with limited success, due to the importance of the hydroxy group of DHAP for donor binding. Thus instead of using phosphorylated donors, recent attention has turned towards the exceptional Type I aldolases deoxyribose-5-phosphate aldolase (DERA) and fructose-6-phosphate aldolase (FSA). These two enzymes and their sequence homologs have the unique ability to utilize simple aldehydes and ketones as donor substrates, allowing the formation of aldose or ketose sugars in a single step (*Figure 4.1B*).

DERA natively functions in the metabolism of deoxynucleotides, using acetaldehyde as donor and D-glyceraldehyde-3-phosphate (G3P) as acceptor. Although evolutionarily related to Type I fructose-1,6-bisphosphate aldolases, it has lost the requirement for a phosphorylated donor. Additionally tolerated donors include propionaldehyde, acetone, cyclobutanone, cyclopentanone, glycolaldehyde, and dihydroxyacetone [15,16]. Many aldehydes can replace G3P, in line with the general acceptor promiscuity of aldolases. Sequential aldol additions catalyzed by DERA have been used industrially to synthesize the chiral side-chain of atorvastatin, a blockbuster cholesterol-lowering drug [17]. The enzyme FSA, first discovered in 2001, evolved from photosynthetic

transaldolases and its structure is neither related to DERA nor any canonical family of true aldolases [18]. The native substrates of FSA are dihydroxyacetone and G3P, which parallels the reaction of DERA but with a preference for hydroxylated donors. As such, FSA-catalyzed reactions of glycolaldehyde have been employed for the synthesis of numerous sugars and iminosugars. [19,20]. Engineering of FSA has additionally broadened the donor scope to include non-hydroxylated aldehydes and ketones [21,22].

The following work describes our investigation into fluorosugar synthesis catalyzed by DERA and FSA. The compounds fluoroacetone, 1,3-difluoroacetone, 1-fluoro-3-hydroxyacetone, and fluoroacetaldehyde were tested as fluoro-donors. In the process, we also developed a novel chemoenzymatic synthesis of 1-fluoro-3-hydroxyacetone based upon the selective enzymatic oxidation of 3-fluoro-1,2-propanediol with a glycerol dehydrogenase. The preliminary results detailed in this chapter, showing that DERA and FSA do perform the desired fluoro-aldol reactions in many cases, provides a foundation for future optimization.

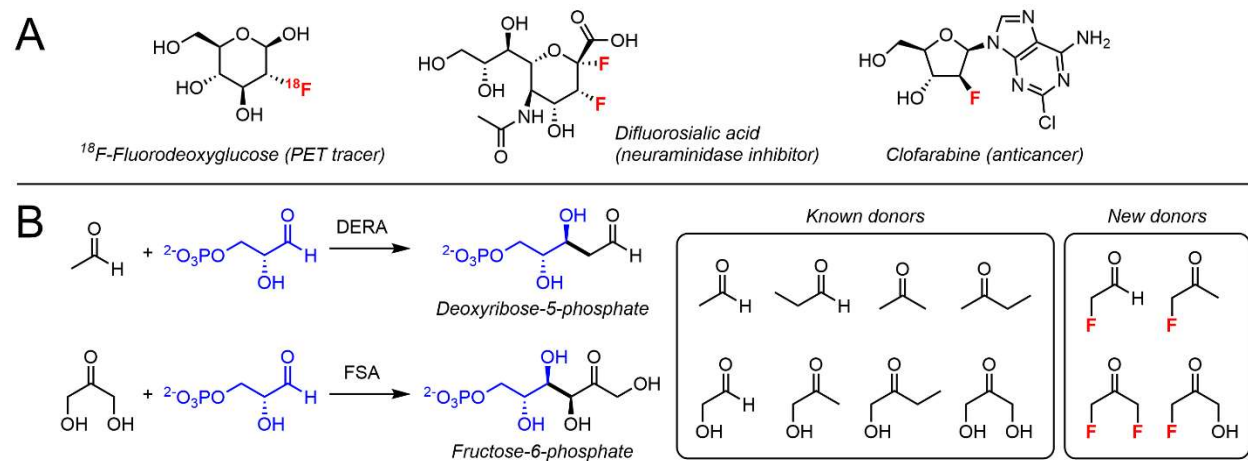


Figure 4.1. Motivation and outline for the aldolase-catalyzed synthesis of fluorinated sugars. A) Examples of fluorosugars or compounds containing fluorosugars with applications in chemical biology and pharmaceuticals. B) Native reactions catalyzed by aldolases DERA and FSA, which can utilize simple ketones and aldehydes as donors. In panels are shown the known donors for either aldolase, and the new fluorinated donors to be investigated.

4.2. Materials and methods

Commercial materials. Acetaldehyde, acetic anhydride, alcohol dehydrogenase from *Saccharomyces cerevisiae* (ADH), ammonium persulfate, cesium fluoride, 1,3-difluoroacetone, dihydroxyacetone dimer, dihydroxyacetone phosphate hemimagnesium salt hydrate, Dowex 50WX8 hydrogen form, epifluorohydrin, fluoroacetone, 2-fluoroethanol, fructose-1,6-bisphosphate trisodium salt hydrate, D-glyceraldehyde, D-glyceraldehyde-3-phosphate solution (G3P, 8-13 mg/mL), glycolaldehyde dimer, hydroxyacetone (90%), DL-1,2-isopropylidenglycerol, L-lactate dehydrogenase from rabbit muscle (LDH), lysozyme, malic dehydrogenase from porcine heart (MDH), 2-mercaptoethanol, methanesulfonyl chloride, methoxyamine hydrochloride, 1-methylimidazole, β -nicotinamide adenine dinucleotide hydrate (NAD⁺), oxaloacetic acid, phenylmethanesulfonyl fluoride (PMSF), polyethyleneimine (PEI), polymyxin B sulfate, pyridine, pyridinium dichromate, sodium carbonate, sodium pyruvate, sodium dodecyl sulfate, TEMED, tetrabutylammonium fluoride, *p*-toluenesulfonyl chloride, and zinc chloride were purchased from Sigma-Aldrich (St. Louis, MO). Acetone, agarose, Bicine,

bromophenol blue, calcium chloride dihydrate, carbenicillin disodium salt, deoxynucleotides (dNTPs), dithiothreitol (DTT), formic acid (LC-MS), glycerol, glycine, HEPES, imidazole, O'GeneRuler 1 kb Plus DNA Ladder, PageRuler Plus Prestained Protein Ladder, PEG 3350, PEG 8000, potassium chloride, sulfuric acid, *tert*-butanol, TCEP, and urea were purchased from Thermo Fisher Scientific (Waltham, MA). LB Miller agar, LB Miller broth, magnesium chloride hexahydrate, and Terrific broth (TB) were from EMD-Millipore (Burlington, MA). InstantBlue Protein Stain was from Expedeon (San Diego, CA). Isopropyl β -D-1-thiogalactopyranoside (IPTG) was from Santa Cruz Biotechnology (Dallas, TX). Acrylamide-bis 30% solution, Bradford reagent, and ethidium bromide were from Bio-Rad (Hercules, CA). Chloroform-*d* and deuterium oxide were from Cambridge Isotope Laboratories (Tewksbury, MA). DNase I, Phusion polymerase, restriction enzymes, Taq ligase, and T5 exonuclease were from New England Biolabs (Ipswich, MA). Ni-NTA agarose resin and DNA purification kits were from Qiagen (Redwood City, CA). Oligonucleotides and gBlocks gene fragments were synthesized by Integrated DNA Technologies (Coralville, IA). All chemicals were used as purchased without further purification.

Bacterial strains. *Escherichia coli* DH10B-T1^R was used for plasmid construction and *E. coli* BL21(DE3)- T1^R was used for protein production. The strains were made chemically competent by the method of [23].

Construction of expression plasmids. Plasmids were constructed by Gibson isothermal assembly and verified by sequencing (Quintara Biosciences; South San Francisco, CA and Genewiz; South Plainfield, NJ). *In silico* sequence alignments were performed on Benchling (Benchling, San Francisco, CA). All DNA sequences used in this study are listed in *Appendix 3*. Plasmids encoding N-terminally His₁₀-tagged enzymes were derived from pET16hp-IMDH which was linearized with NdeI/BamHI. The vector fragments were purified by extraction from 1% agarose gel. The genes *deoC*, *fsaA*, *fbaA*, *tpiA*, and *gldA* from *E. coli* were amplified by polymerase chain reaction (PCR) from the genome of *E. coli* DH10B-T1^R. Primers were reconstituted to 100 μ M in water. The 50 μ L PCR mixtures contained 5 \times Phusion HF buffer, 0.2 mM dNTPs, 1 μ L cell suspension (a single colony in 50 μ L water), 2 units Phusion HF polymerase, and 0.5 μ M each primer. Thermal cycling was as follows: 3 min at 98 $^{\circ}$ C, followed by 35 cycles of 30 s at 98 $^{\circ}$ C, 30 s at (T_m -5) $^{\circ}$ C, and 1 min at 72 $^{\circ}$ C, followed by a final extension of 5 min at 72 $^{\circ}$ C. A 5 μ L sample was visualized on 1% agarose gel and the remaining material was purified by spin column.

Purified vectors and inserts were combined (9:1 v/v) in 5 μ L volume, to which 15 μ L of Gibson master mix, prepared according to [24], was added. The mixture was incubated in a thermocycler at 50 $^{\circ}$ C for 1 h, then added to chemically competent *E. coli* DH10B-T1^R along with 20 μ L of 5 \times KCM solution (0.5 M KCl, 0.15 M CaCl₂, 0.25 M MgCl₂) and water to 200 μ L total. After incubation on ice for 20 min, the cells were heat-shocked at 42 $^{\circ}$ C for 90 s and returned to ice for 2 min. Cells were diluted with 1 mL of LB, recovered with shaking at 37 $^{\circ}$ C for 45 min, and pelleted by centrifugation (7,000 $\times g$, 1 min). The supernatant was partially removed (0.8 mL), then the cells were resuspended and plated on LB agar with 50 μ g/mL carbenicillin. Plates were incubated at 37 $^{\circ}$ C overnight. Liquid cultures (5 mL LB with 50 μ g/mL carbenicillin) were inoculated with single colonies and grown at 37 $^{\circ}$ C overnight, following which the plasmids were isolated and confirmed by sequencing from the T7 promoter and T7 terminator.

Point mutations in the plasmid pET16hp-DeoC were introduced by the two-step modification of the Quikchange site-directed mutagenesis method [25]. Mutagenic primers (NNK codon was used to obtain most variants of residue Cys-47, with targeted primers for other residues) were reconstituted to 100 μ M in water. Single primer extension reactions (50 μ L) were prepared

containing 5X Phusion HF buffer, 0.2 mM dNTPs, 20 ng template, 1 unit Phusion HF polymerase, and 0.5 μ M forward or reverse primer. Thermal cycling was as follows: 3 min at 98 °C, followed by 10 cycles of 30 s at 98 °C, 30 s at 65 °C, and 3 min at 72 °C, followed by a final extension of 5 min at 72 °C. Then, 25 μ L of each primer extension reaction was combined and supplemented with 1 unit Phusion HF polymerase. Thermal cycling was continued for 25 cycles. DpnI (2 units) was added to digest the template for 1 h at 37 °C. *E. coli* DH10B-T1^R was transformed with 5 μ L of the mixture (KCM method), then plated on LB agar with 50 μ g/mL carbenicillin and incubated at 37 °C overnight. Liquid cultures were inoculated with single colonies and grown at 37 °C overnight. The plasmids were isolated and mutations confirmed by sequencing.

Expression and purification of His-tagged proteins. Expression plasmids (10 ng) were transformed into *E. coli* BL21(DE3)-T1^R (KCM method) and a single colony was used to inoculate 25 mL of TB media with 50 μ g/mL carbenicillin. The seed culture was grown overnight with shaking at 37 °C (200 RPM). Then, 1 L of TB in an Ultra Yield baffled flask (Thomson Instrument Company; Oceanside, CA) was inoculated with the seed culture and growth was continued to OD600 of 0.8-1.2. For zinc-dependent proteins FbaA and GldA, the culture was supplemented with ZnCl₂ (0.3 mM) before inoculation. The culture was chilled in an ice bath for 15 min and protein expression was induced with IPTG (1 mM). Expression proceeded at 16 °C overnight, after which the cells were harvested (7,000 \times g, 5 min, 4 °C) and either stored at -80 °C or subjected to protein purification immediately.

Harvested cells were resuspended in 5 mL/(g wet cell wt.) of lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% v/v glycerol). To the cells was added PMSF (1 mM) and lysozyme (1 mg/mL). After incubation at r.t. for 30 min, the sample was sonicated (QSonica Q700) with the following program: 10 s on, 20 s off, 1 min total process time, amplitude 50. Cell debris was removed by centrifugation (15,000 \times g, 20 min, 4 °C). To the supernatant was added PEI (0.05% w/v) and the precipitated nucleic acids were removed by centrifugation (15,000 \times g, 20 min, 4 °C). Ni-NTA agarose resin (50% suspension in 20% EtOH, 0.2 mL/(g wet cell wt.)) was added to the samples, which were shaken gently on an orbital shaker at r.t. for 30 min and then poured into a glass column. The resin was washed with >20 column volumes (CV) of wash buffer (same as lysis buffer with 20 mM imidazole) until the effluent tested negative by the Bradford dye-binding assay. Protein was eluted with >5 CV of elution buffer (same as lysis buffer with 250 mM imidazole) until the effluent tested negative again. The elution fraction was supplemented with β -ME (0.1% v/v). For FbaA and GldA, ZnCl₂ (0.3 mM) was also added. Fractions and purified proteins were analyzed by SDS-PAGE.

The proteins were concentrated in Amicon Ultra-15 centrifugal filter units (EMD Millipore; Burlington, MA) with 10 kDa molecular weight cutoff. Note that GldA has poor solubility at 4 °C, and should be concentrated at room temperature. Buffer exchange was conducted with PD-10 columns (GE Healthcare; Chicago, IL) into HEPES storage buffer (20 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 10% v/v glycerol, 1 mM TCEP). GldA was sequentially desalted three times using HEPES storage buffer without glycerol, in order to achieve full removal of glycerol from the sample. Desalted FbaA and GldA were again supplemented with ZnCl₂ (0.3 mM). The proteins were finally concentrated to the following target values: 250 μ M DeoC, 50 μ M FsaA, 250 μ M FbaA, 250 μ M TpiA, 200 μ M GldA. Protein concentrations were determined by the absorbance at 280 nm, using extinction coefficients DeoC 14440 M⁻¹cm⁻¹, FsaA 25946 M⁻¹cm⁻¹, FbaA 42096 M⁻¹cm⁻¹, TpiA 29553 M⁻¹cm⁻¹, GldA 41661 M⁻¹cm⁻¹ predicted with ExpASY ProtParam (Swiss Institute of Bioinformatics).

Parallel expression and purification of DERA variants. The following procedure was used for simultaneous, medium-scale preparation of multiple protein variants. Expression plasmids (10 ng) were transformed into *E. coli* BL21(DE3)-T1^R (KCM method) and a single colony was used to inoculate 5 mL of TB media with 50 µg/mL carbenicillin. The seed culture was grown overnight with shaking at 37 °C (200 RPM). Then, 100 mL of TB was inoculated with the seed culture and growth was continued for 2 h. Protein expression was induced with IPTG (1 mM). Expression proceeded at 16 °C overnight, after which the cells were harvested (7,000 × g, 5 min, 4 °C) and either stored at -80 °C or subjected to protein purification immediately.

Enzymatic lysis buffer was prepared by supplementing lysis buffer from the previous section with 1 mg/mL lysozyme, 1 mM PMSF, 50 µg/mL polymyxin B sulfate, 1 U/mL DNaseI, 5 mM MgCl₂, and 1 mM CaCl₂. Harvested cells were resuspended in 10 mL of enzymatic lysis buffer and incubated on a shaker for 1 h. Cell debris was removed by centrifugation (15,000 × g, 20 min, 4 °C). Ni-NTA agarose resin (50% suspension in 20% EtOH, 1 mL) was added to the samples, which were shaken gently on an orbital shaker at r.t. for 30 min. The resin was collected by centrifugation (5000 × g, 1 min, 4 °C) and the supernatant removed. The resin was washed with wash buffer (3 × 10 mL, same as lysis buffer with 20 mM imidazole) and then eluted with elution buffer (2 × 1.25 mL, same as lysis buffer with 250 mM imidazole). Each time the resin was collected by centrifugation, and the wash fraction discarded or the elution fraction transferred to a separate tube. The combined elution fractions (2.5 mL) were desalted with PD-10 columns (GE Healthcare; Chicago, IL) into HEPES storage buffer (20 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 10% v/v glycerol, 1 mM TCEP) and then concentrated in Amicon Ultra-4 centrifugal filter units (EMD Millipore; Burlington, MA) with 10 kDa molecular weight cutoff. DERA variants were concentrated to a target value of 250 µM. Protein concentrations were determined by the absorbance at 280 nm, using extinction coefficient 14440 M⁻¹cm⁻¹ predicted with ExPASy ProtParam (Swiss Institute of Bioinformatics).

Safety note on handling of organofluorine compounds. 2-Fluoroethanol used in the synthesis of fluoroacetaldehyde is an extremely toxic (due to metabolism to fluoroacetate) and volatile liquid. It must be used only in the fume hood with protective equipment and utmost caution. Fluoroacetone, 1,3-difluoroacetone, and epifluorohydrin are toxic and volatile liquids that should be handled in the fume hood. Dilute aqueous solutions of fluoroacetaldehyde and 1-fluoro-3-hydroxyacetone should be handled with care. Other organofluorines including synthetic intermediates towards 1-fluoro-3-hydroxyacetone and novel synthesized fluorosugars may have uncharacterized toxicity and warrant general handling precautions.

General remarks for 1D-NMR spectroscopy. NMR was performed at the UC Berkeley College of Chemistry NMR Facility. Characterization of isolated compounds (¹H and ¹⁹F NMR) and quantitative ¹⁹F NMR assays were performed on a Bruker AV-600 (600 MHz) equipped with a Z-gradient broadband cryoprobe (supported by NIH S10OD024998). All experiments were conducted at 298 K. Chemical shifts were referenced to the solvent and expressed relative to tetramethylsilane (¹H) or trichlorofluoromethane (¹⁹F). The spectral center of ¹⁹F NMR was -150 ppm and the spectral width was 300 ppm. For characterization of isolated compounds, the number of ¹H or ¹⁹F scans was 8. Scan parameters for different ¹⁹F NMR assays are described in their respective sections. Data was processed in MestreNova (MestreLab Research). Baseline distortion in the -150 to -200 ppm region of ¹⁹F NMR due to fluoropolymer in the probes was corrected with backward linear prediction (Toeplitz) of the first 128 data points, followed by manual phase correction and baseline correction (Whittaker smoother).

Synthesis of fluoroacetaldehyde. 2-Fluoroethanol (0.15 mL, 2.55 mmol) was dissolved in CH₂Cl₂ (2 mL). Pyridinium dichromate (1.0 g, 2.66 mmol) was added and the reaction stirred at r.t. overnight. Using a water-jacketed short-path distillation apparatus, the temperature was gradually raised to 80 °C to distill the liquid portion to dryness, collecting the distillate in a receiving flask containing water (1 mL). The aqueous layer was separated and an aliquot (0.1 mL) was added to D₂O (0.6 mL) with 5-fluorouracil (1.0 mg) for ¹⁹F NMR quantification (number of scans: 8, recycle delay: 30 s). The typical concentrations of 2-fluoroethanol and fluoroacetaldehyde were 30-50 mM and 5-15 mM respectively. ¹⁹F NMR (565 MHz, D₂O) 5-Fluorouracil δ -169.4 (d, *J* = 5.3 Hz); 2-Fluoroethanol δ -224.6 (tt, *J* = 47.4, 31.9 Hz); Fluoroacetaldehyde δ -231.3 (td, *J* = 46.6, 9.9 Hz).

Synthesis of 3-fluoro-1,2-propanediol (FPDO) from epifluorohydrin. Epifluorohydrin (152 mg, 143 μL, 2 mmol) was suspended in water (0.5 mL). Concentrated sulfuric acid (2 μL) was added and the mixture stirred overnight. Aliquots of the reaction (15 μL) were added to D₂O (0.7 mL) for measurement of conversion by ¹⁹F NMR (single scan). After 2 h, the fluorinated species consisted of 22% product and 78% starting material. After 24 h, the fluorinated species consisted of 96% product and 4% of an unidentified minor impurity. The overnight reaction was neutralized with 1 M NaOH and water was added to a total volume of 1 mL for a 2 M solution of FPDO. ¹⁹F NMR (565 MHz, D₂O) Epifluorohydrin δ -228.5 (td, *J* = 47.5, 14.1 Hz); FPDO δ -232.7 (td, *J* = 46.9, 22.2 Hz); Unidentified impurity δ -232.0 (td, *J* = 45.4, 21.9 Hz).

Synthesis of 3-fluoro-1,2-propanediol (FPDO) from solketal.

Solketal mesylate. Solketal (1.5 mL, 12.1 mmol) was dissolved in dichloromethane (15 mL). Pyridine (2 mL, 24.8 mmol) and MsCl (1 mL, 12.9 mmol) were added and the reaction stirred overnight. The mixture was diluted with EtOAc (60 mL) and washed with 5% HCl (30 mL) and 5% sodium bicarbonate (30 mL). The combined aqueous layers were back-extracted with EtOAc (2 × 30 mL). The combined organic layers were washed with brine (30 mL), dried over sodium sulfate, and evaporated. The residue was purified by flash chromatography (2:1 hexanes/EtOAc) to give the product as a clear oil (2.38 g, 94%). ¹H NMR (600 MHz, CDCl₃) δ 4.42-4.38 (m, 1H), 4.26-4.24 (m, 2H), 4.13 (dd, *J* = 8.7, 6.5 Hz, 1H), 3.85 (dd, *J* = 8.7, 5.5 Hz, 1H), 3.09 (s, 3H), 1.47 (s, 3H), 1.39 (s, 3H).

3-Fluoro-1,2-propanediol (FPDO). Solketal mesylate (0.5 g, 2.38 mmol) was dissolved in *t*-BuOH (7 mL). TBAF (1 M in THF, 3.5 mL) was added, and the reaction was heated to 80 °C and stirred overnight. The mixture was diluted with diethyl ether (40 mL), washed with water (2 × 20 mL) and brine (20 mL), and dried over sodium sulfate. The organic layer was concentrated without heating, then methanol (20 mL) and Dowex-50W (500 mg) were added and the mixture stirred overnight. After evaporation, the residue was purified by flash chromatography (3:1 DCM/EtOAc) to give the product as a clear oil (34 mg, 15%). ¹H NMR (600 MHz, CDCl₃) δ 4.59-4.44 (m, 2H), 4.01 (dt, *J* = 18.2, 5.8, 4.0 Hz, 1H), 3.80-3.69 (m, 2H). ¹⁹F NMR (565 MHz, CDCl₃) δ -232.7 (td, *J* = 47.1, 18.2 Hz).

Enzymatic oxidation of FPDO to FHA. Reactions (0.5 mL) contained 100 mM FPDO (25 μM of 2 M), 30/60/120 mM cosubstrate (50 μM of 0.3/0.6/1.2 M), 10 mM NAD⁺ (50 μL of 100 mM), 25 μg recycling enzyme (25 μL of 1 mg/mL), 100 μM GldA (250 μL of 200 μM), and 100

μL of 20 mM HEPES pH 7.5 with 1.5 mM ZnCl_2 (final concentration 0.3 mM of ZnCl_2). Combinations of co-substrate and recycling enzyme were pyruvate/lactate dehydrogenase (LDH), oxaloacetate/malic dehydrogenase (MDH), and acetaldehyde/alcohol dehydrogenase (ADH). The stock solution of oxaloacetic acid was neutralized carefully with 3 M NaOH prior to use. The reactions were incubated at r.t. for 48 h, then particulates were removed by centrifugation ($20,000 \times g$, 5 min, 4 °C) and the samples transferred to NMR tubes with a glass capillary of D_2O (100 μL) for quantitative ^{19}F NMR analysis (single scan). ^{19}F NMR (565 MHz, $\text{H}_2\text{O}+\text{D}_2\text{O}$) FHA hydrate δ -230.3 (t, $J = 47.0$ Hz); FHA keto δ -236.5 (t, $J = 46.0$ Hz); fluoride ion elimination δ -119.4 (s).

For reactions conducted with pyruvate/LDH at higher pH, the reaction buffer was replaced with 50 mM Bicine pH 8.5 or 50 mM carbonate pH 9.5, and GldA enzyme was exchanged with PD-10 columns into corresponding storage buffers without glycerol. Control reactions at pH 7.5, 8.5, and 9.5 with the volume of GldA replaced by corresponding storage buffers showed no product formation. Reactions at pH 7.5 containing 5 mM FPDO/5 μM GldA or 100 mM FPDO/100 μM GldA but without cosubstrate or recycling enzyme showed no product formation. For most of the oxidation reactions, FPDO was synthesized from epifluorohydrin. However, FPDO synthesized from solketal was also verified to be a competent substrate and showed identical results when tested in the pH 7.5, pyruvate/LDH conditions.

Initial rate assay of GldA. Reactions (1 mL) contained 100 mM alcohol substrate, 1 mM NAD^+ , 1 or 5 μM GldA, and 0.3 mM ZnCl_2 . Three substrates were used: glycerol, 1,2-propanediol (PDO), or 3-fluoro-1,2-propanediol (FPDO). Three buffer conditions were used: 20 mM HEPES pH 7.5, 50 mM Bicine pH 8.5, or 50 mM carbonate pH 9.5. Control reactions were performed at each pH condition without substrate. Reactions were initiated by addition of GldA and the rate of change in absorbance at 340 nm was monitored for 30 s. Estimates of the turnover rate k_{cat} were calculated using the 340 nm extinction coefficient of $6220 \text{ M}^{-1}\text{cm}^{-1}$ for NADH.

Aldol assay and derivatization-LC-MS analysis of sugar phosphates. Reactions (100 μL) contained 1 mM donor substrate, 0.5 mM G3P (1 μL of ~ 50 mM commercial solution), 5 μM DERA or FSA (10 μL of 50 μM), and 20 mM HEPES pH 7.5 (remaining volume). Non-fluorinated donors used as positive controls were acetaldehyde (ALD), glycolaldehyde (hydroxyacetaldehyde, HAL), acetone (ACE), hydroxyacetone (monohydroxyacetone, MHA), and dihydroxyacetone (DHA). Novel fluorinated donors tested were fluoroacetone (monofluoroacetone, MFA), 1,3-difluoroacetone (DFA), fluoroacetaldehyde (FAL), and 1-fluoro-3-hydroxyacetone (FHA). (1 μL of 100 mM). Most donor substrates were applied by adding 1 μL of a 100 mM stock solution. FAL was applied by adding 15 μL of synthesized aqueous solution (containing 5~10 mM FAL). FHA was applied by adding 15 μL of GldA reaction (pH 7.5, pyruvate/LDH condition, containing 5~10 mM FHA) that had been spin-filtered (3000 kDa molecular weight cutoff) to remove proteins.

The reaction samples were incubated at r.t. for 1 h, then frozen at -80 °C and lyophilized to dryness in a vacuum concentrator. Methoxyamine reagent (18 μL , 2% methoxyamine hydrochloride in pyridine) was added and the samples incubated at r.t. overnight. Acetic anhydride (12 μL) and 1-methylimidazole (6 μM) were added and the samples incubated at 37 °C for 1 h, then evaporated to dryness in a vacuum concentrator. The residues were dissolved in 0.1% formic acid (1 mL) and clarified by centrifugation ($20,000 \times g$, 5 min, 4 °C) for high resolution LC-MS analysis of derivatized sugar phosphates. An Agilent 1290 Infinity LC system was equipped with an Agilent Poroshell 120 EC-C18 column (2.7 μm , 2.1 \times 50 mm) and coupled to an Agilent 6530 Q-TOF MS operated in positive full-scan mode (capillary voltage 3500 V, nozzle voltage 500 V, fragmentor voltage 175 V, scan range 100-1700 m/z). Reverse-phase chromatography conditions

were as follows. Solvent A: 0.1% v/v formic acid; solvent B: acetonitrile; flow rate 0.6 mL/min; maximum pressure 600 bar; timetable (%B): 0% at 0 min, 0% at 1 min, 100% at 6 min, 100% at 6.5 min, 0% at 7.5 min.

Aldol assay of fluorinated donors for ^{19}F NMR analysis. For qualitative assessment of fluorosugar formation, reactions (0.5 mL) contained 5 mM donor substrate (MFA, DFA, or FAL), 2.5 mM G3P, 25 μM DERA, and 20 mM HEPES pH 7.5. Unsuccessful reactions attempting to use D-glyceraldehyde as acceptor contained 10 mM MFA or DFA, 20 mM D-glyceraldehyde, 100 μM DERA, and 20 mM HEPES pH 7.5. The reactions were incubated at r.t. for 24 h, then particulates were removed by centrifugation ($20,000 \times g$, 5 min, 4 $^{\circ}\text{C}$) and the samples transferred to NMR tubes with a glass capillary of D_2O (100 μL) for ^{19}F NMR analysis (number of scans: 32, recycle delay: 1 s). This assay was also used to assess the relative activity to wild-type of the DERA variants C47A, C47D, C47S, C47V, L20A, V73A, F76A, and I139A for aldol addition of the donors MFA and DFA.

For quantitative analysis of fluorosugar formation, reactions (0.5 mL) contained 20 mM MFA or DFA, 20 mM fructose-1,6-bisphosphate, 100 μM DERA, 2.5 μM triose phosphate isomerase (TpiA), 2.5 μM fructose-1,6-bisphosphate aldolase (FbaA), 0.3 mM ZnCl_2 , and 20 mM HEPES pH 7.5. The reactions were incubated at r.t. for 24 h, then particulates were removed by centrifugation ($20,000 \times g$, 5 min, 4 $^{\circ}\text{C}$) and the samples transferred to NMR tubes with a glass capillary of D_2O (100 μL) for ^{19}F NMR analysis (number of scans: 8, recycle delay: 30 s). This assay was used to accurately compare the extent of conversion for MFA and DFA aldol reactions catalyzed by different variants of DERA.

4.3. Results and discussion

To embark on our investigations of aldolase-catalyzed synthesis of fluorinated sugars, we began with cloning, expression, and purification of deoxyribose-5-phosphate aldolase (DERA, *deoC* gene) and fructose-6-phosphate aldolase (FSA, *fsaA* gene) from *E. coli*. Both aldolases have well-characterized biochemical properties and extensive studies of biocatalytic potential due to substrate promiscuity [15-22]. From the known behavior of these two functionally similar but structurally unrelated aldolases, we hypothesized that DERA would prefer donor substrates without an α -hydroxy group such as aliphatic aldehydes and ketones, while FSA would prefer donor substrates with an α -hydroxy group. In the case of asymmetrical ketones that have two nucleophilic sites, the regioselectivity of aldol addition might then be controlled by the choice of enzyme. We were most excited by the prospect of regioselective attack from a fluorinated α -carbon, as this would create a new and valuable fluorine chiral center. It would also be the most difficult regioselectivity to accomplish according to electronic principles, due to the high electronegativity of fluorine and resulting low nucleophilicity of the α -carbon.

With the privileged aldolases DERA and FSA in hand, we desired to procure four simple fluorinated carbonyl compounds to test as novel aldolase donors. These included fluoroacetone (monofluoroacetone, MFA), 1,3-difluoroacetone (DFA), fluoroacetaldehyde (FAL), and 1-fluoro-3-hydroxyacetone (FHA). MFA and DFA are commercially available as pure liquids, which has facilitated the prior study of these potential substrates. In the seminal study of DERA as a biocatalyst, MFA was mentioned as a donor substrate but the product characterization including regioselectivity was unclear [15]. DFA has only been studied as an acceptor substrate for aldolases, for which its high electrophilicity is advantageous [26]. FAL is a known fluorinated natural product

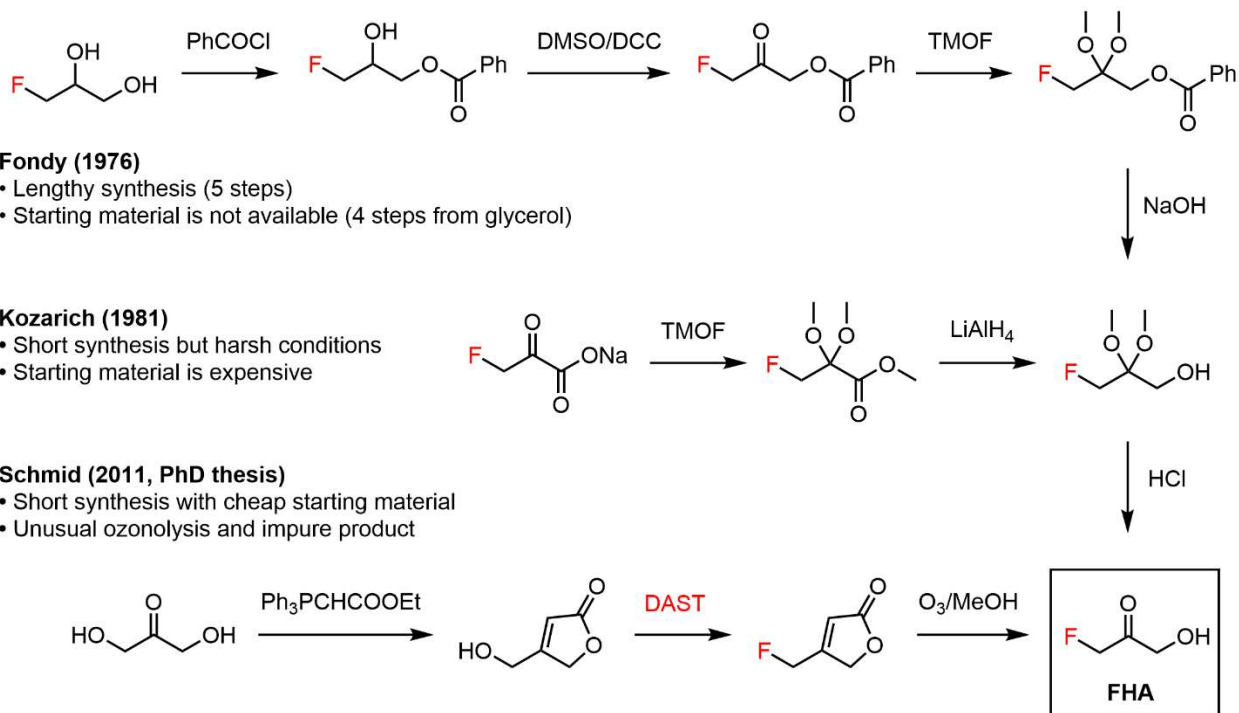


Figure 4.2. Previous synthetic routes towards 1-fluoro-3-hydroxyacetone (FHA), and the advantages and drawbacks of each method. FHA cannot be isolated but can be obtained in solution by deprotection of a suitable precursor.

and the immediate precursor to toxic defense molecules fluoroacetate and fluorothreonine in organofluorine biosynthesis pathways [27,28]. In fluorothreonine biosynthesis, FAL acts as the acceptor substrate of a PLP-dependent transaldolase, and therefore similarly to DFA it has not been studied as a potential donor substrate. Due to the high volatility and toxicity of FAL, it is generally not isolated. Instead, a dilute aqueous solution is prepared from 2-fluoroethanol via oxidation with pyridinium dichromate followed by distillation of all volatiles into a receiving flask containing water, thus trapping FAL as its aldehyde hydrate [29]. Following this procedure provided aqueous solutions containing 5-15 mM FAL along with 30-50 mM 2-fluoroethanol. Such mixtures have been used in enzymatic reactions, as 2-fluoroethanol is generally inert towards protein structure [27].

The asymmetrical fluoroketone FHA is especially interesting as it would provide the most straightforward synthesis of aldoses and ketoses with a single OH to F substitution. FHA is known to be unstable in isolation and has therefore been synthesized in the literature by deprotection in solution of a suitable precursor (*Figure 4.2*). The first synthesis was accomplished in five steps from 3-fluoro-1,2-propanediol (FPDO), but this route is lengthy and the starting material is not commercially available [30]. Another early route prepared FHA in only three steps from fluoropyruvate, however, this starting material is expensive and the harsh reaction conditions were not reproducible in our hands [31]. Both of these routes deprotect FHA dimethyl ketal as the final step to produce a clean solution of the desired product. In contrast, the most recent synthesis of FHA used dihydroxyacetone as an inexpensive starting material and an innovative ozonolysis step to liberate FHA from the precursor [32]. However, due to difficulties with ozonolysis-derived impurities, the FHA was not implemented in enzymatic reactions.

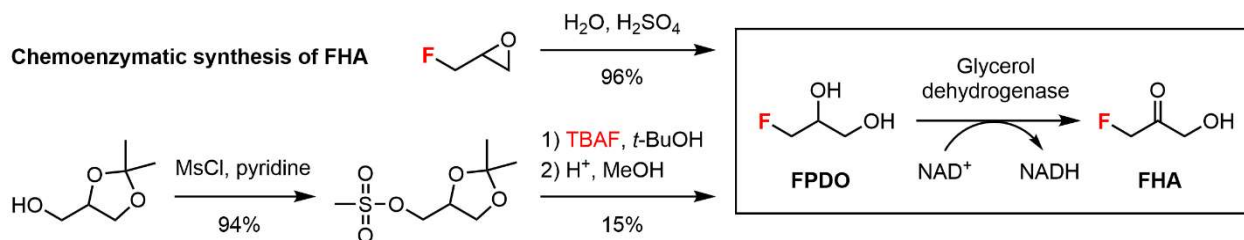


Figure 4.3. Novel method to obtain 1-fluoro-3-hydroxyacetone (FHA) by enzymatic oxidation of 3-fluoro-1,2-propanediol (FPDO) with a glycerol dehydrogenase. Two synthetic routes for obtaining the FPDO substrate are shown, utilizing epifluorohydrin or solketal as a precursor.

Seeking a more amenable preparation of FHA, we hypothesized that a selective enzymatic oxidation of 3-fluoro-1,2-propanediol (FPDO) with a glycerol dehydrogenase could provide FHA in a single step (Figure 4.3). The zinc-dependent glycerol dehydrogenase GldA from *E. coli* natively performs a redox reaction between glycerol and dihydroxyacetone using NAD^+/NADH cofactors, which was similar to our desired reaction aside from a single OH to F substitution [33,34]. GldA was thus cloned, expressed, and purified for investigation (Figure 4.4). Glycerol was thoroughly removed from GldA samples by multiple rounds of buffer exchange into storage buffer without glycerol. For the oxidation substrate, FPDO was mainly synthesized from the acidic hydrolysis of epifluorohydrin [35]. Although this method is very simple and efficient, the expensive nature of epifluorohydrin also led us to develop a synthesis of FPDO from solketal (isopropylidene glycerol). Conversion of solketal to its mesylate was followed by hydrogen-bond promoted nucleophilic fluorination with TBAF, and deprotection of the acetonide to furnish FPDO [36]. In our optimizations of this sequence, TBAF was found to be superior to cesium fluoride which provided little product, and solketal mesylate superior to solketal tosylate due to difficulties in separating an impurity of glycerol 1-tosylate from the FPDO product.

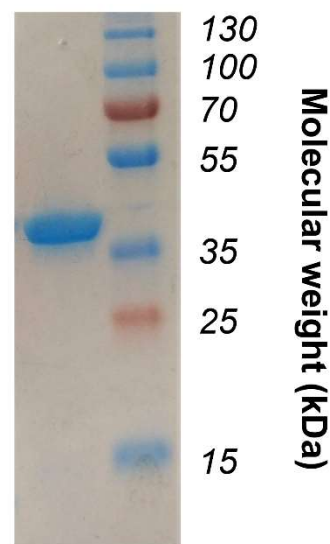


Figure 4.4. SDS-PAGE analysis of glycerol dehydrogenase GldA.

FPDO (100 mM) was incubated with GldA (0.1 mol%), NAD^+ (10 mM), a co-substrate to serve as terminal oxidant (30-120 mM), and a cofactor recycling enzyme for 48 h to then assess the extent of oxidation by ^{19}F NMR (Figure 4.5). Three systems of co-substrate and recycling enzyme were compared: pyruvate/lactate dehydrogenase (LDH), oxaloacetate/malic dehydrogenase (MDH), and acetaldehyde/alcohol dehydrogenase (ADH). At assay pH 7.5 (HEPES buffer), the LDH condition achieved 7-10% conversion of FPDO to FHA, along with 3% fluoride elimination. The MDH and ADH conditions resulted in lower or no product formation along with higher fluoride elimination. The observation of fluoride is likely to occur from non-specific oxidation of the primary alcohol in FPDO, leading to a β -fluoro aldehyde that undergoes facile β -elimination. The LDH condition was also tested at pH 8.5 (Bicine buffer) which showed only a slight increase to 10-13% FHA, while pH 9.5 (carbonate buffer) resulted in detriment to FHA production and a large amount of fluoride elimination. Control reactions without GldA at all three pH conditions had neither FHA nor fluoride elimination, showing that GldA is responsible for both products. As the highest conversion to FHA achieved was not significantly greater than

the supplied NAD^+ , it was unclear if the recycling system with co-substrate had taken effect. However, reactions run without co-substrate or recycling enzyme at either 5 mM or 100 mM FPDO also had no product formation, suggesting that the redox equilibrium of FPDO and FHA lies strongly toward the reduced species and requires a coupling enzyme to drive the oxidation.

To probe if the slow rate of oxidation of FPDO to FHA was related to enzyme kinetic properties, the initial rate of oxidation was measured by UV-vis spectroscopy, monitoring the production of NADH as an increase in absorbance at 340 nm (*Figure 4.6*). FPDO was compared to glycerol and 1,2-propanediol (PDO), with all substrates applied at 100 mM. The measurements were also performed at three pH conditions, and we were surprised to find that the initial rate of oxidation at pH 8.5 for all substrates was far higher than at pH 7.5 or 9.5. This was contrary to the pH trend in the extended progress of FPDO oxidation reactions. The observation of a faster rate for PDO than the native substrate glycerol was unexpected, but corroborated by literature [33]. More surprisingly, the initial rate for FPDO was only slightly lower than glycerol. This follow-up experiment suggests that the low oxidation yield of FPDO is not due to kinetic limitations of a fluorine substituent, but due to thermodynamic limitations of GldA that may be solved in the future by thorough screening of reaction conditions or consideration of other enzyme homologs. For the time being, GldA reactions were filtered to remove proteins and provide solutions containing 7-13 mM FHA for use as a potential aldolase substrate.

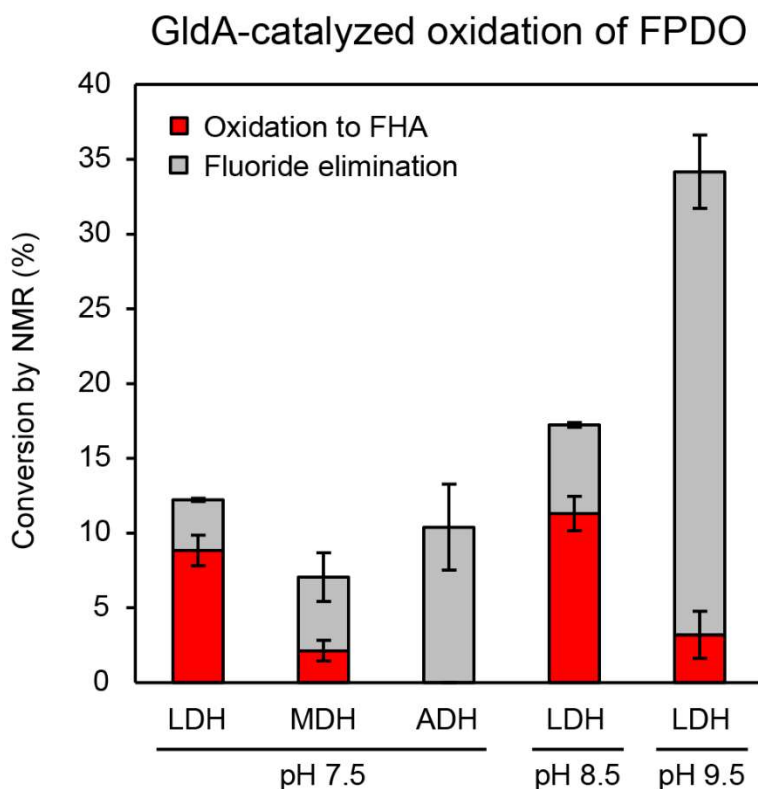


Figure 4.5. Assessment by ^{19}F NMR of glycerol dehydrogenase GldA-catalyzed oxidation of FPDO to FHA. Reactions contained 100 mM FPDO, 10 mM NAD^+ , 100 μM GldA, 0.3 mM ZnCl_2 , 25 μg recycling enzyme, and a co-substrate (pyruvate, oxaloacetate, or acetaldehyde) corresponding to each recycling system. For each condition, the co-substrate was applied at 30, 60, or 120 mM. An average of the results is displayed, as none of the FHA yields significantly exceeded 10 mM and no clear effect of the co-substrate concentration was observed.

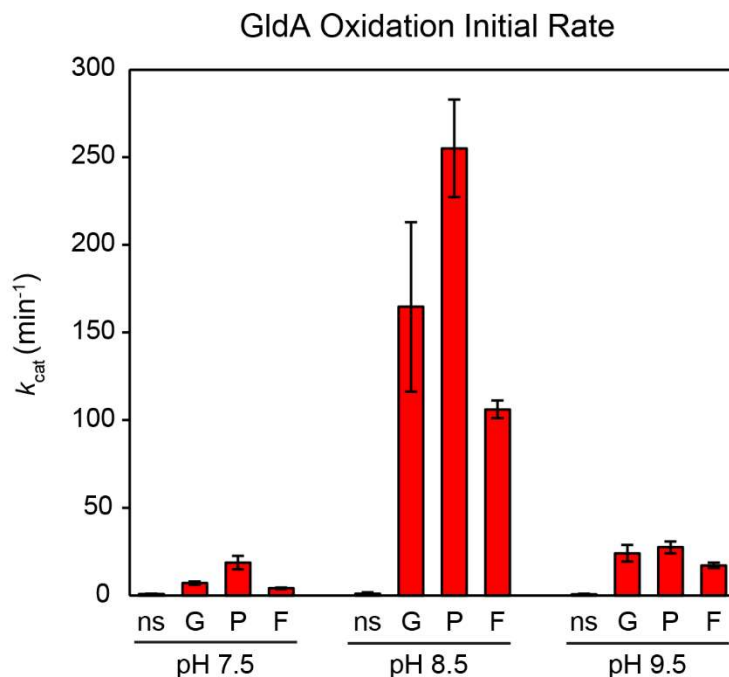


Figure 4.6. Measurement of the oxidation kinetics of glycerol dehydrogenase GldA on different alcohol substrates. (ns) No substrate control, (G) glycerol, (P) 1,2-propanediol, (F) 3-fluoro-1,2-propanediol. The assay was performed at pH 7.5, 8.5, and 9.5 using HEPES, Bicine, and carbonate buffers, respectively.

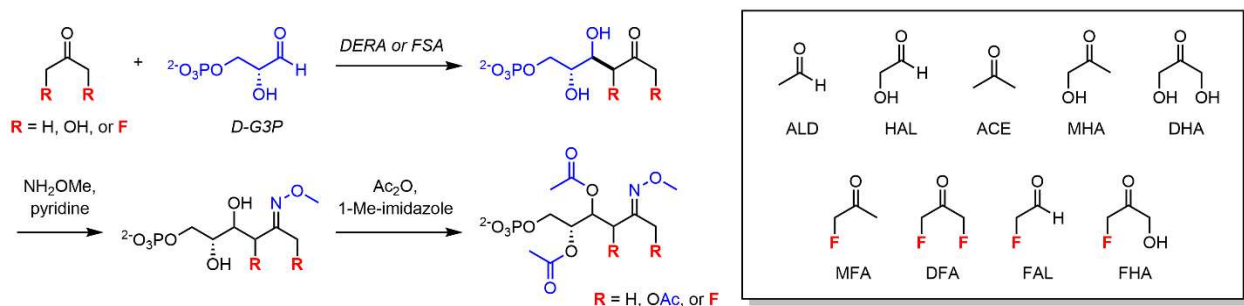


Figure 4.7. Derivatization-LC-MS analysis of fructose-6-phosphate analogs resulting from the aldol addition of various donor substrates. The oxime formation and acylation steps allow for chromatography on a reverse-phase column. Five non-fluorinated control substrates and four fluorinated substrates of interest were used.

Analytical scale aldol activity assays of DERA and FSA were performed in 100 μL volume by incubating a donor substrate (1 mM) with the native acceptor D-glyceraldehyde-3-phosphate (G3P, 0.5 mM) and the aldolase together in HEPES buffer pH 7.5. The solvent was removed and the sugar phosphates subjected to a two-step derivatization for high resolution LC-MS with a reverse-phase column. Using minor modifications of an established protocol [37], ketone or aldehyde moieties were converted to oximes with methoxyamine, and all hydroxy groups were acetylated with acetic anhydride (Figure 4.7). The exact mass of each derivatized sugar phosphate could be extracted from its chromatogram to assess the formation of aldol products (Figure 4.8). Native activity of DERA and FSA was ascertained on non-fluorinated substrates including acetaldehyde (ALD), glycolaldehyde (hydroxyacetaldehyde, HAL), acetone (ACE), hydroxyacetone (monohydroxyacetone, MHA), and dihydroxyacetone (DHA). In accordance with the known

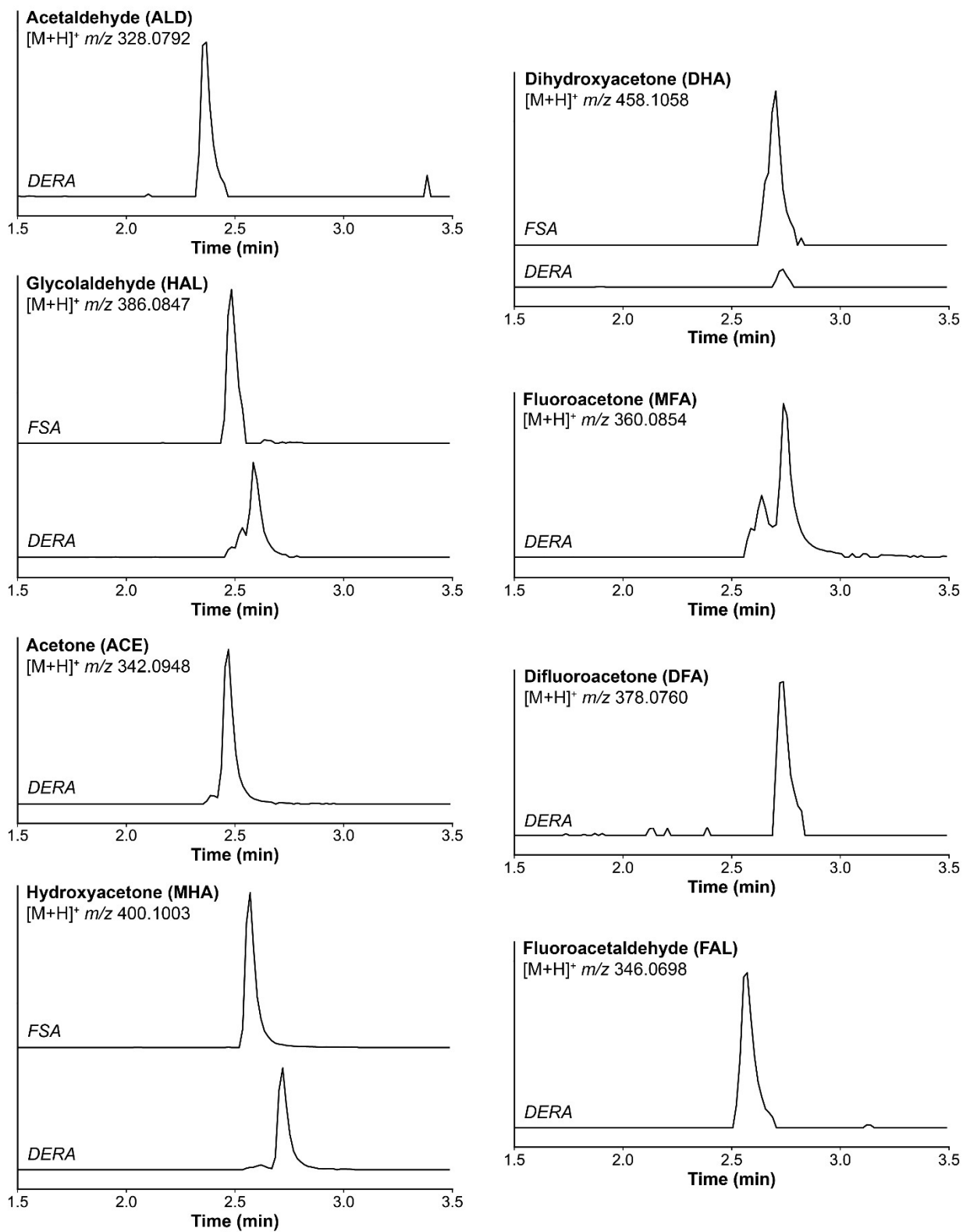


Figure 4.8. Extracted ion chromatograms from the derivatization-LC-MS analysis of aldol reactions. No product peaks were detected in the absence of aldolase. The compounds ALD, ACE, MFA, DFA, and FAL were not substrates for aldolase FSA. The compound 1-fluoro-3-hydroxyacetone (FHA) was not a substrate for aldolases DERA or FSA.

donor preferences, non-hydroxylated donors were utilized by DERA but not FSA, while hydroxylated donors were tolerated by both aldolases but with FSA producing more product. With the asymmetrical ketone MHA, a stark difference in retention times indicated a difference in regioselectivity, with DERA promoting attack from the methyl side and FSA from the hydroxymethyl side. A slight retention time difference for the products of HAL can be ascribed to varying hydroxyl stereoselectivity. Having validated the assay procedure and the enzyme activity, the procedure was repeated for fluoro-donors MFA, DFA, FAL, and FHA. When using DERA, the products of MFA, DFA, and FAL were detected. No fluorinated products were detected with the aldolase FSA. Future engineering of aldolases or consideration of other enzyme homologs may be required to achieve activity with the fluoro-donor FHA.

Reactions of MFA, DFA, and FAL catalyzed by DERA were then scaled up to 5 mM donor and 2.5 mM acceptor in 0.5 mL volume to enable qualitative ^{19}F NMR analysis (*Figure 4.9*). Unfortunately, no NMR peaks corresponding to a product of FAL were observed, such that the previously seen LC-MS peak may correspond to a trace product lower than the limit of detection in the NMR experiment ($\sim 0.5\%$). However, we were pleased to see NMR product peaks of significant intensity for MFA and DFA. Due to the potential lack of fluorine stereoselectivity and multiple speciation of fluorosugars (open-chain keto form, open-chain hydrated form, α - and β -furanose forms), the spectra are complex and a full assignment of product peaks has not yet been possible. However, ^{19}F splitting patterns easily allow the assignment of primary and secondary

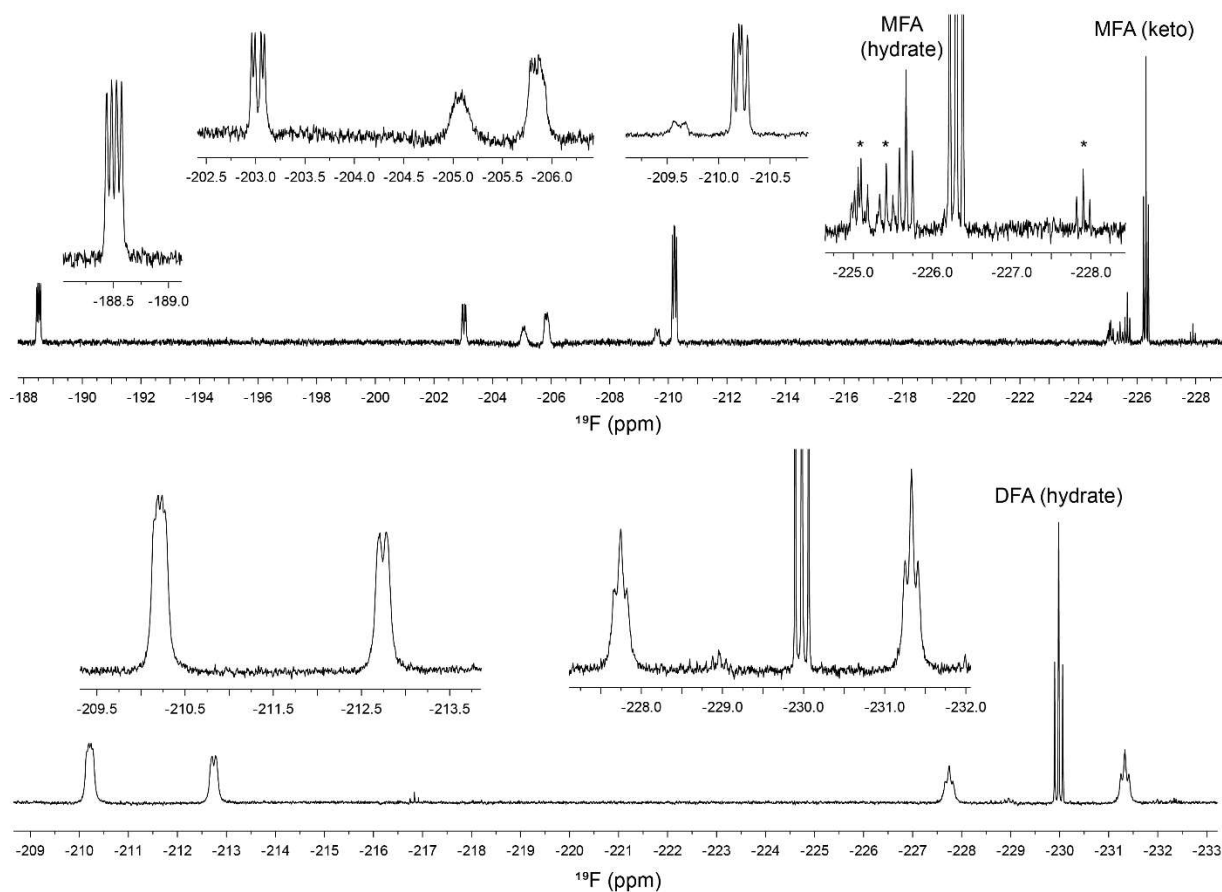


Figure 4.9. ^{19}F NMR spectra for the DERA-catalyzed aldol addition of MFA to G3P (upper spectrum) and DFA to G3P (lower spectrum). Asterisks indicate peaks corresponding to the minor product of MFA.

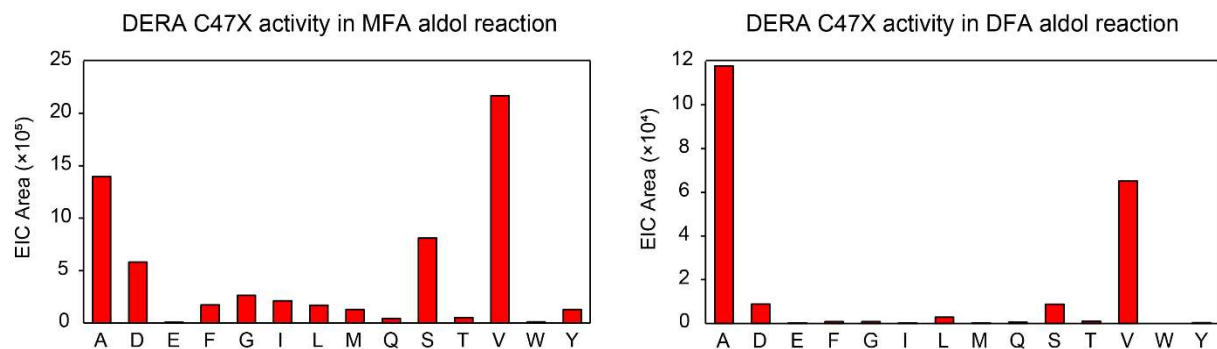


Figure 4.10. Screening of DERA mutants at Cys-47 for the aldol addition of MFA or DFA to G3P. Samples were analyzed by the derivatization-LC-MS procedure. All mutants produced fewer product ion counts than wild-type DERA (9.5×10^6 for MFA, 2.2×10^5 for DFA).

fluorides. The reaction of MFA produced secondary fluorides as a major product (attack from fluoromethyl side, doublet of doublet peaks) and primary fluorides as a minor product (attack from methyl side, triplet peaks). This distribution was surprising since the methyl side of MFA was expected to be more nucleophilic and native substrate-like than the fluoromethyl side. In literature on organocatalytic aldol reactions of MFA, the methyl attack is exclusively observed when water is used as solvent, but inclusion of organic solvents can favor the fluoromethyl attack [38]. Therefore, the encapsulation of MFA in the active site of DERA to achieve an environment more similar to that of an organic solvent may explain the observed regioselectivity for the biocatalytic aldol reaction of MFA.

Seeking to further optimize the observed conversion of MFA and DFA to 3-fluoro-1,3-dideoxyfructose-6-phosphate and 1,3-difluoro-1,3-dideoxyfructose-6-phosphate respectively, our attention was drawn to the DERA residue Cys-47 which has been implicated in irreversible enzyme inhibition at high concentrations of the native donor acetaldehyde. An aldol addition of acetaldehyde to an acceptor followed by dehydration forms a Michael acceptor tethered to the catalytic Lys-167, which upon Michael addition of Cys-47 effectively crosslinks the two residues and disables catalysis [39]. Mutation of Cys-47 has been employed to make inhibition-resistant DERA variants, but generally decreases the specific activity as a drawback [40]. We surmised that a Cys-47 related phenomenon could occur for fluoroketones, and so we prepared a purified protein collection that covered most possible mutations of this residue (14 out of 19 amino acids). Screening all variants on small scale with the derivatization-LC-MS procedure showed that only C47A, C47D, C47S, and C47V had significant activity with either MFA or DFA (Figure 4.10). However, all product ion counts were lower than with wild-type DERA, suggesting either that Cys-47 inhibition is not in effect, or that any such benefit was counterbalanced by intrinsically lower activity.

Next, we performed an alanine scan of other nearby hydrophobic residues, which could hinder the binding of fluorinated substrates due to fluorine's larger size or unpredictable effect on substrate hydrophobicity (Figure 4.11). The mutants L20A, V73A, F76A, and I139A were prepared, as these residues were observed to be close to the C-C bond forming pocket in a crystal structure of DERA [41]. The four hydrophobic-to-alanine mutants along with the four active Cys-47 mutants from before were screened on MFA and DFA by qualitative ¹⁹F NMR analysis. From these spectra, the peak integrals of reactants (R) and products (P) were used to calculate $(P_{mut}/[P_{mut}+R_{mut}])/(P_{wt}/[P_{wt}+R_{wt}])$ as a rough estimate of each mutant's relative activity to wild-

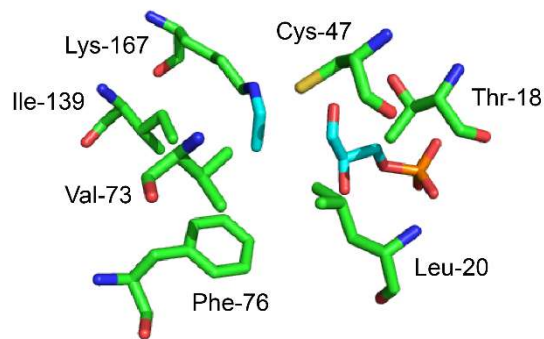


Figure 4.11. Active site of DERA from crystal structure RCSB PDB 6Z9I [Ref. 41]. Lysine-bound acetaldehyde and G3P are displayed in cyan.

DERA	MFA activity	DFA activity
WT	+++	+++
C47A	+++	+++
C47D	++	-
C47S	++	-
C47V	++	+
L20A	+*	-
V73A	+++	+++
F76A	++	-
I139A	+*	-

Figure 4.12. Relative activity to wild-type of DERA mutants for aldol addition of MFA or DFA. (-) No reaction, (+) <25% of wild-type, (++) 25-75% of wild-type, (+++) >75% of wild-type, (+*) only trace amount of primary fluoride product detected.

reaction system containing fluoro-donor, fructose-1,6-bisphosphate, and three enzymes (TpiA, FbaA, and DERA) in HEPES with $ZnCl_2$ (for the zinc-dependent FbaA) was successful in recapitulating the previously observed ^{19}F NMR peaks at higher intensity. Quantitative ^{19}F NMR was enabled by including a long recycle delay of 30 s between scans for atom relaxation, and performing measurements in triplicate (Figure 4.14). For the DERA-catalyzed reaction of MFA, no significant differences were seen among wild-type, C47A, and V73A, which all formed the secondary fluoride in 64-70% conversion and the primary fluoride in 4-8% conversion. However, using DFA as fluoro-donor gave noticeable differences in conversion: 34% with wild-type, 42% with C47A, and 81% with V73A. The remarkable benefit of V73A for DFA but not MFA is surprising owing to the two substrates' apparent similarity, but may be attributable to the fact that DFA is completely hydrated as a geminal diol in aqueous solution, while MFA is mostly present as the keto form. For the aldol reaction, DFA would require dehydration within the enzyme to its keto form, and non-catalytic active site residues may play a role in the ease of dehydration. Future studies are necessary to probe the mechanism of beneficial mutations and further optimize these fluoroketones as aldolase substrates.

type (Figure 4.12). All mutants were active on MFA and produced the secondary fluoride as a major product, with the exception of L20A and I139A which gave no secondary fluoride and a small amount of primary fluoride. When screened on DFA, only the mutants C47A, C47V, and V73A were active. Overall, C47A and V73A had the most promising relative activity on both MFA and DFA, with a relative activity at least 75% that of wild-type DERA.

Finally, to accurately compare the performance of DERA C47A and V73A against wild-type DERA, an improved assay was designed to allow for quantitative ^{19}F NMR analysis. To achieve this it was necessary to increase the concentration of MFA or DFA to 20 mM and use the fluoro-donor as the limiting reagent. The expensive nature of G3P, which is commercially provided only as dilute solutions, precluded its direct use at the desired concentration. Our attempts to replace G3P with D-glyceraldehyde were unsuccessful, resulting in no product formation. Instead, fructose-1,6-bisphosphate was applied at 20 mM along with a triose phosphate isomerase and a Type II fructose-1,6-bisphosphate aldolase, for a multi-enzymatic system supplying up to 40 mM G3P. In contrast to G3P, fructose-1,6-bisphosphate is readily commercially available in gram quantities as a solid. The necessary enzymes TpiA and FbaA were amplified from the genome of *E. coli*, cloned, expressed, and purified (Figure 4.13). The final

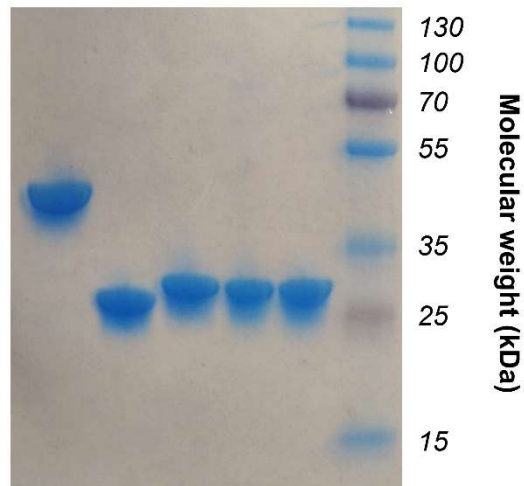


Figure 4.13. SDS-PAGE analysis of purified proteins used in the quantitative ^{19}F NMR aldol assay. From left to right: fructose-1,6-bisphosphate aldolase FbaA, triose phosphate isomerase TpiA, DERA V73A, DERA C47A, wild-type DERA, molecular weight marker.

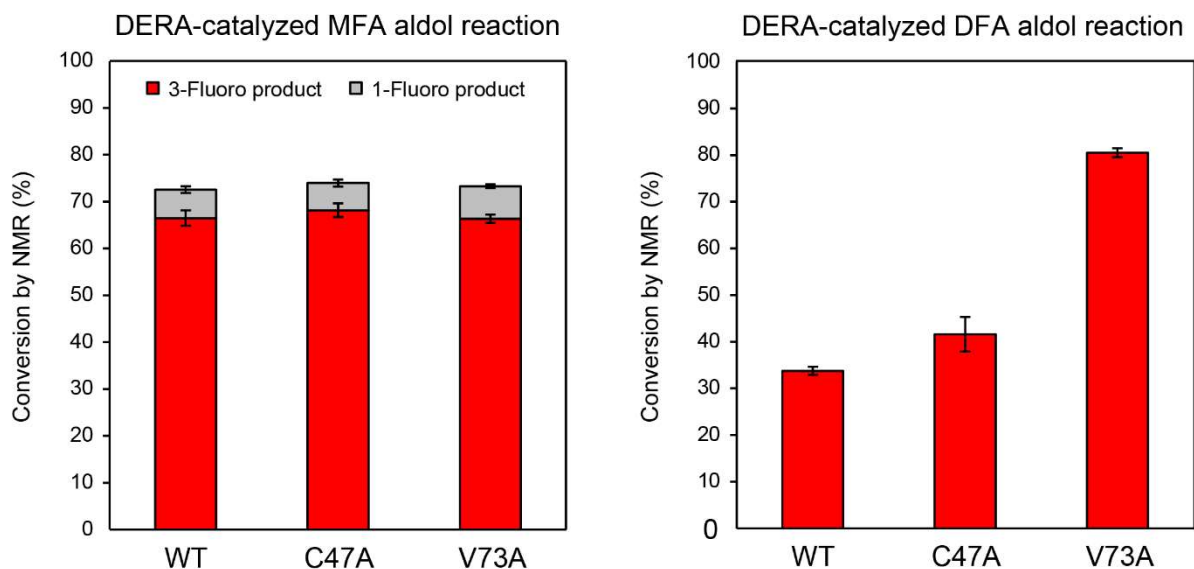


Figure 4.14. Results of quantitative ^{19}F NMR assessment for fluoroketone aldol addition to G3P, catalyzed by wild-type DERA, DERA C47A, or DERA V73A. Reactions contained 20 mM fluoro-donor, 20 mM fructose-1,6-bisphosphate, 100 μM DERA, 2.5 μM FbaA, 2.5 μM TpiA, 0.3 mM ZnCl_2 , and 20 mM HEPES pH 7.5.

4.4. Conclusion

This final chapter described our study of four simple organofluorine compounds as donor substrates for the aldolases DERA and FSA. These two aldolases are unique in their ability to use simple, uncharged carbonyl compounds as donors. Unlike the β -fluoro- α -ketoacids studied in prior chapters, fluorinated aldehydes and ketones as aldolase donors represent an efficient route towards fluorinated sugars, which have many applications in drug design and chemical biology. Our efforts to acquire potential substrate 1-fluoro-3-hydroxyacetone via a novel enzymatic oxidation of 3-fluoro-1,2-propanediol led to a deep study of the glycerol dehydrogenase enzyme, although more improvement of the oxidation yield is needed. Upon attempting the fluoro-aldol reactions, the

combination of DERA with fluoroacetone and 1,3-difluoroacetone formed fluorinated analogs of fructose-6-phosphate in appreciable conversion. Further optimizations by active-site engineering of DERA led to identification of beneficial mutations that provided 60-80% conversion of fluoro-donor to fluorosugar. Unfortunately, the other two fluoro-donors investigated, fluoroacetaldehyde and 1-fluoro-3-hydroxyacetone, have thus far not formed products above trace amounts. Future enzyme engineering or homolog discovery of aldolases may enable these remaining fluoro-donors, as our promising results with fluoroacetone and 1,3-difluoroacetone have conceptually proven the feasibility of the enzymatic fluoro-aldol reaction of simple fluoroketones. We hope that the preliminary analytical results in this study can be later expanded to a general methodology for synthesis and purification of diverse fluorinated sugars.

4.5. References

1. Linclau, B., Ardá, A., Reichardt, N.-C., Sollogoub, M., Unione, L., Vincent, S. P., Jiménez-Barbero, J. Fluorinated carbohydrates as chemical probes for molecular recognition studies. Current status and perspectives. *Chem Soc Rev* **2020**, *49* (12) 3863-3888.
2. Denavit, V., Lainé, D., St-Gelais, J., Johnson, P. A., Giguère, D. A chiron approach towards the stereoselective synthesis of polyfluorinated carbohydrates. *Nat Commun* **2018**, *9* 4721.
3. Pankiewicz, K. W. Fluorinated nucleosides. *Carbohydr Res* **2000**, *327* (1-2) 87-105.
4. Liu, P., Sharon, A., Chu, C. K. Fluorinated nucleosides. Synthesis and biological implication. *J Fluor Chem* **2008**, *129* (9) 743-766.
5. McCarter, J. D., Yeung, W., Chow, J., Dolphin, D., Withers, S. G. Design and synthesis of 2'-deoxy-2'-fluorodisaccharides as mechanism-based glycosidase inhibitors that exploit aglycon specificity. *J Am Chem Soc* **1997**, *119* (25) 5792-5797.
6. Tu, Z., Lin, Y.-N., Lin, C.-H. Development of fucosyltransferases and fucosidase inhibitors. *Chem Soc Rev* **2013**, *42* (10) 4459-4475.
7. Kim, J.-H., Resende, R., Wennekes, T., Chen, H.-M., Bance, N., Buchini, S., Watts, A. G., Pilling, P., Streltsov, V. A., Petric, M., Liggins, R., Barrett, S., McKimm-Breschkin, J. L., Niikura, M., Withers, S. G. Mechanism-based covalent neuraminidase inhibitors with broad-spectrum influenza antiviral activity. *Science* **2013**, *340* (6128) 71-75.
8. Buchini, S., Gallat, F.-X., Greig, I. R., Kim, J.-H., Wakatsuki, S., Chavas, L. M. G., Withers, S. G. Tuning mechanism-based inactivators of neuraminidases: Mechanistic and structural insights. *Angew Chem Int Ed* **2014**, *53* (13) 3382-3386.
9. Millett, F., Raftery, M. A. Fluorine-19 nuclear magnetic resonance study of the binding of trifluoroacetylglucosamine oligomers to lysozyme. *Biochemistry* **1972**, *11* (9) 1639-1643.
10. O'Connell, T. M., Gabel, S. A., London, R. E. Anomeric dependence of fluorodeoxyglucose transport in human erythrocytes. *Biochemistry* **1994**, *33* (36) 10985-10992.
11. Beuthien-Baumann, B., Hamacher, K., Oberdorfer, F., Steinbach, J. Preparation of fluorine-18 labelled sugars and derivatives and their application as tracer for positron-emission-tomography. *Carbohydr Res* **2000**, *327* (1-2) 107-118.

12. Li, A., Cai, L., Chen, Z., Wang, M., Wang, N., Nakanishi, H., Gao, X.-D., Li, Z. Recent advances in the synthesis of rare sugars using DHAP-dependent aldolases. *Carbohydr Res* **2017**, *452* 108-115.
13. Schümperli, M., Pellaux, R., Panke, S. Chemical and enzymatic routes to dihydroxyacetone phosphate. *Appl Microbiol Biotechnol* **2007**, *75* (1) 33-45.
14. Fischer, M., Kählig, H., Schmid, W. Gram scale synthesis of 3-fluoro-1-hydroxyacetone phosphate: a novel donor substrate in rabbit muscle aldolase-catalyzed aldol reactions. *Chem Commun (Camb)* **2011**, *47* (23) 6647-6649.
15. Barbas III, C. F., Wang, Y. F., Wong, C.-H. Deoxyribose-5-phosphate aldolase as a synthetic catalyst. *J Am Chem Soc* **1990**, *112* (5) 2013-2014.
16. Chambre, D., Guérard-Hélaine, C., Darii, E., Mariage, A., Petit, J.-L., Salanoubat, M., de Berardinis, V., Lemaire, M., Hélaine, V. 2-Deoxyribose-5-phosphate aldolase, a remarkably tolerant aldolase towards nucleophile substrates. *Chem Commun* **2019**, *55* (52) 7498-7501.
17. Ma, S. K., Gruber, J., Davis, C., Newman, L., Gray, D., Wang, A., Grate, J., Huisman, G. W., Sheldon, R. A. A green-by-design biocatalytic process for atorvastatin intermediate. *Green Chem* **2010**, *12* 81-86.
18. Schurmann, M., Sprenger, G. A. Fructose-6-phosphate aldolase is a novel class I aldolase from *Escherichia coli* and is related to a novel group of bacterial transaldolases. *J Biol Chem* **2001**, *276* (14) 11055-11061.
19. Garrabou, X., Castillo, J. A., Guérard-Hélaine, C., Parella, T., Joglar, J., Lemaire, M., Clapés, P. Asymmetric self- and cross-aldol reactions of glycolaldehyde catalyzed by D-fructose-6-phosphate aldolase. *Angew Chem Int Ed* **2009**, *48* (30) 5521-5525.
20. Szekrenyi, A., Soler, A., Garrabou, X., Guérard-Hélaine, C., Parella, T., Joglar, J., Lemaire, M., Bujons, J., Clapés, P. Engineering the donor selectivity of D-fructose-6-phosphate aldolase for biocatalytic asymmetric cross-aldol additions of glycolaldehyde. *Chem Eur J* **2014**, *20* (39) 12572-12583.
21. Roldán, R., Sanchez-Moreno, I., Scheidt, T., Hélaine, V., Lemaire, M., Parella, T., Clapés, P., Fessner, W.-D., Guérard-Hélaine, C. Breaking the dogma of aldolase specificity: Simple aliphatic ketones and aldehydes are nucleophiles for fructose-6-phosphate aldolase. *Chem Eur J* **2017**, *23* (21) 5005-5009.
22. Junker, S., Roldán, R., Joosten, H.-J., Clapés, P., Fessner, W.-D. Complete switch of reaction specificity of an aldolase by directed evolution in vitro: Synthesis of generic aliphatic aldol products. *Angew Chem Int Ed* **2018**, *57* (32) 10153-10157.
23. Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., Vidal, M. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol* **2000**, *328* 575-592.
24. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A., Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009**, *6* (5) 343-345.

25. Wang, W., Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions, and insertions using QuikChange site-directed mutagenesis. *Biotechniques* **1999**, *26* (4) 680-682.
26. Laurent, V., Goubeyre, L., Uzel, A., H elaine, V., Nauton, L., Tra ikia, M., de Berardinis, V., Salanoubat, M., Gefflaut, T., Lemaire, M., Gu erard-H elaine, C. Pyruvate aldolase catalyze cross-aldol reactions between ketones: Highly selective access to multi-functionalized tertiary alcohols. *ACS Catal* **2020**, *10* (4) 2538-2543.
27. Moss, S. J., Murphy, C. D., Hamilton, J. T. G., McRoberts, W. C., O'Hagan, D., Schaffrath, C., Harper, D. B. Fluoroacetaldehyde: A precursor of both fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*. *Chem Comm* **2000**, *22* 2281-2282.
28. Murphy, C. D., O'Hagan, D., Schaffrath, C. Identification of a PLP-dependent threonine transaldolase: A novel enzyme involved in 4-fluorothreonine biosynthesis in *Streptomyces cattleya*. *Angew Chem Int Ed* **2001**, *40* (23) 4479-4481.
29. Morris, O., McMahon, A., Boutin, H., Grigg, J., Prenant, C. Automation of [18F]fluoroacetaldehyde synthesis: Application to a recombinant human interleukin-1 receptor antagonist (rhIL-1RA). *J Labelled Comp Radiopharm* **2016**, *59* (7) 277-283.
30. Pero, R. W., Babiarz-Tracy, P., Fondy, T. P. 3-Fluoro-1-hydroxypropan-2-one (fluorohydroxyacetone) and some esters. Syntheses and effects in BDF mice. *J Med Chem* **1977**, *20* (5) 644-647.
31. Chari, R. V. J., Kozarich, J. W. Synthesis and properties of halohydroxyacetones and halomethylglyoxals. *J Org Chem* **1982**, *47* (12) 2355-2358.
32. Fischer, M. Organic chemistry as a useful tool in structural biology (Ph.D. thesis). University of Vienna **2011**.
33. Tang, C.-T., Ruch Jr., F. E., Lin, E. C. C. Purification and properties of a nicotinamide adenine dinucleotide-linked dehydrogenase that serves an *Escherichia coli* mutant for glycerol catabolism. *J Bacteriol* **1979**, *140* (1) 182-187.
34. Subedi, K. P., Kim, I., Kim, J., Min, B., Park, C. Role of GldA in dihydroxyacetone and methylglyoxal metabolism of *Escherichia coli* K12. *FEMS Microbiol Lett* **2008**, *279* (2) 180-187.
35. Wang, X. J., Lee, H. S., Li, H., Yang, X. Q., Huang, X. J. The effects of substituting groups in cyclic carbonates for stable SEI formation on graphite anode of lithium batteries. *Electrochem Commun* **2010**, *12* 386-389.
36. Lee, J.-W., Oliveira, M. T., Jang, H. B., Lee, S., Chi, D. Y., Kim, D. W., Song, C. E. Hydrogen-bond promoted nucleophilic fluorination: concept, mechanism, and applications in positron emission tomography. *Chem Soc Rev* **2016**, *45* (17) 4638-4650.
37. Rende, U., Niittyl a, T., Moritz, T. Two-step derivatization for determination of sugar phosphates in plants by combined reversed phase chromatography/tandem mass spectrometry. *Plant Methods* **2019**, *15* (127).
38. Chen, X.-H., Luo, S.-W., Tang, Z., Cun, L.-F., Mi, A.-Q., Jiang, Y.-Z., Gong, L.-Z. Organocatalyzed highly enantioselective direct aldol reactions of aldehydes with

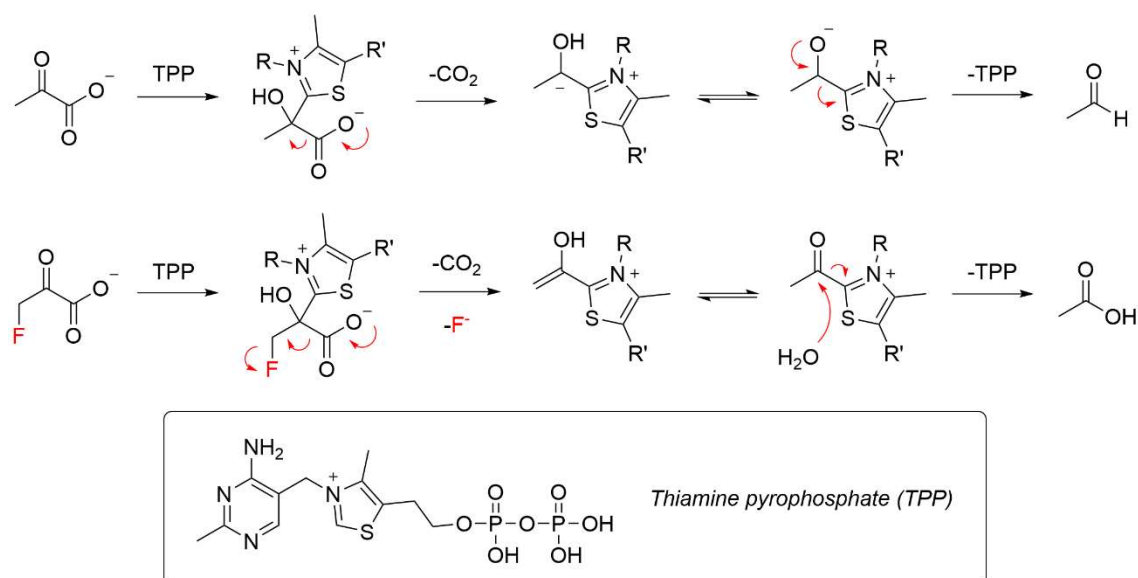
- hydroxyacetone and fluoroacetone in aqueous media: The use of water to control regioselectivity. *Chemistry* **2007**, *13* (2) 689-701.
39. Dick, M., Hartmann, R., Weiergräber, O. H., Bisterfeld, C., Classen, T., Schwarten, M., Neudecker, P., Willbold, D., Pietruszka, J. Mechanism-based inhibition of an aldolase at high concentrations of its natural substrate acetaldehyde: structural insights and protective strategies. *Chem Sci* **2016**, *7* (7) 4492-4502.
 40. Bramski, J., Dick, M., Pietruszka, J., Classen, T. Probing the acetaldehyde-sensitivity of 2-deoxy-ribose-5-phosphate aldolase (DERA) leads to resistant variants. *J Biotechnol* **2017**, *258* 56-58.
 41. Voutilainen, S., Heinonen, M., Andberg, M., Jokinen, E., Maaheimo, H., Pääkkönen, J., Hakulinen, N., Rouvinen, J., Lähdesmäki, H., Kaski, S., Rousu, J., Penttilä, M., Koivula, A. Substrate specificity of 2-deoxy-D-ribose 5-phosphate aldolase (DERA) assessed by different protein engineering and machine learning methods. *Appl Microbiol Biotechnol* **2020**, *104* (24) 10515-10529.

Appendix 1: *Summary of miscellaneous experiments*

Investigating the enzymatic synthesis of α -fluoroaldehydes

In this thesis, we have described a platform for chiral organofluorine synthesis based on the fluoropyruvate-tolerant pyruvate aldolase HpcH as an entryway to fluorine-bearing chiral centers. The initial aldol adducts, β -fluoro- α -ketoacids, were transformed into various downstream products using chemical and enzymatic methods. For example, chemical decarboxylation by hydrogen peroxide provided α -ketoacids, while enzymatic tailoring with lactate dehydrogenase or transaminase provided β -fluoro- α -hydroxyacids and β -fluoro- α -aminoacids respectively. We also desired a robust method for synthesis of α -fluoroaldehydes. Many of the novel organofluorines that were efficiently synthesized by HpcH contained multiple hydroxyl groups, thus the ability to transform them into α -fluoroaldehydes could provide fluorinated sugars, which have a wide variety of applications in pharmaceuticals and chemical biology [1]. Our efforts to develop enzymatic reactions for this goal were met with formidable challenges, so we eventually resorted to multi-step chemical methods to produce fluorinated sugars from fluoropyruvate aldol adducts. The results of enzymatic reactions aimed to produce α -fluoroaldehydes and the discussion of lessons learned are described in this appendix section.

We initially envisioned that the most direct route to α -fluoroaldehydes would be the non-oxidative decarboxylation of β -fluoro- α -ketoacids, thus entailing just one step from fluoropyruvate aldol adducts. Unlike the facile decarboxylation of β -ketoacids, the non-oxidative decarboxylation of α -ketoacids is mechanistically prohibited without a catalyst to effect polarity inversion, or umpolung. Typical umpolung catalysts such as cyanide or thiamine must add to the α -keto group, creating an electron sink that is β to the carboxylate to be removed [2]. However, these have the drawback of requiring harsh basic conditions for nucleophilic attack of the catalyst. Along the same mechanistic principles, the cofactor thiamine pyrophosphate (TPP) is employed by pyruvate dehydrogenase enzyme complex to convert pyruvate to acetyl-CoA, and by pyruvate decarboxylase to convert pyruvate to acetaldehyde. Unfortunately, when β -fluoro- α -ketoacids are subjected to this mechanism, the electron sink is necessarily β to the fluorine atom too, such that non-oxidative decarboxylation competes with β -elimination of fluoride [3].

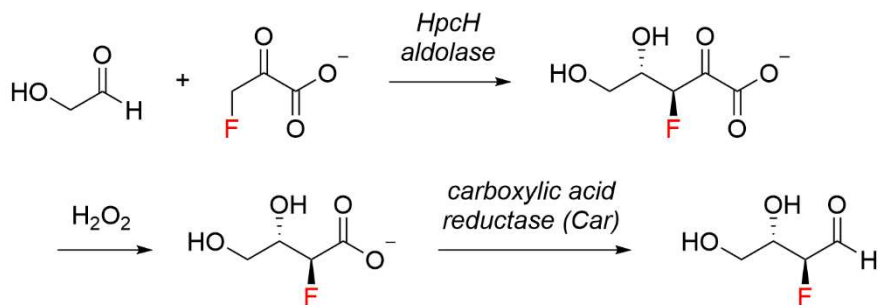


Appendix Figure 1. Divergent mechanisms of TPP-dependent pyruvate decarboxylase, resulting in non-oxidative decarboxylation of pyruvate to acetaldehyde, or fluoride elimination from fluoropyruvate.

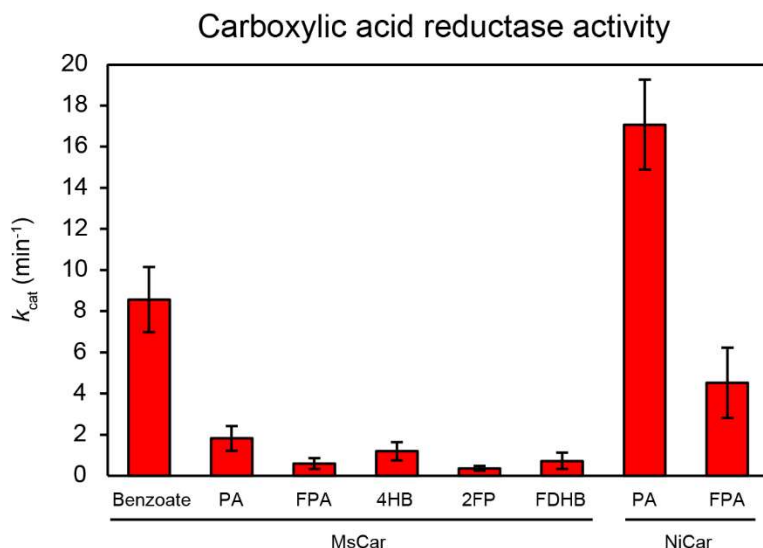
First, the chemocatalytic non-oxidative decarboxylation of β -fluoro- α -ketoacids was tested. Fluoropyruvate and (3*S*,4*S*)-3-fluoro-4,5-dihydroxy-2-oxopentanoate (FDHOP, generated from HpcH-catalyzed reaction of fluoropyruvate and glycolaldehyde) were used as substrates, in hopes of detecting fluoroacetaldehyde and (2*S*,3*S*)-2-fluoro-3,4-dihydroxybutyraldehyde by ^{19}F NMR. Substrate (20 mM) was incubated with thiamine (2 mM) in 20 mM HEPES pH 7.5 or 50 mM borate pH 9, for 1 h at 95 °C. At neutral pH, we observed a small amount of fluoride elimination but no other fluorinated species besides the starting material. The basic pH resulted in complete fluoride elimination, which was also verified to occur in the absence of thiamine. This shows that the harsh basic conditions needed for non-enzymatic thiamine activation cannot be applied to β -fluoro- α -ketoacids, which are unstable in basic aqueous solution. In hopes that an enzymatic system for non-oxidative decarboxylation would circumvent this problem, the TPP-dependent pyruvate decarboxylase from *Zymomonas mobilis* (ZmPdc) was cloned by Gibson assembly, expressed, and purified. Fluoropyruvate or FDHOP (20 mM) was incubated overnight with TPP (2 mM), MgCl_2 (1 mM), and ZmPdc (0.5 mg) in 20 mM HEPES pH 7.5. As expected, only fluoride elimination was observed for fluoropyruvate, while FDHOP underwent no reaction, likely due to not binding in the active site of ZmPdc. Thus, the non-oxidative decarboxylation of β -fluoro- α -ketoacids was deemed unsuitable for production of α -fluoroaldehydes.

An alternative two-step route was hypothesized, in which β -fluoro- α -ketoacids could first undergo oxidative decarboxylation to α -fluoroacids, then be reduced by a carboxylic acid reductase (Car) to provide α -fluoroaldehydes. Oxidative decarboxylation by hydrogen peroxide is rapid and selective, leaving only the second step to be investigated. Car enzymes are large multi-domain proteins that utilize ATP, NADPH, and a phosphopantetheine prosthetic group to reduce carboxylic acids to aldehydes [4]. The best substrates for Car are hydrophobic, such as fatty acids or benzoic acid, but some have activity on polar substrates [5]. In recent years, the biocatalytic community has developed robust syntheses of valuable aldehydes through discovery and engineering of Car proteins. However, the issue of whether or not α -fluoroacids could be reduced by Car had not been described in literature.

For our study, Car proteins from *Mycobacterium smegmatis* (MsCar) and *Nocardia iowensis* (NiCar) were cloned by Gibson assembly, expressed, and purified. MsCar was obtained from the genome of the organism, while NiCar was constructed from three synthetic gene fragments. Co-expression of phosphopantetheinyl transferase (Sfp) from *Bacillus subtilis* is needed for maturation of the prosthetic group of Car. Therefore, each Car was inserted into the upstream multiple cloning site (MCS) of the vector pETDuet-1 using restriction sites BamHI/NotI, followed by insertion of Sfp into the downstream MCS using sites NdeI/KpnI. Benzoate, phenylacetate (PA), 2-fluoro-2-phenylacetate (FPA), 4-hydroxybutyrate (4HB), 2-fluoropropionate (2FP), and 2-fluoro-3,4-



Appendix Figure 2. Investigated route for the conversion of fluoropyruvate aldol adducts into fluorinated sugars, via oxidative decarboxylation followed by enzymatic reduction by Car enzymes.



Appendix Figure 3. Activity assay of MsCar and NiCar for the reduction of different carboxylic acids. Fluorinated and hydroxylated substrates were reduced more slowly than simple hydrophobic substrates.

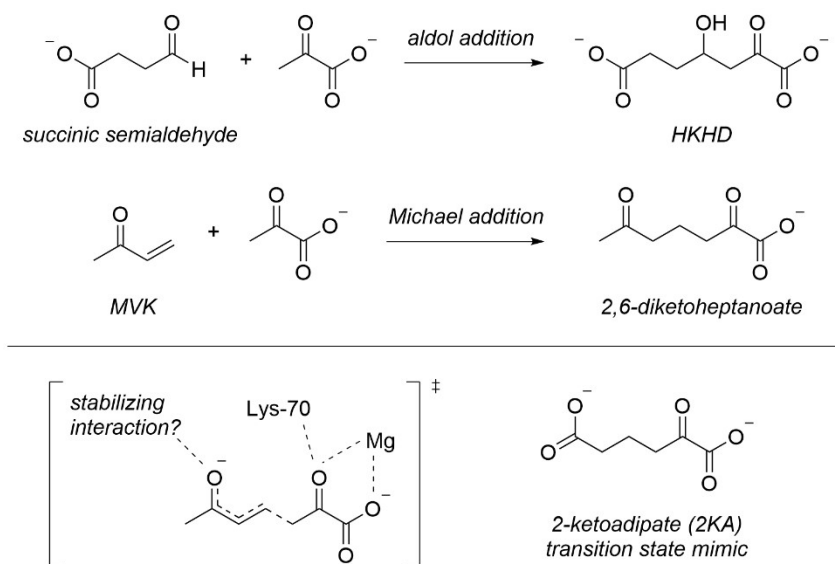
dihydroxybutyrate (FDHB, generated from HpcH-catalyzed reaction of fluoropyruvate and glycolaldehyde followed by H_2O_2 decarboxylation) were tested as substrates. Each substrate (0.5 mM) was incubated with ATP (1 mM), NADPH (0.2 mM), MgCl_2 (1 mM), and MsCar (0.5 or 2.5 μM) in 20 mM HEPES pH 7.5. Due to poor expression yield of NiCar (< 0.5 mg/L), only PA and FPA were used as substrates, and less enzyme was applied (50 nM). The disappearance of NADPH was monitored as a decrease in absorbance at 340 nm, and converted to an estimated turnover rate for each substrate and enzyme combination. Rates of reduction for hydroxylated or fluorinated substrates were consistently lower than for hydrophobic substrates benzoate and PA. Although the rejection of increased polarity may be averted by protein engineering, the influence of fluorine on the reduction rate may be harder to avoid. This is because the rate-determining step of Car is the nucleophilic attack of the carboxylate onto ATP, and α -fluoro substitution greatly decreases the carboxylate nucleophilicity. Taking these obstacles into consideration, the two-step route to α -fluoroaldehydes or fluorosugars based on Car reduction was not implemented.

Engineering towards an aldolase-catalyzed Michael reaction

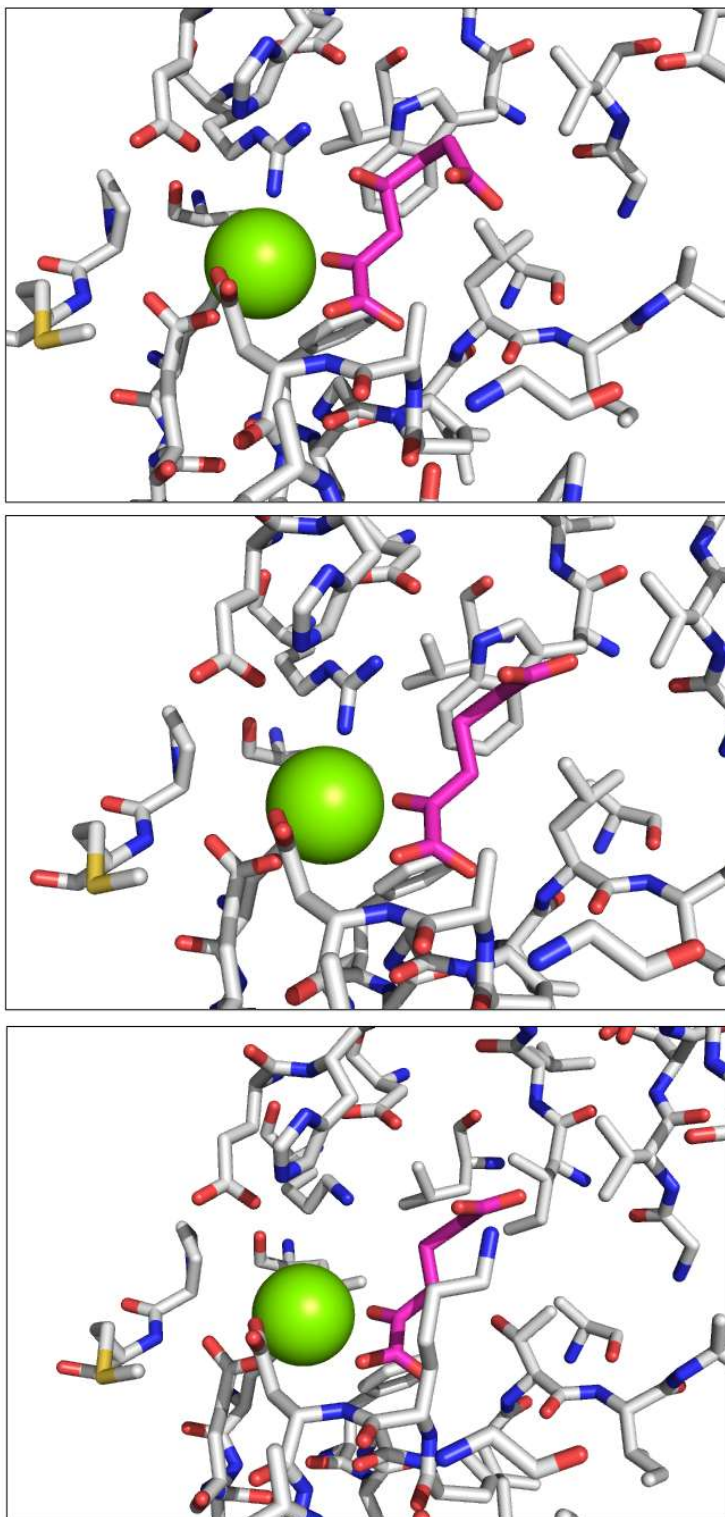
The Michael reaction is the conjugate addition of a carbon nucleophile to an α,β -unsaturated carbonyl compound, and is an important method of C-C bond formation in synthetic chemistry. Few native enzyme mechanisms are known to involve Michael reactions, and those that do such as macrophomate synthase typically operate upon complex substrates in natural product biosynthesis [6]. On the other hand, enzymatic Michael reactions tractable for biocatalysis have been achieved with promiscuous enzymes such as lipases and 4-oxalocrotonate tautomerase, as well as non-natural enzymes such as a *de novo* designed retro-aldolase [7-9]. Further expanding the limited toolbox of Michaelases is ideal for combining the sustainability of enzymes with the utility of C-C bond formation. Hoping to transfer the excellent properties of HpcH aldolase (eg. protein stability, nucleophile specificity for pyruvate analogs, and electrophile promiscuity) to a fundamentally different reaction, we undertook a computational approach for engineering of HpcH towards catalysis of the Michael reaction. Unfortunately, we were ultimately unsuccessful in

generating Michaelase activity. In this section, the computational and experimental procedures are briefly outlined, and the lessons learned are discussed.

A computational redesign of the HpcH active site was investigated using the PyRosetta 3.6 software suite [10]. First, the crystal structure of HpcH in complex with 4-hydroxy-2-ketoheptane-1,2-dioate (HKHD, the native aldol product of this enzyme) was obtained from RCSB PDB 4B5V [11]. Atom labels and coordinates of all atoms in HKHD were recorded. The PDB file was cleaned manually by deletion of all water molecules, ligands, and subunit chains of the homohexamer except for the single chain corresponding to the location of HKHD. To design a protein that might enable the Michael addition of pyruvate to methyl vinyl ketone (MVK) forming 2,6-diketoheptanoate, the ligand 2-ketoadipate (2KA) was thought to be a close mimic for the negatively-charged transition state. The MOL file of 2KA (obtained from RCSB PDB ligand identifier OOG) was converted to a PyRosetta ligand parameter set and a PDB file using the molfile_to_params script. The PDB file of 2KA was then concatenated to the cleaned PDB file of HpcH. To manually dock 2KA into the active site, the coordinates of carbon atoms C1, C2, C3, C5, and all their attached oxygen or hydrogen atoms were mapped to the corresponding atoms of HKHD. Atoms C4 and C6 of 2KA were mapped to C4 and C6 of HKHD, and appropriate coordinates of C4 hydrogen atoms and C6 oxygen atoms were calculated with standard geometry and bond length. Next, the mutate_residue function was used to change Arg-70 to Lys-70, a mutation that is known to drastically lower the rate of aldol addition without greatly affecting the binding of pyruvate [11]. Finally, the sidechain packer tool of PyRosetta was used to sample all possible mutations and sidechain conformations at the non-catalytic residues Trp-19, Ala-174, and Leu-212, followed by energy optimization with the MinMover minimization algorithm and energy scoring of all protein variants with the full-atom scorefunction ref2015. In addition to the stipulated R70K mutation, the lowest energy pose outputted contained the mutations W19I, A174K, and L212T. Examination of the pose revealed that the amine sidechain of A174K formed a new hydrogen bond with the C6 carboxylate of 2KA.



Appendix Figure 4. Comparison of the native reaction of HpcH aldolase with the desired Michael reaction. To computationally probe mutations for stabilizing the transition state of the Michael reaction, 2-ketoadipate (2KA) was used as a transition state mimic.



Appendix Figure 5. Models of the HpcH active site during computational design for Michaelase activity. Upper panel: the native aldol adduct HKHD co-crystallized in wild-type HpcH. Middle panel: transition state mimic 2KA manually docked in the active site of HpcH using the coordinates of HKHD. Lower panel: HpcH R70K subjected to sampling of non-catalytic sidechains, resulting in additional mutations W19I, A174K, and L212T.

After computational design revealed A174K as a potential key mutation, the HpcH variants R70K, A174K, and R70K/A174K were cloned, expressed, and purified for experimental comparison to wild-type. Michael reactions were attempted by incubating pyruvate (50 mM), MVK (100 mM), MgCl₂ (1 mM), and HpcH variant (0.1 mol%) in 20 mM HEPES pH 7.5 overnight. Decarboxylation was effected with H₂O₂ (10 eq.), then the samples were evaporated in a vacuum concentrator and the residues dissolved in D₂O for ¹H NMR analysis. Compared to a commercial standard of the expected product 5-ketohexanoate (from the decarboxylation of 2,6-diketoheptanoate) no products of a Michael reaction were detected. A more sensitive LC-MS method for detection of 5-ketohexanoate was developed that combined HILIC chromatography with ESI- detection of the 129.1→59.0 *m/z* transition (acetate fragmentation). However, desired products were still not detected in the reaction samples.

Two possibilities were considered. The lack of reactivity could simply mean that MVK was not bound in a productive orientation, without affecting the necessary interactions of pyruvate. Alternatively, the lysine mutations could be detrimental to binding or α -deprotonation of pyruvate. A deuterium exchange experiment was conducted by incubating pyruvate (50 mM), MgCl₂ (1 mM), and HpcH variant (0.1 mol%) in 20 mM HEPES-buffered D₂O (pD 7.5) for 1 h. The samples were decarboxylated by adding a solution of urea-hydrogen peroxide in D₂O (10 eq.) and then spiked with equal amounts of DMSO (~20 mM) as an internal standard for ¹H NMR. The peak integral ratio of acetate protons (1.81 ppm) to DMSO protons (2.71 ppm) was used to calculate the amount of α -protons remaining. Values were normalized to a control reaction without enzyme that was decarboxylated immediately, serving as a reference point for 100% proton retention. The proton retention was 80% for incubation in D₂O without aldolase, 0% with wild-type HpcH, 13% with R70K, 52% with A174K, and 70% for R70K/A174K. In the double mutant, the activity of deuterium exchange is no longer significantly above the background rate. Therefore, the enzyme interactions responsible for either binding or α -deprotonation of pyruvate are slightly perturbed by R70K and severely perturbed by A174K. The detrimental effect of these lysine mutations could not be predicted from our simplified PyRosetta computational model, and a more refined model allowing for ligand flexibility of a transition state mimic may be required. The overall complexity of large enzymes like HpcH also poses a challenge to the current state of computational enzyme design, thus combining the approach with others such as structure-guided directed evolution may be appropriate. On account of the encountered difficulties, the investigation to engineer a native aldolase into a Michaelase was not continued.

References

1. Linclau, B., Ardá, A., Reichardt, N.-C., Sollogoub, M., Unione, L., Vincent, S. P., Jiménez-Barbero, J. Fluorinated carbohydrates as chemical probes for molecular recognition studies. Current status and perspectives. *Chem Soc Rev* **2020**, *49* (12) 3863-3888.
2. Hanson, R. W. Decarboxylation of α -keto acids. *J Chem Educ* **1987**, *64* (7) 591.
3. Gish, G., Smyth, T., Kluger, R. Thiamin diphosphate catalysis. Mechanistic divergence as a probe of substrate activation of pyruvate decarboxylase. *J Am Chem Soc* **1988**, *110* (18) 6230-6234.
4. Kalim Akhtar, M., Turner, N. J., Jones, P. R. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proc Natl Acad Sci U S A* **2013**, *110* (1) 87-92.

5. Winkler, M. Carboxylic acid reductase enzymes (CARs). *Curr Opin Chem Biol* **2018**, *43* 23-29.
6. Guimarães, C. R. W., Udier-Blagović, M., Jorgensen, W. L. Macrophomate synthase: QM/MM simulations address the Diels-Alder versus Michael-Aldol reaction mechanism. *J Am Chem Soc* **2005**, *127* (10) 3577-3588.
7. Cai, J.-F., Guan, Z., He, Y.-H. The lipase-catalyzed asymmetric C-C Michael addition. *J Mol Catal B Enzym* **2011**, *68* (3-4) 240-244.
8. Geertsema, E. M., Miao, Y., Tepper, P. G., de Haan, P., Zandvoort, E., Poelarends, G. J. Biocatalytic Michael-type additions of acetaldehyde to nitroolefins with the proline-based enzyme 4-oxalocrotonate tautomerase yielding enantioenriched γ -nitroaldehydes. *Chemistry* **2013**, *19* (43) 14407-14410.
9. Garrabou, X., Beck, T., Hilvert, D. A promiscuous de novo retro-aldolase catalyzes asymmetric Michael additions via Schiff base intermediates. *Angew Chem Int Ed* **2015**, *54* (19) 5609-5612.
10. Chaudhury, S., Lyskov, S., Gray, J. J. PyRosetta: A script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **2010**, *26* (5) 689-691.
11. Coincon, M., Wang, W., Sygusch, J., Seah, S. Y. K. Crystal structure of reaction intermediates in pyruvate class II aldolase: substrate cleavage, enolate stabilization, and substrate specificity. *J Biol Chem* **2012**, *287* (43) 36208-36221.

Appendix 2: *Reaction tables for ^{19}F NMR assays*

Table 1. ^{19}F NMR conversion data for enzymatic reactions of fluoropyruvate in Chapter 2. Reactions contained 50 mM fluoropyruvate, 55 mM acceptor, 1 mM MgCl_2 , and the indicated amount of enzyme. The buffer was 20 mM HEPES pH 7.5 or 50 mM phosphate pH 7.5. Reactions were incubated for 16 h and decarboxylated with H_2O_2 . Reaction sets are grouped by acceptor. The ratio of diastereomers is expressed as % *syn*; the % *anti* is equal to $100-(\% \text{ syn})$.

Acceptor: D-glyceraldehyde

Enzyme	mol %	Buffer	Conv. (%)	% <i>syn</i>
GarL	0.05	Phosphate	96	48
RhmA	0.05	HEPES	93	27
EcHpcH	0.05	HEPES	100	19
EcHpcH V118F	0.02	HEPES	84	16
EcHpcH A174F	0.02	HEPES	34	7
EcHpcH A174F	0.05	HEPES	52	6
EcHpcH L212F	0.02	HEPES	16	86
EcHpcH L212F	0.05	HEPES	38	85
SwHpcH1	0.05	Phosphate	100	7
SwHpcH2	0.02	HEPES	83	13

Acceptor: L-glyceraldehyde

Enzyme	mol %	Buffer	Conv. (%)	% <i>syn</i>
GarL	0.05	Phosphate	97	6
RhmA	0.05	HEPES	92	0
EcHpcH	0.05	HEPES	81	0
EcHpcH V118F	0.02	HEPES	76	0
EcHpcH A174F	0.02	HEPES	19	0
EcHpcH L212F	0.02	HEPES	34	7
SwHpcH1	0.05	Phosphate	90	2
SwHpcH2	0.02	HEPES	86	0

Acceptor: D-lactaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	97	90
RhmA	0.05	HEPES	90	62
EcHpcH	0.05	HEPES	100	57
EcHpcH V118F	0.02	HEPES	93	42
EcHpcH A174F	0.02	HEPES	36	21
EcHpcH A174F	0.05	HEPES	52	25
EcHpcH L212F	0.02	HEPES	17	78
EcHpcH L212F	0.05	HEPES	32	78
SwHpcH1	0.05	Phosphate	100	59
SwHpcH2	0.02	HEPES	72	57

Acceptor: L-lactaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	94	7
RhmA	0.05	HEPES	97	5
EcHpcH	0.05	HEPES	99	4
EcHpcH V118F	0.02	HEPES	99	0
EcHpcH A174F	0.02	HEPES	36	0
EcHpcH L212F	0.02	HEPES	18	0
SwHpcH1	0.05	Phosphate	100	5
SwHpcH2	0.02	HEPES	87	0

Acceptor: glycolaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	HEPES	10	35
GarL	0.05	Phosphate	99	41
GarL	0.03	Phosphate	97	31
GarL	0.01	Phosphate	83	29
RhmA	0.05	HEPES	71	24
RhmA	0.05	Phosphate	95	42
EcHpcH	0.05	HEPES	98	61
EcHpcH	0.01	HEPES	85	7
EcHpcH	0.05	Phosphate	100	60
EcHpcH V118F	0.02	HEPES	96	7
EcHpcH V118F	0.05	HEPES	94	17
EcHpcH A174F	0.02	HEPES	37	10
EcHpcH L212F	0.02	HEPES	19	25
SwHpcH1	0.05	HEPES	88	33
SwHpcH1	0.05	Phosphate	99	63
SwHpcH2	0.02	HEPES	100	32
SwHpcH2	0.02	Phosphate	98	30

Acceptor: 2,2-dimethoxyacetaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	HEPES	6	52
GarL	0.05	Phosphate	51	55
RhmA	0.05	HEPES	95	59
RhmA	0.03	HEPES	61	41
RhmA	0.01	HEPES	33	29
RhmA	0.05	Phosphate	90	58
EcHpcH	0.05	HEPES	97	41
EcHpcH	0.01	HEPES	67	16
EcHpcH	0.05	Phosphate	97	44
EcHpcH V118F	0.02	HEPES	86	20
EcHpcH V118F	0.05	HEPES	91	20
EcHpcH A174F	0.02	HEPES	35	71
EcHpcH L212F	0.02	HEPES	18	83
SwHpcH1	0.05	HEPES	59	14
SwHpcH1	0.05	Phosphate	79	10
SwHpcH2	0.02	HEPES	96	39
SwHpcH2	0.02	Phosphate	97	33

Acceptor: glyoxylic acid

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	HEPES	43	85
GarL	0.05	Phosphate	99	99
RhmA	0.05	HEPES	30	47
RhmA	0.05	Phosphate	26	42
EcHpcH	0.05	HEPES	99	8
EcHpcH	0.05	Phosphate	100	5
EcHpcH V118F	0.02	HEPES	86	7
EcHpcH A174F	0.02	HEPES	35	19
EcHpcH L212F	0.02	HEPES	24	36
SwHpcH1	0.05	HEPES	58	38
SwHpcH1	0.05	Phosphate	100	64
SwHpcH2	0.02	HEPES	86	9
SwHpcH2	0.02	Phosphate	87	4

Acceptor: 2-pyridinecarboxaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	91	68
RhmA	0.05	HEPES	79	58
EcHpcH	0.05	HEPES	93	60
EcHpcH V118F	0.02	HEPES	94	56
EcHpcH A174F	0.02	HEPES	43	57
EcHpcH L212F	0.02	HEPES	41	55
SwHpcH1	0.05	Phosphate	93	60
SwHpcH2	0.02	HEPES	83	59

Acceptor: acetaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	16	15
RhmA	0.05	HEPES	18	36
EcHpcH	0.05	HEPES	12	35
EcHpcH V118F	0.02	HEPES	21	38
EcHpcH A174F	0.02	HEPES	14	39
EcHpcH L212F	0.02	HEPES	13	39
SwHpcH1	0.05	Phosphate	12	37
SwHpcH2	0.02	HEPES	9	38

Acceptor: propionaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	21	69
RhmA	0.05	HEPES	70	69
EcHpcH	0.05	HEPES	56	70
EcHpcH V118F	0.02	HEPES	82	67
EcHpcH A174F	0.02	HEPES	35	66
EcHpcH L212F	0.02	HEPES	15	66
SwHpcH1	0.05	Phosphate	55	69
SwHpcH2	0.02	HEPES	72	69

Acceptor: butyraldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	29	81
RhmA	0.05	HEPES	28	69
EcHpcH	0.05	HEPES	26	69
EcHpcH V118F	0.02	HEPES	16	38
EcHpcH A174F	0.02	HEPES	17	69
EcHpcH L212F	0.02	HEPES	8	69
SwHpcH1	0.05	Phosphate	26	68
SwHpcH2	0.02	HEPES	16	68

Acceptor: benzaldehyde

Enzyme	mol %	Buffer	Conv. (%)	% syn
GarL	0.05	Phosphate	16	14
RhmA	0.05	HEPES	21	39
EcHpcH	0.05	HEPES	38	39
EcHpcH V118F	0.02	HEPES	33	27
EcHpcH A174F	0.02	HEPES	20	39
EcHpcH L212F	0.02	HEPES	13	67
SwHpcH1	0.05	Phosphate	35	40
SwHpcH2	0.02	HEPES	29	39

Table 2. ^{19}F NMR conversion data for enzymatic reactions of fluoropyruvate (**1**) in Chapter 3. Reactions contained 50 mM fluoropyruvate, 100 mM acceptor, 1 mM MgCl_2 , and 0.1 mol% enzyme, in 20 mM HEPES pH 7.5 or 20 mM 2:1 CHES/HEPES pH 8.5. Notation: tr. = trace product of <1%; n.d. = not detected. The ratio of diastereomers is expressed as % syn; the % anti is equal to 100-(% syn). (A) Optimization of reactions on formaldehyde. (B) Reactions with other aldehydes using the high pH modification.

A

EcHpch	Acceptor	pH	Time (h)	Conv. (%)
No enzyme	Formaldehyde	7.5	4	tr.
No enzyme	Formaldehyde	7.5	16	4
No enzyme	Formaldehyde	8.5	4	tr.
WT	Formaldehyde	7.5	4	69
W19A	Formaldehyde	7.5	4	9
F170A	Formaldehyde	7.5	4	5
L212A	Formaldehyde	7.5	4	7
F170A/L212A	Formaldehyde	7.5	4	2
F170V	Formaldehyde	7.5	4	6
F170V/L212A	Formaldehyde	7.5	4	4
F170L	Formaldehyde	7.5	4	30
F170I	Formaldehyde	7.5	4	7
L212I	Formaldehyde	7.5	4	57
L212V	Formaldehyde	7.5	4	33
WT	Formaldehyde	7.5	16	95
WT	Formaldehyde	8.5	1	54
WT	Formaldehyde	8.5	4	98
MBP-fused	Formaldehyde	7.5	4	76
MBP-fused	Formaldehyde	7.5	16	97
MBP-fused	Formaldehyde	8.5	4	96

B

EcHpcH	Acceptor	pH	Time (h)	Conv. (%)	% syn
WT	Acetaldehyde	8.5	1	54	57
WT	Acetaldehyde	8.5	4	87	42
WT	Propionaldehyde	8.5	1	81	66
WT	Isobutyraldehyde	8.5	1	52	80
WT	Isobutyraldehyde	8.5	4	61	80
WT	Isobutyraldehyde	8.5	16	67	80
WT	Isobutyraldehyde	8.5	48	68	81
WT	Isovaleraldehyde	8.5	1	58	66
WT	Isovaleraldehyde	8.5	4	92	65
MBP-fused	2-Ethylbutyraldehyde	8.5	1	17	80
MBP-fused	2-Ethylbutyraldehyde	8.5	4	21	81
MBP-fused	2-Ethylbutyraldehyde	8.5	16	29	81
MBP-fused	2-Ethylbutyraldehyde	8.5	48	30	84
MBP-fused	Cyclohexanecarboxaldehyde	8.5	1	49	79
MBP-fused	Cyclohexanecarboxaldehyde	8.5	4	72	82
WT	Furfural	8.5	1	56	39
WT	Furfural	8.5	4	59	39
WT	Furfural	8.5	16	61	39
WT	Furfural	8.5	48	65	39
WT	2-Thiophenecarboxaldehyde	8.5	1	24	43
WT	2-Thiophenecarboxaldehyde	8.5	4	27	43
WT	2-Thiophenecarboxaldehyde	8.5	16	30	42
WT	2-Thiophenecarboxaldehyde	8.5	48	36	42
WT	D-Glyceraldehyde	8.5	1	71	16
WT	L-Glyceraldehyde	8.5	1	95	0
WT	Glycolaldehyde	8.5	1	85	9
WT	Dimethoxyacetaldehyde	8.5	1	83	14
MBP-fused	Benzaldehyde	8.5	1	14	35
MBP-fused	Benzaldehyde	8.5	4	36	39
MBP-fused	Benzaldehyde	8.5	16	54	39
MBP-fused	Benzaldehyde	8.5	48	55	39
WT	2-Pyridinecarboxaldehyde	8.5	1	80	61
WT	3-Pyridinecarboxaldehyde	8.5	1	84	39
WT	4-Pyridinecarboxaldehyde	8.5	1	82	64

Table 3. ^{19}F NMR conversion data for enzymatic reactions of extended 3-fluoro-2-oxoacids in Chapter 3. Reactions contained 50 mM fluoro-donor, 100 mM acceptor, 1 mM MgCl_2 , and 0.1 mol% enzyme, in 20 mM HEPES pH 7.5 or 20 mM 2:1 CHES/HEPES pH 8.5. Notation: tr. = trace product of <1%; n.d. = not detected. Due to kinetic resolution, 100% conversion is equal to 50% consumption of the racemic donor. The ratio of diastereomers is expressed as % syn; the % anti is equal to 100-(% syn). (A) Optimization of reactions of **2a** and **2b** on formaldehyde. (B) Final optimization of **2b** with higher enzyme loading. (C) Reactions of **2a** and **2b** with other aldehydes. (D) Unsuccessful reactions of **2c** and **2d** even under forcing conditions; unsuccessful reactions of **2e** which reacted non-enzymatically.

A

Donor	EcHpcH	Acceptor	pH	Time (h)	Conv. (%)
2a	No enzyme	Formaldehyde	7.5	16	n.d.
2a	No enzyme	Formaldehyde	7.5	48	n.d.
2a	No enzyme	Formaldehyde	8.5	16	n.d.
2b	No enzyme	Formaldehyde	7.5	48	n.d.
2a	WT	Formaldehyde	7.5	16	50
2a	W19A	Formaldehyde	7.5	16	tr.
2a	F170A	Formaldehyde	7.5	16	2
2a	L212A	Formaldehyde	7.5	16	16
2a	F170A/L212A	Formaldehyde	7.5	16	tr.
2a	F170V	Formaldehyde	7.5	16	tr.
2a	F170V/L212A	Formaldehyde	7.5	16	tr.
2a	F170L	Formaldehyde	7.5	16	24
2a	F170I	Formaldehyde	7.5	16	2
2a	L212I	Formaldehyde	7.5	16	tr.
2a	L212V	Formaldehyde	7.5	16	n.d.
2a	WT	Formaldehyde	7.5	48	82
2a	WT	Formaldehyde	8.5	16	32
2b	WT	Formaldehyde	7.5	48	2
2b	W19A	Formaldehyde	7.5	48	tr.
2b	F170A	Formaldehyde	7.5	48	6
2b	L212A	Formaldehyde	7.5	48	10
2b	F170A/L212A	Formaldehyde	7.5	48	2
2b	F170V	Formaldehyde	7.5	48	10
2b	F170V/L212A	Formaldehyde	7.5	48	2
2b	F170L	Formaldehyde	7.5	48	13
2b	F170I	Formaldehyde	7.5	48	13
2b	L212I	Formaldehyde	7.5	48	tr.
2b	L212V	Formaldehyde	7.5	48	tr.

B

Donor	EcHpcH	Acceptor	pH	Time (h)	Conv. (%)
2b	WT (0.3%)	Formaldehyde	7.5	48	3
2b	F170L (0.3%)	Formaldehyde	7.5	48	22
2b	F170V (0.3%)	Formaldehyde	7.5	48	29
2b	F170A (0.3%)	Formaldehyde	7.5	48	29
2b	L212A (0.3%)	Formaldehyde	7.5	48	20
2b	WT (0.3%)	Formaldehyde	7.5	1 week	3
2b	F170L (0.3%)	Formaldehyde	7.5	1 week	22
2b	F170V (0.3%)	Formaldehyde	7.5	1 week	40
2b	F170A (0.3%)	Formaldehyde	7.5	1 week	32
2b	L212A (0.3%)	Formaldehyde	7.5	1 week	27

C

Donor	EcHpcH	Acceptor	pH	Time (h)	Conv. (%)	% syn
2a	WT	Acetaldehyde	7.5	48	87	44
2a	WT	Propionaldehyde	7.5	48	84	50
2a	WT	Isovaleraldehyde	7.5	48	71	54
2a	WT	D-Glyceraldehyde	7.5	48	63	4
2a	WT	L-Glyceraldehyde	7.5	48	85	31
2a	WT	Glycolaldehyde	7.5	48	77	18
2a	WT	Dimethoxyacetaldehyde	7.5	48	88	81
2a	WT	2-Pyridinecarboxaldehyde	7.5	48	87	22
2a	WT	3-Pyridinecarboxaldehyde	7.5	48	77	38
2a	WT	4-Pyridinecarboxaldehyde	7.5	48	86	39
2b	F170V (0.3%)	Acetaldehyde	7.5	1 week	98	50
2b	F170V (0.3%)	Propionaldehyde	7.5	1 week	98	50
2b	F170V (0.3%)	Isovaleraldehyde	7.5	1 week	89	56
2b	F170V (0.3%)	2-Pyridinecarboxaldehyde	7.5	1 week	50	55
2b	F170V (0.3%)	3-Pyridinecarboxaldehyde	7.5	1 week	98	37
2b	F170V (0.3%)	4-Pyridinecarboxaldehyde	7.5	1 week	98	37

D

Donor	EcHpCH	Acceptor	pH	Time (h)	Conv. (%)
2c	No enzyme	Formaldehyde	7.5	48	n.d.
2d	No enzyme	Formaldehyde	7.5	48	n.d.
2e*	No enzyme*	Formaldehyde	7.5	48	30*
<i>*Conversion expressed relative to total donor (not kinetic resolution). Identical outcome observed with inclusion of any enzyme, due to the background rate.</i>					
2c	WT	Formaldehyde	7.5	48	n.d.
2c	W19A	Formaldehyde	7.5	48	n.d.
2c	F170A	Formaldehyde	7.5	48	n.d.
2c	L212A	Formaldehyde	7.5	48	n.d.
2c	F170A/L212A	Formaldehyde	7.5	48	n.d.
2c	F170V	Formaldehyde	7.5	48	n.d.
2c	F170V/L212A	Formaldehyde	7.5	48	n.d.
2c	F170L	Formaldehyde	7.5	48	n.d.
2c	F170I	Formaldehyde	7.5	48	n.d.
2c	L212I	Formaldehyde	7.5	48	n.d.
2c	L212V	Formaldehyde	7.5	48	n.d.
2d	WT	Formaldehyde	7.5	48	n.d.
2d	W19A	Formaldehyde	7.5	48	n.d.
2d	F170A	Formaldehyde	7.5	48	n.d.
2d	L212A	Formaldehyde	7.5	48	n.d.
2d	F170A/L212A	Formaldehyde	7.5	48	n.d.
2d	F170V	Formaldehyde	7.5	48	n.d.
2d	F170V/L212A	Formaldehyde	7.5	48	n.d.
2d	F170L	Formaldehyde	7.5	48	n.d.
2d	F170I	Formaldehyde	7.5	48	n.d.
2d	L212I	Formaldehyde	7.5	48	n.d.
2d	L212V	Formaldehyde	7.5	48	n.d.
2c	WT (0.3%)	Glycolaldehyde	7.5	48	n.d.
2c	F170L (0.3%)	Glycolaldehyde	7.5	48	n.d.
2c	F170V (0.3%)	Glycolaldehyde	7.5	48	n.d.
2c	F170A (0.3%)	Glycolaldehyde	7.5	48	n.d.
2c	L212A (0.3%)	Glycolaldehyde	7.5	48	n.d.
2d	WT (0.3%)	Glycolaldehyde	7.5	48	n.d.
2d	F170L (0.3%)	Glycolaldehyde	7.5	48	n.d.
2d	F170V (0.3%)	Glycolaldehyde	7.5	48	n.d.
2d	F170A (0.3%)	Glycolaldehyde	7.5	48	n.d.
2d	L212A (0.3%)	Glycolaldehyde	7.5	48	n.d.

Table 4. ^{19}F NMR conversion data for non-enzymatic reactions in Chapter 3. The reaction of fluoropyruvate (**1**) or extended 3-fluoro-2-oxoacids (**2a**, **2b**) with formaldehyde was catalyzed by buffers or metals. Reactions contained 50 mM fluoro-donor, 100 mM formaldehyde, 20 mM buffers, and 1 mM divalent metal chlorides. Inclusion of a secondary amine buffer (Tricine, TAPS, or CHES) was found to suppress non-enzymatic reactions at the higher pH of 8.5. Notation: *tr.* = trace product of <1%; *n.d.* = not detected. Reactions of **2a** and **2b** in presence of cobalt or nickel were accompanied by ~5% fluoride elimination unless otherwise specified. Bolded conditions were employed for the synthesis of racemic standards of 2-fluoroesters.

Donor	Buffer	Metal	Time (h)	Conv. (%)	Notes
1	HEPES, pH 7.5	none	16	2	
1	HEPPS, pH 8.5	none	16	15	
1	Bicine, pH 8.5	none	16	7	
1	Tricine, pH 8.5	none	16	<i>n.d.</i>	
1	TAPS, pH 8.5	none	16	<i>n.d.</i>	
1	CHES/HEPES, pH 8.5	none	16	2	2:1 CHES pH 9.3 / HEPES pH 7.5
1	HEPPS, pH 8.5	Mg	16	28	Trace dimerization of 1
1	HEPPS, pH 8.5	Mn	16	25	Line broadening in ^{19}F NMR
1	HEPPS, pH 8.5	Fe	16	10	
1	HEPPS, pH 8.5	Co	16	44	
1	HEPPS, pH 8.5	Ni	16	50	Trace dimerization and double aldol of 1
1	HEPPS, pH 8.5	Cu	16	10	
1	HEPPS, pH 8.5	Zn	16	35	Trace dimerization of 1
1	HEPPS, pH 8.5	Co	48	77	Trace dimerization and double aldol of 1
2a	HEPPS, pH 8.5	Co	48	<i>tr.</i>	
2a	CAPS, pH 10.5	Co	48	<i>tr.</i>	
2a	Carbonate, pH 10.5	Co	48	9	
2a	Phosphate, pH 12.5	Co	48	44	17% F-elim; 14% unidentified side pdt
2a	HEPPS, pH 8.5	Ni	48	<i>tr.</i>	
2a	CAPS, pH 10.5	Ni	48	3	
2a	Carbonate, pH 10.5	Ni	48	11	
2a	Phosphate, pH 12.5	Ni	48	24	14% F-elim; 47% unidentified side pdt
2b	Phosphate, pH 12.5	Co	48	83	8% unidentified side pdt

Table 5. ^{19}F NMR conversion data for oxidations of FPDO catalyzed by glycerol dehydrogenase in Chapter 4. Reactions contained 100 mM FPDO, 10 mM NAD^+ , 25 μg recycling enzyme (LDH, MDH, or ADH), 30-120 mM co-substrate (pyruvate, oxaloacetate, or acetaldehyde), 0.1 mol% GldA, 20 mM HEPES pH 7.5, and 0.3 mM ZnCl_2 . All reactions were incubated at room temperature for 48 h. For reactions at higher pH, the buffer was replaced with 50 mM Bicine pH 8.5 or 50 mM carbonate pH 9.5. Entry 16 contained no recycling system; entry 17 contained 5 mM FPDO with no recycling system; entries 18-20 contained no GldA. The % conversion to the oxidized product FHA and % fluoride elimination are indicated; *n.d.* = not detected.

Entry	Recycling	Co-substrate (mM)	pH	FHA (%)	F elim. (%)
1	LDH	120	7.5	7.7	3.3
2	LDH	60	7.5	9.1	3.4
3	LDH	30	7.5	9.7	3.5
4	MDH	120	7.5	2.8	6.7
5	MDH	60	7.5	2.2	4.6
6	MDH	30	7.5	1.4	3.5
7	ADH	120	7.5	n.d.	7.1
8	ADH	60	7.5	n.d.	12.3
9	ADH	30	7.5	n.d.	11.8
10	LDH	120	8.5	11.2	5.8
11	LDH	60	8.5	12.5	5.9
12	LDH	30	8.5	10.2	6.1
13	LDH	120	9.5	4.3	33.5
14	LDH	60	9.5	3.9	30.8
15	LDH	30	9.5	1.4	28.6
16	none	none	7.5	n.d.	n.d.
17	none	none	7.5	n.d.	n.d.
18	LDH	120	7.5	n.d.	n.d.
19	LDH	120	7.5	n.d.	n.d.
20	LDH	120	7.5	n.d.	n.d.

Appendix 3: *Plasmids, oligonucleotides, and gBlocks*

A. Plasmids

Plasmid	Chapter	MCC#	UniProt	Genotype
pET16hp-IMDH	2, 3, 4	2378	-	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-IMDH T7t
pET16hp-DeoC	4	2402	P0A6L0	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-DeoC T7t
pET16hp-DeoC-C47A	4	3679	-	C47A
pET16hp-DeoC-V73A	4	3680	-	V73A
pET16hp-EcHpcH	2, 3	3178	B1IS70	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-EcHpcH T7t
pET16hp-EcHpcH-A174F	2	3195	-	A174F
pET16hp-EcHpcH-V118F	2	3196	-	V118F
pET16hp-EcHpcH-L212F	2	3197	-	L212F
pET16hp-EcHpcH-W19A	3	3304	-	W19A
pET16hp-EcHpcH-F170A	3	3311	-	F170A
pET16hp-EcHpcH-L212A	3	3305	-	L212A
pET16hp-EcHpcH-F170A-L212A	3	3312	-	F170A, L212A
pET16hp-EcHpcH-F170V	3	3313	-	F170V
pET16hp-EcHpcH-F170V-L212A	3	3314	-	F170V, L212A
pET16hp-EcHpcH-F170L	3	3315	-	F170L
pET16hp-EcHpcH-F170I	3	3316	-	F170I
pET16hp-EcHpcH-L212I	3	3317	-	L212I
pET16hp-EcHpcH-L212V	3	3318	-	L212V
pET16hp-EcHpcH-R70K	App.	3379	-	R70K
pET16hp-FbaA	4	3378	P0AB71	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-FbaA T7t
pET16hp-FsaA	4	3183	P78055	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-FsaA T7t
pET16hp-GarL	2	3176	P23522	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-GarL T7t
pET16hp-GldA	4	3677	P0A9S5	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-GldA T7t
pET16hp-Rhma	2	3177	P76469	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-Rhma T7t
pET16hp-SwHpcH1	2	3179	A5VH82	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-SwHpcH1 T7t
pET16hp-SwHpcH2	2	3180	A5VAX1	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-SwHpcH2 T7t
pET16hp-TpiA	4	3181	P0A858	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-TpiA T7t
pET16hp-ZmPdc	App.	3186	P06672	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₁₀ -PreSc-ZmPdc T7t
pET24a	2	0016	-	ColE1, Km ^R , <i>lacI</i> , pT7 [MCS]-His ₆ T7t
pET24a-VfAT	2	3259	F2XBU9	ColE1, Km ^R , <i>lacI</i> , pT7 VfAT-His ₆ T7t
pETDuet-1	App.	0004	-	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₆ -[MCS] pT7 [MCS] T7t
pETDuet-MsCar-Sfp	App.	3198	I7GER2	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₆ -MsCar pT7 Sfp T7t
pETDuet-NiCar-Sfp	App.	3199	Q6RKB1	ColE1, Ap ^R , <i>lacI</i> , pT7 His ₆ -NiCar pT7 Sfp T7t
pSV272.1	3	0390	-	ColE1, Km ^R , <i>lacI</i> , pT7 His ₆ -MBP-[MCS] T7t
pSV272.1-EcHpcH	3	3310	B1IS70	ColE1, Km ^R , <i>lacI</i> , pT7 His ₆ -MBP-EcHpcH T7t

B. Oligonucleotides

Oligonucleotide	Sequence
pET16hp-DeoC-F	AGCGGCCATCTAGAAGTGCTTTTTTCAGGGCCCGCATATGACTGATCTGAAAGCAAGCAGC
pET16hp-DeoC-R	TCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTAGTAGCTGCTGGCGCTCTTACC
pET16hp-FbaA-F	CTAGAAGTGCTTTTTTCAGGGCCCGCATATGTCTAAGATTTTTGATTCGTA AACCTGGC
pET16hp-FbaA-R	TCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTACAGAACGTCGATCGCGTTACG
pET16hp-FsaA-F	CCATCTAGAAGTGCTTTTTTCAGGGCCCGCATATGGAACGTATCTGGATACTTCAGACGT
pET16hp-FsaA-R	TCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTAAATCGACGTTCTGCCAAACGC
pET16hp-GarL-F	GAAGTGCTTTTTTCAGGGCCCGCATATGAATAACGATGTTTTCCCGAATAAATTCAAAGCC
pET16hp-GarL-R	CGGGCTTTGTTAGCAGCCGGATCCTTATTTTTTAAAGGTATCAGCCAGTTTCTGAGTGCC
pET16hp-GldA-F	GCAGCGGCCATCTAGAAGTGCTTTTTTCAGGGCCCGCATATGGACCGCATTATTCAATCAC
pET16hp-GldA-R	CTCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTATTCCTACTCTTG CAGGAAAC
pET16hp-RhmA-F	CGGCCATCTAGAAGTGCTTTTTTCAGGGCCCGCATATGAACGCATTATTAAGCAATCCCTT
pET16hp-RhmA-R	CAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTCAATAACTACCTTTTATGCGTGCC
pET16hp-TpiA-F	CCATCTAGAAGTGCTTTTTTCAGGGCCCGCATATGCGACATCCTTTAGTGATGGGTA ACTG
pET16hp-TpiA-R	TCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTAAAGCCTGTTTAGCCGCTTCTGC
pET16hp-ZmPdc-F	CTAGAAGTGCTTTTTTCAGGGCCCGCATATGAGTTATACTGTGCGTACCTATTTAGCGGAG
pET16hp-ZmPdc-R	CAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCCTTAGAGGAGCTTGCCCCATTTGACC
pETDuet-MsCar-F	TACCATGGGCAGCAGCCATCACCATCATCACCACAGCCAGGTGCACCAGCTCACGGTCAC
pETDuet-MsCar-R	AATACGATTACTTTCTGTTTCGACTTAAGCATTATGCTCAGATCAGACCGAACTCACGCAG
pETDuet-Sfp-F	TATTAGTTAAGTATAAGAAGGAGATATAACAATGAAGATTTACGGAATTTATATGGACCCG
pETDuet-Sfp-R	AAATTTTCGCAGCAGCGGTTTCTTTACCAGACTTATAAAAAGCTCTTCGTACGAGACCATTG
pSV272.1-EcHpcH-F	GGATCGAGGAAAACCTGTATTTTTTCAGGGCATGGAAAACAGTTTTTAAAGCGGCGCTGAAA
pSV272.1-EcHpcH-R	TGGCTGGCTAGCCCGTTTGATCTCGAGTGCGGCCGCATTAATACACGCCGGGCTTACGG
DeoC-C47X-F	ACCGCCGCTATC NNKATCTATCCTCGC
DeoC-C47X-R	GATAGCGGCGGTATTGCC
DeoC-V73A-F	CGTATCGCTACGGCGACCAACTTCCCACAC
DeoC-V73A-R	CGTAGCGATACGGATTTCCGG
EcHpcH-W19A-F	GGCCGCCCCGAGATCGGATTAGCGCTGGGGCTGAGTAGCAGCTAC
EcHpcH-W19A-R	GTAGCTGCTACTCAGCCCCAGCGCTAATCCGATCTGCGGGCGGCC
EcHpcH-R70K-F	AGCCGGTGGTAAAACCGTCGTGGAACGATCCGGTGCAAATCAAACA ACTGCTGGACGTCG
EcHpcH-R70K-R	TTCCACGACGGTTTTTACCACCGGCTGGCTGGGGTAGGGCGCAATCGCCTG
EcHpcH-V118F-F	GCCGGTATTCGCGGTTTTTGGCAGTGCCTGGCCCCG
EcHpcH-V118F-R	CGGGCCAGCGCACTGCCAAAACCGCAATACCGGC
EcHpcH-F170A-F	GAAGGCGTCGACGGCGTGGCGATCGGCCCGGCGGATCTG
EcHpcH-F170A-R	CAGATCCGCCGGGCCGATCGCCACGCCGTCGACGCCTTC
EcHpcH-F170I-F	GAAGGCGTCGACGGCGTATTATCGGCCCGGCGGATCTG
EcHpcH-F170I-R	CAGATCCGCCGGGCCGATAATCACGCCGTCGACGCCTTC
EcHpcH-F170L-F	GAAGGCGTCGACGGCGTGTGATCGGCCCGGCGGATCTG
EcHpcH-F170L-R	CAGATCCGCCGGGCCGATCAGCACGCCGTCGACGCCTTC
EcHpcH-F170V-F	GAAGGCGTCGACGGCGTGGTGTGATCGGCCCGGCGGATCTG
EcHpcH-F170V-R	CAGATCCGCCGGGCCGATCACCACGCCGTCGACGCCTTC
EcHpcH-A174F-F	GACGGCGTGTATTCGCCCCGTTTGATCTGAGCGCCGATATGGG
EcHpcH-A174F-R	CCCATATCGGCGCTCAGATCAAACGGGCCGATAAACACGCCGTC

EcHpcH-L212A-F	GGCAAAGCGCCGGGGATCGCGATCGCCAATGAGCAACTG
EcHpcH-L212A-R	CAGTTGCTCATTGGCGATCGCGATCCCCGGCGCTTTGCC
EcHpcH-L212F-F	GGCAAAGCGCCGGGGATCTTTATCGCCAATGAGCAACTG
EcHpcH-L212F-R	CAGTTGCTCATTGGCGATAAAGATCCCCGGCGCTTTGCC
EcHpcH-L212I-F	GGCAAAGCGCCGGGGATCATTATCGCCAATGAGCAACTG
EcHpcH-L212I-R	CAGTTGCTCATTGGCGATAATGATCCCCGGCGCTTTGCC
EcHpcH-L212V-F	GGCAAAGCGCCGGGGATCGTGATCGCCAATGAGCAACTG
EcHpcH-L212V-R	CAGTTGCTCATTGGCGATCACGATCCCCGGCGCTTTGCC
EcHpcH-seqL83-R	CAGCAGTTGTTGATTTGCACCGGATC

C. gBlocks

gBlock	Sequence
pET16hp-EcHpcH	ctagaagtgcctttttcagggcccgcacatATGGAAAACAGTTTTAAAGCGGCGCTGAAAGCAGGCCGCC CGCAGATCGGATTATGGCTGGGGCTGAGTAGCAGCTACAGCGCAGAGTTACTGGCCGGAGCAGGATT CGACTGGTTATTGATCGACGGTGAGCACGCGCCGAATAACGTGCAAACCGTGCTCACCCAGCTACAG GCGATTGCGCCCTACCCAGCCAGCCGGTGGTACGTCCGTGTTGGAACGATCCGGTGCAAATCAAAC AACTGCTGGACGTCGGCACACAAACCTTGTGGTGCCGATGGTACAAAACGCTGACGAAGCCCGTGA AGCGGTACGCGCCACCCGTTATCCCCCGCCGGTATTGCGGGTGTGGGCAGTGCGCTGGCCCCGAGCC TCGCGCTGGAATCGTATTCCCTGATTACCTGCAAAAAGCCAACGATCAAATGTGCGTGTGGTGCAGA TCGAAACGCGTGAGGCAATGAAGAACTTACCGCAGATTCTGGACGTGGAAGGCGTCGACGCGGTGTT TATCGGCCCGCGGATCTGAGCGCGATATGGGTTATGCCGGTAATCCGCAGCACCCGGAAGTACAG GCCGCCATTGAGCAGGCGATCGTGCAGATCCGTGAATCGGGCAAAGCGCCGGGGATCCTGATCGCCA ATGAGCAACTGGCAAACGCTATCTGGAAGTGGGCGCGCTGTTTGTGCGCGTCGGCGTTGACACCAC CCTGCTCGCCCGCGCCGCTGAAGCGCTGGCAGCACGGTTTTGGTGCAGGCCACCCGCGTGAAGCCC GGCGTGTATTAAggatccggctgctaacaaagcccgaag
pET16hp-SwHpcH1	ctagaagtgcctttttcagggcccgcacatATGAACAAGGTTTCGCACTTGTGGAACGAGGGTCGCCAG CTTTGGCAGGCTGGCTGCAGCTTCCCGGTAGCTTACACGCGGAAGCGTTGGCAGCTTTAGACTATGA TGCAGTGGTTATTGATATGCAGCATTTCGCTATTGACTTCGGACAAGTGGCCCTATGTTGATCGCG ATTGAGTTAGGGGGCGCAGAGCCATTTCGTCCTACGCAAGTAAACGACCCCTCAGACATTATGAAGC TGTTAGATGCAGGGGCATATGGGATTATCGCACCTATGGTCAATACGCGCGCAGAAAGCTCAAACCT GGCATCTGCTCTTCAATTACTCGCCCCGTGGGCTGCGCTCCTTCGGACCCCGCCCGCCGCTGCGT TATGGGTCAGGCTATCTGGCGCAAGCCTCGGAGACGGTCTGAGGCTTGGCTATGATCGAGACCCGTG AGGCGCTTGCAAATATCGATGAGATCCTGAGCGTAGATGGGATCGATGGCGTGTTCATTGGACCCAC AGACCTGGCTTTGGACCTGGGTCACGCCCACTTGTGGATACGGAGGAAGCAGAAGTAGTTAGTGCC ATCGCACATGTGCGTGAACGTGCTCACGCTGCTGGTAAACGTGTAGGGATCTTTTGGCGAAGTGGCG GCTTCGCACGCGTCAAGCTGGCAGAGGGTTTTGATTTTGTGACGGCAGCACCTGACTTGGCAATGCT GTCTGCGGCGGCCGTCAGGTTATTGCTGACGCTCGCGCATATGAggatccggctgctaacaaagcc cgaaag
pET16hp-SwHpcH2	ctagaagtgcctttttcagggcccgcacatATGAATGACTTTAAAACAGCTTTGCTGCAAGACCGTGTAC AGATCGGCCTTTGGCAGGCTTTGGCGAATCCTTATACAGCGGAAATCTGCGCAGGTGCAGGCTTTGA TTGGCTGCTGCTGGATGGGGAACACGCTCCAACGACTTACCACTGCTTCTGCGCAATTACAAGCC GTAGCGGCTTACCCAGTAGAGCCAGTAGTACGCCTTCCAACCTGGCGATGCAGTCTGTTGAAACAGA TGCTGGACATTGGGGCTCGTAGTTTACTGGTGGCGATGGTAGAATCTGCCGCGCAAGCCGAGGCAAT GGTCCGTGCAACTCGTTACCCGCTCAAGGCACTCGTGGAGTAGGTAGCGCTATCGGTGCGGCTTCC CGCTGGAACCGCACACATAATTATCTTAATGAGGCCGAAGATGAAATTTGCTTATGGTACAGATTG AGAGCGCCGCTGCTCTGGCTGCATTACCGTCTATCGCAGAATTGCCAGGGGTAGATGGAATTTTCAT TGGTCTTCCGACTTGGCGGCCAGCCTGGGACATCTTGGCGACCCCGGTACGCGCCGCTGCAACGT GCAATCGAACGTGCCATTGCAACGGTCCGTAATCCGCGAAGGCTGTTGGCATTCTTGGCGCAGACG AGGGCTTAGCGCGTCTGCTACCTGGAGTTGGGAGCAACATTTGTTGCTGTGCGGTACGGACGTGACCCT GCTGGCCCCGCGTGGGAAGCTTGGCCCCGCTGTTTTAAAGGTGGTCCGATGAGGAACCCCGCGGT GCAACTATCTATTAAggatccggctgctaacaaagcccgaag

pET24a-VfAT

tttgtttaactttaagaagagatatacatATGAATAAGCCCCAGTCGTGGGAAGCGCGCTGAAA
CGTACTCCTTGTACGGCTTCACGGACATGCCAGCCTGCACCAGCGCGGTACGGTGGTAGTCACGCA
TGGGGAAGGACCCTATATCGTGGACGTCAATGGACGCCGCTATCTGGACGCAAACCTCGGGACTTTGG
AATATGGTCGCGGGATTTGACCATAAGGGACTGATTGATGCGGCTAAAGCTCAATATGAACGCTTTC
CTGGTTACCATGCCTTTTTTGGACGCATGTCAGACCAAACAGTTATGCTTAGCGAAAAACTTGTGGA
GGTGTCCCCCTTGTATTCTGGACGCGTCTTTTATACGAACTCGGGATCTGAAGCTAACGATACGATG
GTGAAAATGTTGTGGTTCTTACATGCGGCAGAAGGAAAGCCGCAGAAAACGTAAGATCTTGACTCGTT
GGAATGCCTACCATGGCGTCACGGCCGTATCAGCGTCCATGACCGGGAAGCCCTACAACCTCGGTCTT
TGGATTACCTTACCAGGGTTCGTCCACCTGACGTGCCCTCATTACTGGCGCTACGGGGAAGAAGGC
GAGACCGAGGAACAGTTTCGTAGCGCGCTTAGCGCGTGAGTTAGAGGAGACTATCCAACGCGAGGGGG
CAGATACAATTGCGGGTTTTTTCGCCGAACCCGTTATGGGCGCTGGGGGAGTAATCCCACCCGCGAA
AGGATACTTTCAAGCAATCCTGCCTATCTTACGTAAGTATGATATCCCAGTCATCTCAGACGAGGTC
ATTTGTGGATTGGGCGTACAGGTAACACCTGGGGCTGCGTCACTTACGACTTCACCCGGACGCGA
TCATCAGTAGTAAGAACTTAACTGCCGCTTTTTTCTATGCGGGCTGTGATTTTAGACCGGAGTT
AAGCAAACGCCCTGGAGACCGCCATCGAGGCTATTGAGGAATTTCTCATGGGTTTACGGCATCAGGA
CATCCTGTAGGATGTGCTATTGCTTTAAAGGCGATTGACGTTGTCATGAACGAGGGCTTAGCTGAAA
ACGTACGTCGCTTTCGCGCTCGCTTCGAAGAGCGCCTTAAAGCATATTGCCGAACGTCCAACATTGG
AGAGTACCGTGGTATCGGATTCATGTGGGCACTTGAAGCGGTAAGGATAAGGCCAGCAAACCCCT
TTTGACGGGAATTTATCGGTCAGTGAACGTATCGCGAATACATGTACTGACTTGGGGTTAATCTGTC
GCCCACTGGGCAATCGGTCGTTTATGCCCCTCCATTTCATTTTGACCGAAGCAGATGGATGAAAT
GTTTGATAAATTAGAAAAGGCCCTGGACAAAGTTTTTTCGGGAGGTGGCCctcgagcaccaccaccac
caccactgagatc

pETDuet-NiCar-1

atgggcagcagccatcaccatcatcaccacagccagATGGCTGTCGATAGCCAGATGAGCGTTTAC
AGCGCCGCATTGCGCAGTTGTTTCGCCGAGGACGAGCAAGTCAAAGCTGCGGTCCTTTGGAGGCGGT
TTCGGCAGCTGTCAGTGCTCCTGGTATGCGTCTGGCTCAGATTGCCGCTACAGTTATGGCGGGGTAT
GCGGATCGTCCGGCGGCAGGACAGCGCGCTTTCGAGCTTAATACGGACGACGCTACCGGTCGTACCT
CACTTCGTTTTCCTCCCGTTTGAACCATTACCTACCGTGAGTTATGGCAACGTGTTGGAGAAGT
AGCAGCCGCATGGCATCATGATCCCGAGAACCCCTTGCGCCGAGGAGACTTCGTAGCTTTGCTTGGT
TTCACATCTATCGACTATGCCACTTTAGACCTGGCTGATATTCATTTAGGTGCAGTACCCTGCCG
TTCAGGCATCTGCGGCTGTAAGTCAATTGATTGCGATCTTGACAGAAACATCGCCACGCTTCTGGC
CTCGACGCCGAGCATTTAGATGCGGCAGTCGAGTGCTTACTGGCTGGCCTACACCAGAACGTCCTT
GTAGTCTTTGACTACCACCCAGAGGACGATGACCAACGCGCCGCTTTTGAATCGGCCCGTCGCCGCC
TTGCAGATGCGGGCTCCCTGGTTATTGTGGAGACGTTAGATGCCGTCGGTCCCGCGGCCGCGATTT
ACCAGCTGCCCTTTATTTGTCCCAGACACTGACGACGACCTCTTGCTTTGTTGATTTACTTCC
GGCAGCAGGGGACTCCAAAAGGCGCCATGTATACGAATCGCTTGGCAGCAACAATGTGGCAAGGGA
ATTCTATGCTGCAAGGTAATTCACAACGCGTCGGGATCAACCTTAATTACATGCCAATGAGC

pETDuet-NiCar-2

CAACGCGTCGGGATCAACCTTAATTACATGCCAATGAGCCATATTGCCGGGCGCATTTCCCTTTTTG
GAGTTCTTGC CGCGGAGGCACTGCTTACTTTGCTGCCAAATCTGACATGTCGACTCTGTTTGAGGA
CATTTGGATTGGTTCGCCCTACGGAAATCTTCTTTGTCCTCGCGTCTGGATATGGTGTTCAGCGT
TATCAATCGGAATTAGACCGCCGCTCCGTGGCTGGAGCAGACTTGGATACTCTGGACCGTGAAGTTA
AGGCCGATTTGCGTCAGAATTACCTGGGGGACGTTTTTTGGTAGCTGTAGTTGGCTCTGCACCGCT
TGCTGCCGAAATGAAGACATTCATGGAGTCGGTACTTGACCTGCCACTGCACGACGGTTATGGGTCC
ACCGAAGCGGGTGCTTCGGTTTTATTAGACAATCAAATCCAGCGTCCACCCGTCCTGGACTATAAAT
TGGTCGATGTACCTGAACTGGGCTATTTTTCGTACAGATCGTCTCATCCGCGTGGAGAATCTCTGTT
AAAAGCCGAAACCACGATCCAGGGTATTACAAGCGTCTTGAAGTAACAGCCGAAATTTTGGACGAG
GACGGGTTTTATAAGACTGGAGATATTGTTGCCAATTGGAACACGACCGTCTGGTATATGTAGATC
GTCGCAACAATGTGTTGAACTTAGTCAAGGAGAGTTTCGTTACCCTTGGCCACTTAGAGGCAGTTTT
TGCGAGTAGTCCGTTAATTCGCCAAATCTTTCATCTATGGAAGCAGTGAGCGTTCTTATCTTTGGCA
GTAATCGTTCCAACGACGACGCACTTCGCGGTCGCGACACCGCCACTCTTAAATCGGCCCTGGCGG
AAAGCATTACGCGTATTGCAAAAAGACGCCAATTTACAACCGTACGAGATTCCTCGTGACTTC

pETDuet-NiCar-3

GACGCCAATTTACAACCGTACGAGATTCCTCGTGACTTCTTAATTGAGACCGAACCTTTTACAATCG
CAAACGGACTTTTGTCTAGGCATTGCTAAGCTTCTTCGTCCCAATTTGAAGGAACGTTATGGTGCTCA
ATTAGAACAAAATGTACACGGATCTTGCCACGGGGCAAGCCGACGAGTTGTTAGCCCTGCGCCGTGAG
GCGGCAGATCTCCTGTACTTGAAACGGTTTCTCGTGCCGCAAAGGCAATGTTGGGAGTAGCATCAG
CTGATATGCGTCCCGACGCACATTTACGGACTTAGGCGGGGACTCCCTTAGTGCTCTTTCCTTTAG
CAATTTGCTGCACGAGATTTTGGGGTAGAAGTCCCAGTGGGAGTAGTGGTGTCTCCTGCCAACGAG
CTTCGCGACCTTGCCAATTACATCGAGGCAGAACGCAATAGCGGAGCGAAGCGTCCAAC TTCACCA
GTGTTACGGTGGCGGTTCTGAAATTCGCGCTGCTGACTTGACCCTTGACAAGTTCATCGACGCGCG
CACACTGGCAGCCGCGGACAGTATCCACACGCTCCCCTTCTGCTCAAACGGTGTACTGACGGGC
GCAAACGGCTATCTGGGCCGCTTCCTTTGCTTGGAATGGCTGGAACGCTTGATAAGACGGGTGGCA
CATTGATCTGCGTAGTCCGCGGGTTCGGACGCTGCAGCCGCGCGCAAACGTTTAGATTCGCGGTTTGA
TAGCGGAGACCCTGGTTTGTGGAACATTACCAACAATTAGCCGCACGCACGTTAGAAGTGTGGCC
GGGACATTGGGGACCCTAATCTTGGTCTTGATGATGCAACTTGGCAGCGTCTGGCGGAGACAGTGC
ACTTGATCGTACATCCGGCGGCCCTTGTGAATCACGTA CTCCGTACACACAATTGTTTGGACAAA
CGTAGTCGGGACGGCGGAAATCGTGCGTTTGGCGATTACGGCTCGTCCGCAAGCCCGTAACCTATCTT
AGCACAGTGGGCGTTGCGGATCAGGTCGATCCGGCAGAGTATCAGGAGGACAGCGATGTGCGTGAAA
TGTCTGCAGTCCGCGTTGTGCGCGAGAGTTACGCGAATGGCTATGGCAATAGTAAGTGGGCTGGGGA
AGTACTGTTGCGTGAGGCGCACGATCTGTGCGGGCTGCCTGTTGCTGTGTTCCGCAGTGACATGATC
CTTGCACATTCACGTTATGCCGGCCAAC TGAATGTCCAAGACGCTCTTACC CGCCTTATCTTGTCTT
TGGTAGCTACTGGCATCGCTCCATATTCATTTACC GCACAGACGCAGATGGAATCGCCAGCGTGC
CCACTACGATGGTTTGGCGGACAGCTTACGGCTGCGGCCATTACC GCGCTTGGTATTCAGGCCACA
GAGGGTTTCCGTACTTACGATGTATTGAATCCCTACGATGATGGAATTTTCGTTAGACGAGTTCGTGC
ATTGGTTAGTGAATCTGGGCACCCATCCAACGATATCACCGACTACTCAGACTGGTTCCACCGCTT
TGAGACCGGATTCGTGCGCTTCTTGAGAAACAGCGCCAAGCCAGTGTCTTACCGCTTTTGGATGCA
TACCGTAATCCTTGCCCCGAGTACGTGGAGCTATCTTGGCGGCGAAAGAATTCCAGGCCCGCTCC
AGACCGCGAAAATCGGTCTTGAGCAGGACATTC CGCATTTAAGTGCGCCACTTATCGACAAAATACGT
GTCTGACTTAGAACTGCTTCAGTTGCTTTAAgcataaatgcttaagtcgaacagaaagtaacgta
