DEVELOPMENT OF A NOVEL DEHYDROGENASE AND A

STABLE COFACTOR REGENERATION SYSTEM

A Dissertation
Presented to
The Academic Faculty

by

Eduardo Vázquez-Figueroa

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology
December 2008

DEVELOPMENT OF A NOVEL DEHYDROGENASE AND A

STABLE COFACTOR REGENERATION SYSTEM

Approved by:

Dr. Andreas S. Bommarius, Advisor
School of Chemical and Biomoloecular
Engineering
Georgia Institute of Technology

Dr. Donald F. Doyle
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. William J. Koros
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Dr. Jeffrey C. Moore
Bioprocess Research and Development
Merck & Co., Inc.

Dr. Mark R. Prausnitz
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Date Approved: July 31, 2008

In Santa Barbara, 1933

Life is like riding a bicycle.
To keep your balance you must keep moving.

—Albert Einstein

Para mi familia
(To my family)

ACKNOWLEDGEMENTS

I would like to start by thanking my parents and family, without your guidance and unconditional support I would not have made it this far. Charlene Rincón-Rosenbaum, your company and counsel throughout these past ten years has been invaluable, *muchas gracias por todo*. I would also like to recognize the Bommarius lab members (current and past), you all made our lab what it is and what I will miss most. Thank you all for such stimulating conversations, cheering me up when needed, but most of all, thank you for putting up with me. Dr. Javier F. Chaparro-Riggers, thank you for assisting with ideas, training and the opportunity to work with you—your work ethic is imposing and contagious. Dr. James M. Broering and Dr. Karen M. Polizzi, thank you for training and for proof reading all my documents *ad nauseam*, without your assistance I probably would not have received funding nor my manuscripts submitted/accepted. Dr. Tracey L. Thaler and Dr. Phillip Gibbs, thank you for all your sacrifices in setting up the laboratory and instilling an excellent laboratory environment. Dr. Bernard L. W. Loo your youthful spirit is refreshing and I will miss you a lot. I would like to thank Janna Blum and Dr. Melanie Hall for continuously trying to improve our lab and for much needed assistance and advice with my research. Thomas A. Rogers and Michael J. Abrahamson, it was great doing research with you, thank you for helping me forget about research when needed most. I would like to recognize Victor Yeh for his contributions to the glucose 1-dehydrogenase project. Thanks to Aaron Englehart and Profesor Nicholas V. Hud for providing training, advice and access to the CD. I would also like to thank the estrogen-loaded Doyle lab for providing stimulating scientific conversations, access

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

AADH	Amino acid dehydrogenase
AmDH	Amine dehydrogenase
B-PER	Bacterial protein extraction reagent
CASTing	Combinatorial Active-Site Saturation Test
CD	Circular dichroism
C_x^y	Concentration at which x % activity remains after y minutes incubation at a defined temperature
DAPDH	*meso*-2,6-diaminopimelate dehydrogenase
DNA	Deoxyribonucleic acid
DRA	Direct reductive amination
ep-PCR	Error-prone PCR
FDH	Formate dehydrogenase
GC	Gas chromatography
GDH	Glucose 1-dehydrogenase
gDNA	Genomic DNA
HTS	High-throughput screening
IMAC	Immobilized-metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRA	Indirect reductive amination
ISM	Iterative site-saturated mutagenesis
kDA	kilodaltons
$k_{d,obs}$	Observed deactivation constant
K_m	Michaelis constant

LB	Luria Bertani
meso-DAP	*meso*-2,6-diaminopimelic acid
MIBK	Methyl isobutyl ketone
MSA	Maximum value of exposed surface area
NAD(H)	nicotinamide adenine dinucleotide (reduced)
NADP(H)	nicotinamide adenine dinucleotide phosphate (reduced)
NBD-Cl	4-chloro-7-nitrobenz-2-oxa-1,3-diazole
NBD-F	4-fluoro-7-nitrobenz-2-oxa-1,3-diazole
NMR	Nuclear magnetic resonance
NTA	Ni-nitrilotriacetic acid
OD_{600}	Optical density at 600nm
PCR	Polymerase chain reaction
PGA	Penicillin G acylase
PRALINE	Profile ALIgNEment
PTDH	Phosphite dehydrogenase
RFU	Relative fluorescent units
RSA	Relative solvent accessibility
SA	Solvent accessibility
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPPIDER	Solvent accessibility based Protein-Protein Interface iDEntification and Recognition
T_m	Apparent melting temperature
T_x^y	Temperature at which x % activity remains after y minutes incubation

SUMMARY

Enzyme catalysis is becoming increasingly important for the synthesis of chiral pharmaceuticals and fine chemicals. However, most enzymes have been optimized by nature to perform at specific conditions (i.e., pH, temperature, and medium), generally distinct from the desired process conditions. Advances in molecular biology and protein engineering have allowed researchers to tailor enzymes to meet ever-increasing demands.

The first goal of this work focused on the development of an amine dehydrogenase (AmDH) from a leucine dehydrogenase using site-directed mutagenesis. We aimed at reductively aminating a prochiral ketone to a chiral amine by using leucine dehydrogenase (LeuDH) as a starting template. This initial work was divided into two stages. The first focused mutagenesis to a specific residue (K68) that we know is key to developing the target functionality. Subsequently, mutagenesis focused on residues known to be in close proximity to a key region of the substrate (M65 and K68). This approach allowed for reduced library size while at the same time increased chances of generating alternate substrate specificity. An NAD+-dependent high throughput assay was optimized and will be discussed. The best variants showed specific activity in mU/mg range towards deaminating the target substrate.

The second goal of this work was the development of a thermostable glucose dehydrogenase (GDH) starting with the wild-type gene from *Bacillus subtilis*. Due to the high cost of nicotinamide cofactors an efficient regeneration system is vital to the development of a process utilizing redox enzymes (including an AmDH). GDH is able to carry out the regeneration of both NADH and NADPH cofactors using glucose as a substrate. We applied the structure-guided consensus method to identify 24 mutations that were introduced using overlap extension. 11 of the tested variants had increased thermal stability, and when combined a GDH variant with a half-life ~3.5 days at $65^{o}C$ was generated—a ~10^{6} increase in stability when compared to the wild-type.

The final goal of this work was the characterization of GDH in homogeneous organic-aqueous solvent systems and salt solutions. Organic media and salts are extensively used in organic synthesis to increase substrate stability and solubility as well

as enzyme stability. Engineered GDH variants showed increased stability in all salts and organic solvents tested. Thermal stability had a positive correlation with organic solvent and salt stability. This allowed the demonstration that consensus-based methods can be used towards engineering enzyme stability in uncommon media. This is of significant value since protein deactivation in salts and organic solvents is not well understood, making *a priori* design of protein stability in these environments difficult.

Lastly, future works for further improving LeuDH and GDH and potential applications are discussed.

CHAPTER 1

INTRODUCTION

Enzymes are biological macromolecules that catalyze biochemical reactions of value to Nature. After millions of years of evolution, enzymes have been superbly optimized. Most enzymes have evolved exquisite substrate specificity and selectivity towards the desired natural reactants and/or products under conditions that promote life: aqueous media and mild pH and temperature. Furthermore, due to the inherent low concentrations of substrates in biological systems, some enzymes have evolved to work nearly as fast as is physically possible with second-order rate constants that approach their rates of collisions with the substrate in solution ($\sim 10^9$ $s^{-1}M^{-1}$) (175).

Paradoxically, these evolved traits are what make enzymes both attractive and lacking for industrial applications. Some disadvantages of biocatalysts are: 1) enzymes are provided by Nature in only one enantiomeric form (L-amino acids), 2) enzymes are prone to inhibition, 3) enzymes tend to be extremely substrate specific and selective, 4) enzymes are coupled to their natural cofactors, 5) enzymes require narrow operation parameters and 6) enzymes display their highest catalytic activity in water (39).

Although progress has been achieved toward alleviating these concerns, unfortunately, when most enzymes are removed from their natural environment and placed in the presence of organic solvents, operating concentrations of reactants and/or products, and suboptimal (for the enzyme) but necessary conditions of pH and temperature, most natural (wild-type) enzymes will not perform at the required level. Due to the narrow operating conditions of the biocatalyst and the fact that biocatalysts often represent a significant portion of the process operating costs, once the target

synthetic route is known and the biocatalyst defined, process design and optimization efforts are focused on the biocatalyst (22, 40). However, with increasing substrate costs, advances in biocatalyst production, and an increase in available, suitable biocatalysts, bioprocess design will reach a stage where optimization will center on reaction properties and product output rather than on the biocatalyst (22, 104, 122). This is critical due to fervent competition, ever more complex molecules, environmental pressure and short development times due to patent expiry (22).

In a general sense, this work makes three main contributions to further develop and adapt biocatalysts for industrial applications. The first contribution is a further understanding of an amino acid dehydrogenase with potential application towards the synthesis of chiral amines. Amines are present ~3 times more frequently in synthetic drugs that in natural products (46) with tertiary aliphatic amine functionality the most frequent among drug molecules (58). To date, most synthesis of amines is carried out using chemical routes. Available enzymatic routes rely on racemic amines as starting substrates (10, 136), amino acids as amino donors (ω-transaminases) (146-150), or amine-boranes as reducing agents via chemoenzymatic approaches (2, 24, 25). However, enzymatic synthesis of chiral amines from prochiral ketones has not been exploited (65) and to date the best route to chiral amines from prochiral ketones relies on ω-transaminases (146). Unfortunately, ω-transaminases suffer from severe product inhibition.

In Chapter 2, we describe how a leucine dehydrogenase (LeuDH) is used as a starting scaffold for engineering a novel amine dehydrogenase. Site-saturated mutagenesis was done at two key positions identified after reviewing the catalytic mechanism and available X-ray structures. The screening of over 3700 variants required the development of a high throughput assay. This proved to be a challenge

since our product is a primary amine and amino acids inherent in the screening process (cell lysate), ammonia in the buffer and an aqueous environment rendered most techniques ineffective. Ultimately, an NAD+-dependent fluorescence assay was chosen for screening (163).

The second contribution of this work focuses on the improvement of biocatalyst stability at high temperatures via protein engineering. In Chapter 3, a glucose 1-dehydrogenase (GDH) is engineered using a structure-guided consensus method (120, 156, 168). GDH is an invaluable enzyme used for the recycle of nicotinamide cofactors (NAD(P)(H)), coenzymes necessary for enzymatic redox reactions. Due to the high costs of NAD(P)(H), stoichiometric amounts of cofactors cannot be used and thus a robust, stable cofactor regeneration system is indispensable towards the use of enzymes at a technical scale.

Improvement of protein stability by utilizing the consensus sequence (amino acid sequence alignments of homologous genes) was first employed by Steipe et al. (156) and applied primarily to antibodies (155, 156). The ease of use and proven success rate of consensus-based tools (85, 86, 120, 168) (> 30% success in identifying stable mutations—significantly higher than alternative methods such as directed molecular evolution) make them an attractive approach to stabilizing enzymes. Sequence alignment tools are used to generate a consensus sequence and this sequence is then used to identify recurring amino acids at each position. We then used a set of criteria based on the protein structure to sieve through the amino acids identified via the consensus sequence. For GDH, 9.2% of the protein was mutated and ~46% of the chosen mutations incurred increased stability. This approach will be most useful for complex enzymes (120) where arbitrary randomization of the gene may render the protein inactive, enzymes whose substrates/products are expensive or

difficult to handle and not amenable to high throughput screening and finally as a precursor to the use of directed molecular evolution where stable scaffolds have been shown to improve library results by increasing tolerance to mutagenesis (11, 13).

Chapter 4 characterizes four thermostable GDH variants generated via the structure-guided consensus concept in homogeneous aqueous-organic media and salt solutions. The four variants selected for characterization had significant differences in kinetic stability (as verified by half-lives). Stability in the presence of high salt concentrations and organic solvents would allow enzymes to be employed for transformations of compounds not readily soluble at low temperature or in pure aqueous systems.

In addition to characterizing thermostable GDHs in atypical media, these studies allowed us to test the applicability of the structure-guided consensus concept towards developing stable enzymes towards other denaturants. As previously observed, a good linear correlation between temperature stability and organic solvent stability was observed (31, 113). Furthermore, thermostable variants demonstrated increased salt stability regardless of salt-type and concentration. This is invaluable to protein engineers since deactivation in salt solutions and organic solvents is not well understood, rendering *a priori* design of stable enzymes in this media difficult.

Recommendations towards further protein engineering efforts and conclusions are summarized in Chapter 5. In particular, future mutagenesis on leucine dehydrogenase towards an amine dehydrogenase should focus on substrate-interacting residues and alternative starting protein scaffolds are also considered. Future efforts for further stabilization of glucose 1-dehydrogenase should focus at the protein-protein interface and an initial attempt via engineered disulfide bonds is discussed. Finally, the use of the structure-guided consensus approaches towards stabilization of enzymes relevant to biomass conversion to value-added products is discussed.

CHAPTER 2

DEVELOPMENT OF A NOVEL DEHYDROGENASE

2.1 Introduction

Nature has customized enzymes to meet specific demands which often are incompatible with those desired for industrial processing. An example of such a desired (but lacking) functionality is the transformation of prochiral ketones into chiral amines via reductive amination. Chiral amines are important functional elements in a multitude of biologically active pharmaceuticals and their synthesis is of importance (46, 58). For example, Ghose et al. (46) noticed that tertiary aliphatic amine is the most frequent functional group in drug molecules while Henkel et al. (58) found that amine functionality is found two and three times more frequently in synthetic and drug compounds when compared to natural compounds, respectively.
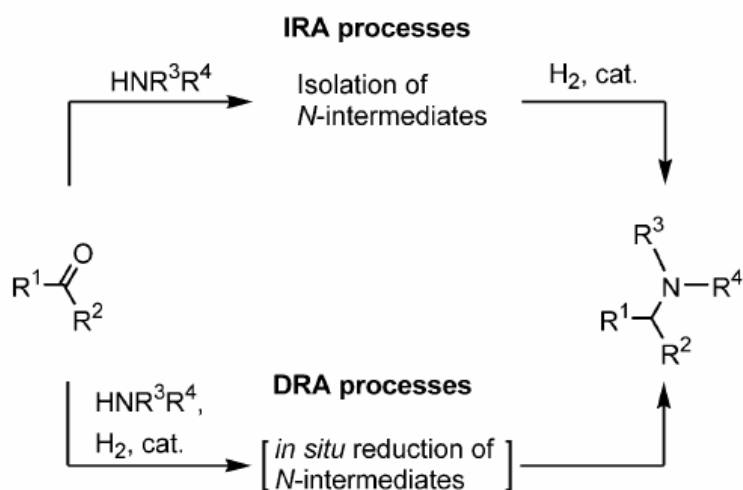
Figure 2.1. General approaches for reductive aminations. From Tararov 2004 (160).
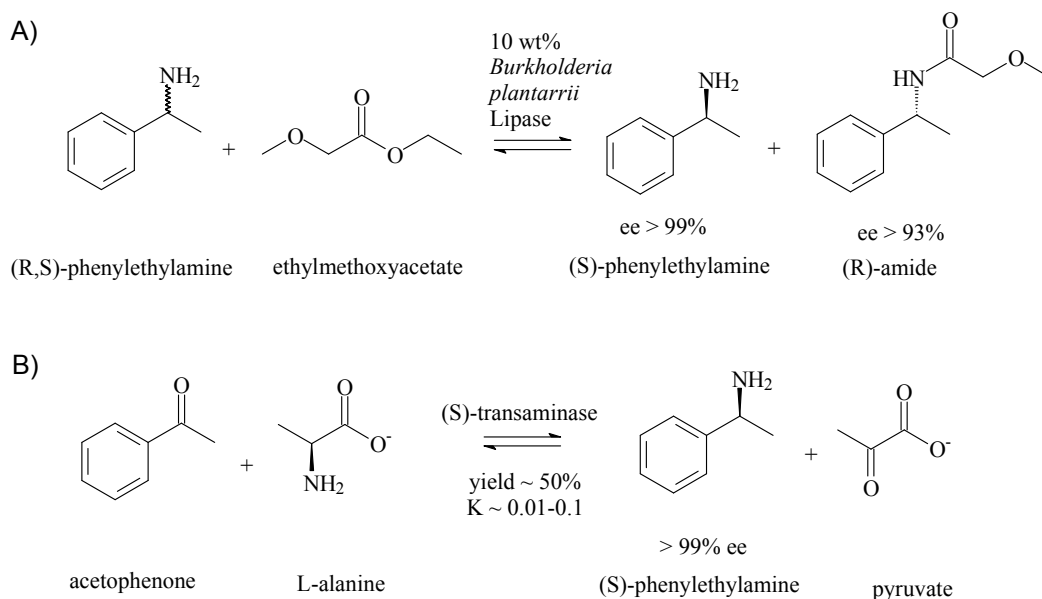
Figure 2.2. A) BASF lipase-based synthesis of enantiomerically pure amines. From Balkenhohl 1997 (10). B) Amination of acetophenone by ω-transaminase (149, 150).

Chemically, amine incorporation is achieved with the hydrogenation of imines or enamine derivatives using late transition metal complexes based on chiral P-ligands, after an initial condensation of the ketone with an appropriate amine (159). This approach requires a two step process. An alternative approach utilizes a reducing agent to directly treat the carbonyl and amine substrates. These methodologies are known as indirect (IRA) and direct reductive amination (DRA), respectively (Figure 2.1).

BASF (Germany) developed a lipase-based route to enantiomerically pure amines by using ethylmethoxyacetate to stereospecifically acetylate one enantiomer. The products (i.e., R-amide and S-amine) can be recovered via distillation at high chemical and optical purities (Figure 2.2A) (10, 136). Although lipase-assisted kinetic resolution of racemic amines has been successful in the development of chiral amines, the same product can be obtained via reductive transamination of ketones (Figure 2.2B) (146-150). The asymmetric synthesis of chiral amines with ω-
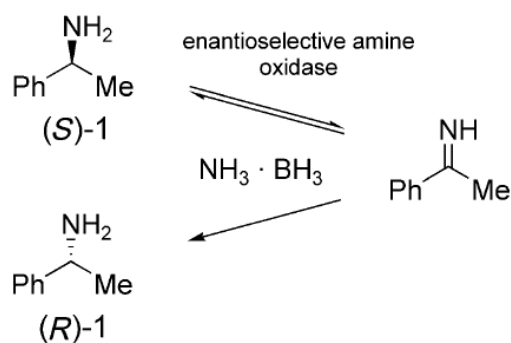
Figure 2.3. Deracemisation of amines with an amine oxidase in combination with ammonia borane. From Carr 2003(25).

transaminases has a significant advantage over kinetic resolution in that it does not need a racemic amine. However, the ω-transaminase reaction is very difficult to perform due to unfavorable thermodynamic equilibrium as well as severe product inhibition, limiting net conversion to ~ 50% (147, 149). Significant advances have been achieved by Turner's group (University of Edinburgh, UK) using the less exploited deracemisation of amines (2, 24, 25) (Figure 2.3). Utilizing a chemo-enzymatic approach using ammonia borane and an amine oxidase mutant from *Aspergillus niger*, Alexeeva et al. (2) were able to achieve a one-pot deracemisation yielding an amine with > 93% ee and 73% yield. Subsequent studies demonstrated that developed amine oxidase mutants had broad substrate specificity and enantioselectivity (25) and were also applied to the deracemisation of cyclic secondary amines (24). Even though amine-boranes posses interesting properties that can overcome the disadvantages of other popular hydridic reducing agents (i.e., sodium borohydride and sodium cyanoborohydride) such as low stability and high toxicity (1), like the BASF lipase-based process this approach requires a source of racemic amines or the initial conversion of a keto acid or an aldehyde to an imine. Thus, this approach does not allow the direct reductive amination of a prochiral ketone to a chiral amine.

7

A) Oxidoreduction of leucine dehydrogenase.
L-leucine, NAD$^+$ + H$_2$O $\rightleftharpoons$ NH$_4^+$ + 4-methyl-2-oxopentanoic acid + NADH + H$^+$

B) Target amine dehydrogenase.
1,3-dimethylbutylamine, NAD$^+$ + H$_2$O $\rightleftharpoons$ NH$_4^+$ + methyl isobutyl ketone + NADH + H$^+$
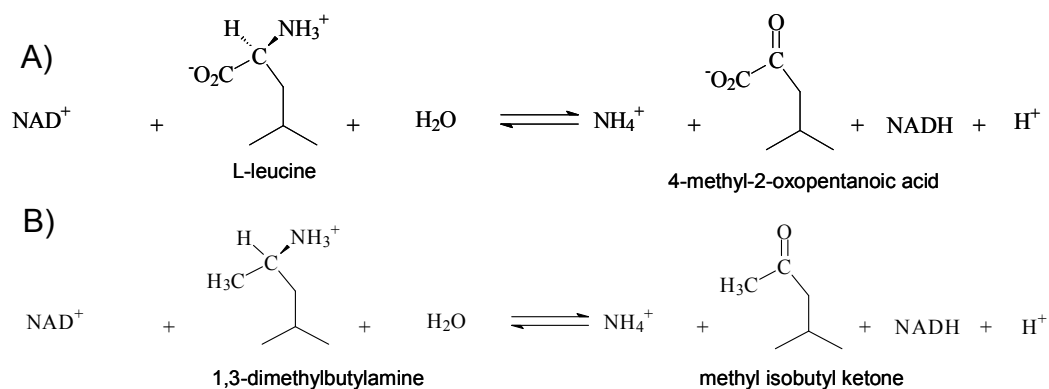
Figure 2.4. A) Oxidoreduction of leucine dehydrogenase. B) Target amine dehydrogenase.

To date, Itoh and coworkers (Toyama Prefectural University, Japan) (65) have been the only researchers to have reported an NAD+-dependent amine dehydrogenase. This enzyme was found in cell extract from *Streptomyces virginiae* IFO 12827 and the authors detect the reductive amination of keto alcohols, keto acids, ketones, and aldehydes. Unfortunately, the enzyme responsible for this activity in the cell extract has not been purified nor further characterized and the DNA and amino acid sequence is unknown.

In this work we intend to develop an NAD(H)-dependent amine dehydrogenase starting from an amino acid dehydrogenase (AADH)—specifically a leucine dehydrogenase (LeuDH) from *(Geo)Bacillus stearothermophilus* strain 10. AADHs (EC 1.4.1.X) are part of the oxidoreductase superfamily (21) that catalyze the removal of the amino group from an L-amino acid, forming the corresponding keto acid with the concomitant reduction of NAD(P)$^+$ (Figure 2.4A). This LeuDH was chosen due to: organism accessibility, known gene sequence, high identity/similarity with respect to the available crystallized structure, available kinetic parameters, and previous success in overexpression and purification (108-110). If successful, the

developed AmDH will reductively aminate methylisobutylketone (MIBK) to (R)-1,3-dimethylbutylamine (Figure 2.4B).

LeuDH is an octameric enzyme (~320 kDa), first isolated from *Bacillus cereus* (134) and crystallized by Baker et al. from *Bacillus sphaericus* (8). The chemical mechanism en route from the keto function to the amine group passes through a carbinolamine and an imine intermediate before the reductive step to the amine group Figure 2.5) (20, 140). Importantly, the key redox chemistry step, involving NADH cofactor, occurs only between the imine and the amine. However, while the imine of an α-keto acid possesses reasonable stability as its carboxyl group is geometrically fixed by a lysine (K68) in the enzyme (Figure 2.5), an imine of a ketone is very unstable. Thus, binding and stabilization of the imine intermediate is a key step towards the creation of an amine dehydrogenase from LeuDH.

A review by Morley et al. (105) compiled results obtained from enzyme mutation studies and noticed that to increase catalytic activity ("catalytic
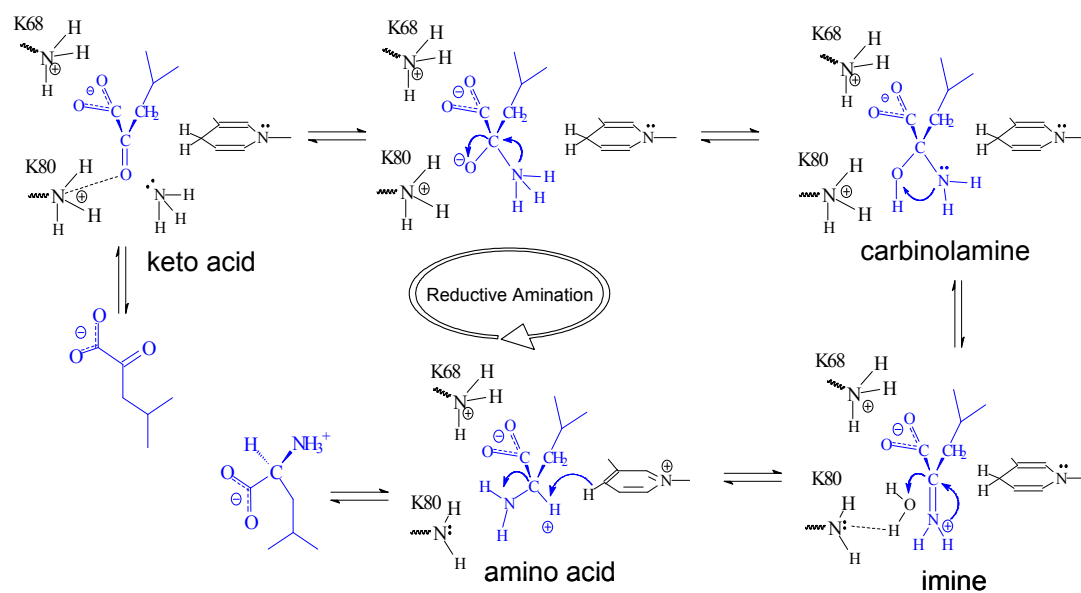
Figure 2.5. LeuDH chemical mechanism.

9

promiscuity"), mutations < 10 Å of the active site proved more effective than those on residues more distant to it. As a result, the authors concluded that in these cases, random mutagenesis targeting the entire enzyme might not be the best approach to find improved variants. With this in mind—and after careful analysis of the mechanism and available crystal structures of amino acid dehydrogenases—it was decided to focus on those amino acids that were in the vicinity of the carboxyl being replaced.

2.2 Materials and Methods

2.2.1 gDNA preparation, gene isolation and overexpression

Geobacillus stearothermophilus strain 10 was kindly donated by Assistant Professor Bert C. Lampson from East Tennessee State University. *Geobacillus stearothermophilus* was grown at $60^{o}C$ in modified Luria Bertani (mLB). mLB contains: 10 g tryptone, 5 g of yeast extract, 5 g of NaCl, 1.25 ml NaOH (10% w/v), 1 mL nitrilotriacetic acid (1.05 M), 1 mL $MgSO_4$ (0.59 M), 1 mL $CaCl_2$ (0.91 M), 1mL $FeSO_4$ (0.04 M) and 995 mL of H_2O. Genomic DNA (gDNA) was obtained using method B described by Mehling et al. (101). The gene was isolated from the gDNA via polymerase chain reaction (PCR) using primers 5'-ATGGAATTGTTCAAATATATGGAAACTTACGATTATGAGC-3' (forward) and 5'- TTAAAATGCGAAACTTCTGCTCTTCCATCGTT-3' (reverse). The reverse primer was chosen to anneal downstream of coding region of the gene due to the rich GC content at the 3' end of the gene. The gene was later isolated from the larger DNA fragment and later cloned into the pPROTet expression vector via the *KpnI* and *PstI* restriction sites and into the pET17b and pET28a expression vectors via the *NdeI*, *HindIII* restriction sites. The pPROTet and pET constructs were tranformed into

10

Escherichia coli (*E. coli*) DH5αPRO and BL21(DE3), respectively. DH5αPRO, containing the construct (in pPROTet) was grown in LB at 37°C to an OD of 0.5. At this time the culture was induced with 1 mg/mL of autoclaved tetracycline (anhydrotetracycline) and incubated overnight at 37°C. Analogously, BL21(DE3) cells containing the construct (in pET28a) were grown in LB at 37°C to an OD of 0.5 and induced with 0.1 mM IPTG and incubated overnight at 37°C. Cultures were then pelleted via centrifugation.

2.2.2. Protein purification

LeuDHs cloned into pET28a contained a N-terminal 6xHis tag for ease of purification. The cell pellet was resuspended in 50 mM phosphate pH 8.0 with 20 mM imidazole and 300 mM NaCl. Cells were then lysed via sonication and cell debris removed via centrifugation. Clarified cell extracts were then purified via immobilized-metal affinity chromatography (IMAC) using a Ni-nitrilotriacetic acid (NTA) matrix (Qiagen). Clarified supernatant was loaded onto the resin and washed with 50 mM phosphate pH 8.0 with 300 mM NaCl and 50 mM imidazole. Pure protein was eluted with 50 mM phosphate pH 8.0 with 300 mM NaCl and 250 mM imidazole. A representative SDS-PAGE can be seen below (Figure 2.6).

L = Protein Markers
1 = Lys68Ala - lysate
2 = Lys68Cys - lysate
3 = Lys68Gly - lysate
4 = Lys68Met - lysate
5 = Lys68Ser - lysate
6 = Lys68Thr - lysate
7 = Lys68 - lysate
8 = Lys68Ala - pure
9 = Lys68Cys - pure
10 = Lys68Gly - pure
11 = Lys68Met - pure
12 = Lys68Ser - pure
13 = Lys68Thr - pure
14 = Lys68 - pure
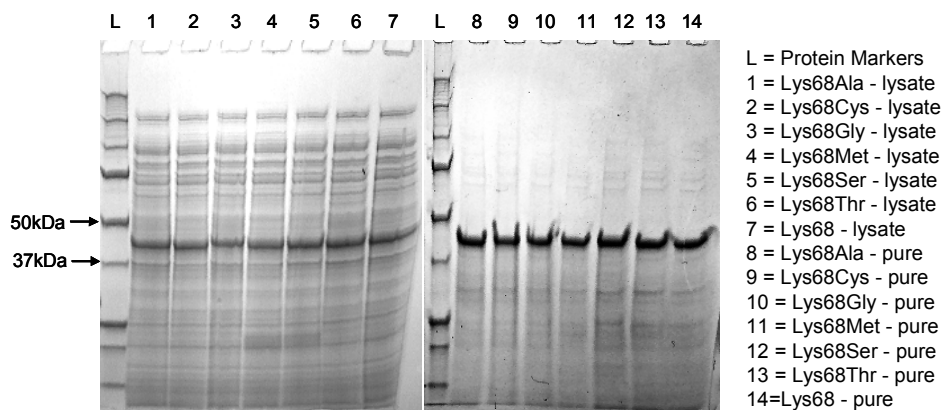
Figure 2.6. Representative SDS-PAGE of LeuDH wildtype (lane 7 and 14) an variants. LeuDH with N-terminal His-tag and thrombin cleavage site has a MW ~43kDa.

2.2.3. Spectrophotometric assay

Purified samples or cell lysate were tested for activity spectrophotometrically via absorbance at 340nm for NAD(H). For the reductive amination, 1mM 4-methyl-2-oxopentanoic acid, 200 mM NADH in 1 M NH_4Cl/NH_4OH pH 9.7 was used. A low concentration (1 mM) of 4-methyl-2-oxopentanoic was required due to its absorbance at the analytical wavelength. For the oxidative deamination, 10 mM L-leucine, 2 mM NAD+ in 0.1 M glycine at pH 10.0 was used. For pH profile, pH was adjusted with NaOH or HCl prior to assay. All assays were carried out at $25^{o}C$ unless otherwise specified.

2.2.4. Mutagenesis and library generation

For the first round position Lys68 was saturated using degenerate NNK primers: 5'-CGCGGCATGACGTACNNKAACGCGGCCGCCGGCCTTAATTTA-3' and 5'- TAAATTAAGGCCGGCGGCCGCGTTMNNGTACGTCATGCCGCG-3'. For the second round of mutagenesis Met65 and Lys68 were saturated using NNK primers: 5'- CGCCTCGCCCGCGGCNNKACGTACNNKAACGCGGCCGCCGGC-3' and 5'-GCCGGCGGCCGCGTTMNNGTACGTMNNGCCGCGGGCGAGGCG -3'. Primers were designed using the guidelines provided by Stratagene (Cedar Creek, Texas) for its QuikChange® Site-Directed Mutagenesis Kit. Mutations were introduced using the degenerate primers via overlap extension polymerase chain reaction (60, 63).

For position Lys68, the degenerate gene was cloned into pPROTet, transformed into DH5αPRO and random colonies were selected and sent for sequencing until all 19 variants were isolated. Variants were stored for further

characterization. Interesting variants were then cloned into pET28a to insert an N-terminal 6xHis-tag for protein purification and further characterization.

For the second round of site-saturated mutagenesis at position Met65 and Lys68, the degenerate gene (randomized cassette) was cloned into pET17b. BL21(DE3) cells were transformed and random colonies (20 total) were selected and sent for sequencing to test library diversity (Figure 2.7).
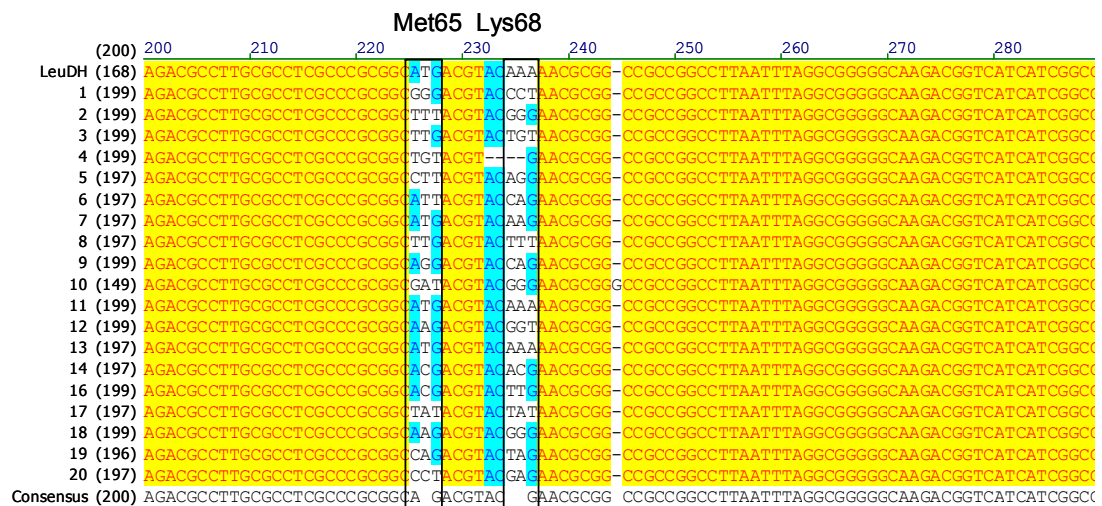
Figure 2.7. DNA sequence alignment of 20 random colonies to test library diversity.

2.2.5. Screening of LeuDH mutants and library

For the first round, wild-type LeuDH and 19 variants (site-saturation at position Lys68) were screened via gas chromatography (GC). 200 mL cultures were grown, pelleted, lysed via sonication and cell lysate used for screening. 2 mL reactions were carried out in 4 mL glass vials overnight. Amination was carried out with 8 mM 4-methyl-2-pentanone, 8 mM NADH in 0.7 M NH₄Cl/NH₄OH pH 8.5. For deamination 7 mM 1,3-dimethylbutylamine, 7 mM NAD+ in 0.1 M glycine pH 10.0. Reaction mixtures were incubated overnight. pH was adjusted > 11 prior to

extraction to facilitate isolation of amine. Extraction was carried out with 1 mL of ethyl ether and samples were then loaded to a GC for assay.

For the second round, after generating the library, 3752 colonies were picked using a Genetix Systems QPix colony picker (Hampshire, UK). This accounts for a 97.4% (equation 2.1) probability of complete library screening. Colonies were transferred to 96-well plates containing 250 μL of LB. Most liquid handling was carried out using a MultiDrop® 384 from Themo Electron Corporation (Milford, MA). Plates were grown overnight at 37°C. Plates were then replicated into 250 μL of MagicMedia^{TM} expression medium (Invitrogen, San Francisco, CA). MagicMedia^{TM} is an expression medium designed to increase yield of recombinant proteins in T7-regulated *E. coli* systems without the need of adding inducer, simplifying the overexpression process (Appendix, Figure A2.1). Plates were incubated and agitated overnight at 37°C. Cultures were then pelleted via centrifugation and store at -80°C until screening.

$$N = \frac{\ln(1-P)}{\ln\left(1-\frac{1}{n}\right)}$$

$N =$ number of clones required $\qquad$ (2.1)

$P =$ probability desired

$\frac{1}{n} =$ fractional proportion of each possible codon

Pelleted cells were resuspended in 50 μL of bacterial protein extraction reagent (B-PER-Pierce, Rockford, IL). B-PER is a cell lysis solution that contains a nonionic detergent in 20 mM Tris·HCl pH 7.5, eliminating the need for mechanical cell disruption or freeze-thaw cycles. 20 μL of resuspended, lysed cells were

transferred to two, distinct black 96-well plates. These plates were later screened via NAD+-dependent fluorescence—one plate was used as background.

LeuDH variants were screened using an adaptation of the plate assay developed by Tsotsou et al. (163) based on work by Kaplan et al. (73) and Lowry et al. (91). This assay is based on NAD+. NAD+ is a product during the reductive amination of MIBK (Figure 2.4B). After the reaction, the pH is lowered using HCl were the remaining NADH is destroyed. Then pH is increased and sample incubated for 2.5 hours. During this time the NAD+ is transformed to a fluorescent, stable compound (73, 91, 163).

To the 20μL cell lysate, 180 μL of reaction mixture was added. This resulted in a total concentration of 8 mM MIBK, 1 mM NADH in 500 mM NH_4Cl/NH_4OH pH 9.7. The reaction was allowed to proceed for at least 2 hours at room temperature (~23°C). pH was then lowered to ~1.0 using 30 μL of 6.0 N HCl. Plates were allowed to sit for at least 15 minutes in the dark. 80 mL of 10.0 N NaOH was then added and plates were then incubated for an additional 2.5 hours in the dark. Fluorescence was then measured using a Gemini Spectramax Microplate Spectrofluorometer (Molecular Devices, Sunnyvale, CA). Plates were excited at 360 nm and the emission measured at 455 nm (163).

2.3 Results and Discussion

2.3.1. Identification of key residues in LeuDH

Due to the lack of a crystal structure of LeuDH with bound substrate (leucine or 4-methyl-2-oxopentanoic acid) it was decided to use the crystal structure of phenylalanine dehydrogenase from *Rhodococcus species* (PheDH) (20). The
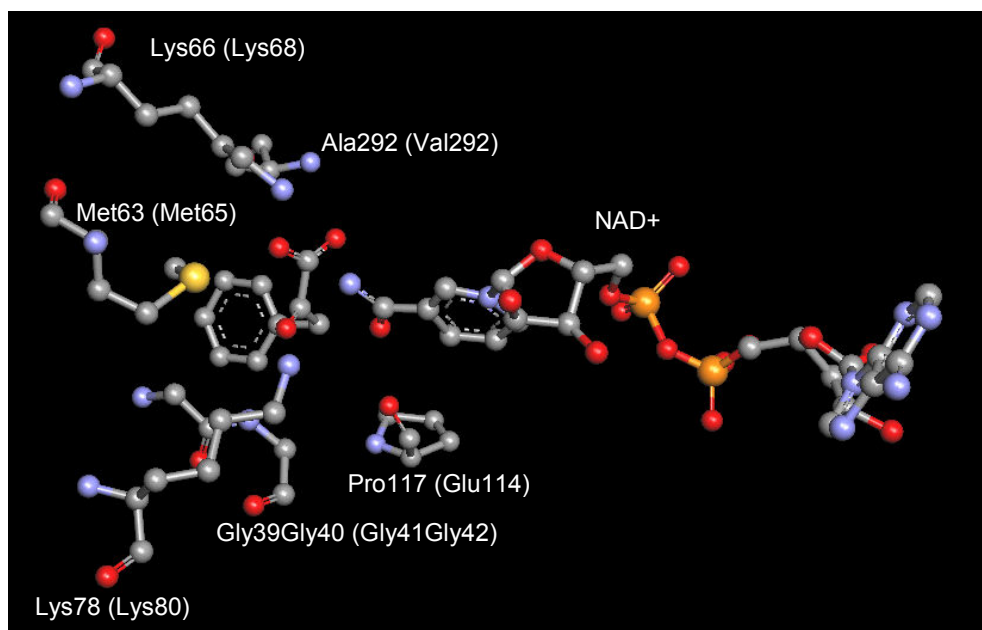
Figure 2.8. Residues ≤ 3.5 Å from the carboxylic acid group of phenylpyruvate of phenylalanine dehydrogenase (PDB ID 1BW9). Residues in parenthesis represent the corresponding amino acids in LeuDH.

available structure for LeuDH is from *Bacillus spahericus* (PDB ID 1LEH) and has an amino acid identity of 80% with the LeuDH from (*Geo*)*Bacillus stearothermophilus*. The available crystal structure for PheDH (PheDH 1BW9) has bound phenylpyruvate (Figure 2.8). LeuDH from *(Geo)Bacillus stearothermophilus* and PheDH from *Rhodococcus species* share 34% amino acid identity. Met63 and Lys66 were identified as key residues in PheDH, their counterparts Met65 and Lys68 were determined via amino acid sequence alignment and crystal structure comparisons in LeuDH. Lys68 is known to stabilize the α-carboxyl group of the α-keto acid by an ionic interaction—a group that will be replaced by a methyl in the target substrate. By focusing on the active site—specifically on the amino acids that face the carboxyl being replaced to a methyl group on our substrate—we intended to minimize library size while at the same time increase our chances of successfully generating the desired activity. Lys80 was not chosen for mutation since this residue is known to

form a hydrogen bond with the carbonyl group of the α-keto acids (20, 140). The ketone is expected to bind in the same orientation.

2.3.2. pH optimization for amination and deamination of wild-type LeuDH

The pH optimum for the amination and deamination of His-tag, wild-type LeuDH was determined to be 9.2 and 10.5 for amination and deamination, respectively at 25°C (Figure 2.9). Interestingly the pKa for free ammonia, the amino group for leucine and for the ε-amino in the side chain of lysine are 9.2, 9.6 and 10.5, respectively. Thus, the content of free ammonia, the protonated α-amino group of L-leucine and the protonated ε-amino groups in Lys68 and Lys80 are critical for the substrate stabilization and catalysis. Vanhooke et al. (20) carried out pH studies on L-phenylalanine dehydrogenase from *Rhodococcus* and determined that in the direction of L-phenylalanine synthesis two enzymic groups with pKas of 9.4 and 8.1 are critical for catalysis. The authors assigned the higher pKa value to Lys78 (Lys80 in LeuDH), a residue known to form a strong hydrogen bond (when protonated) with the carbonyl
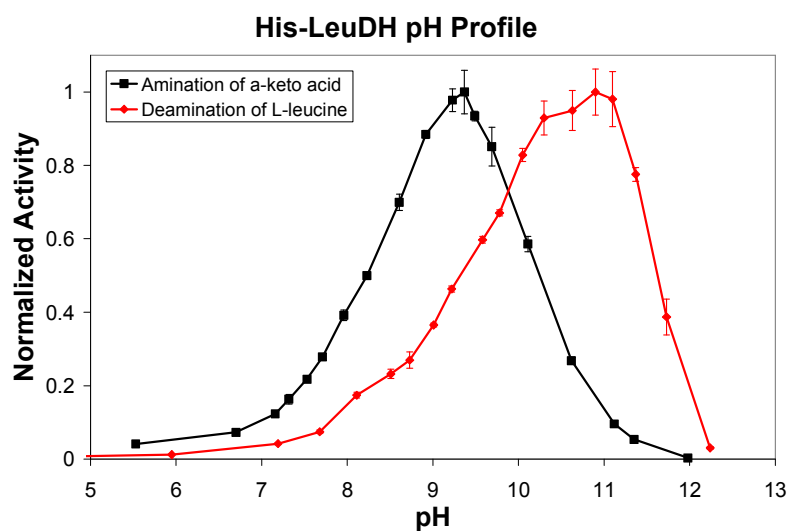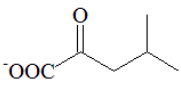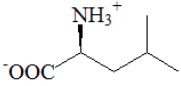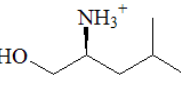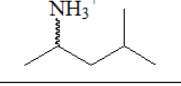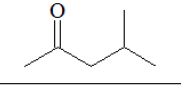
Figure 2.9. pH profile for wild-type, His-tag LeuDH from *(Geo)Bacillus stearothermophilus.*

oxygen of the α-keto acid. Similarly, Sekimoto and coworkers identified Lys68 (138) and Lys80 (140) via pH studies and determine that both need to be protonated during binding and catalysis (Figure 2.9).

Table 2.1. Oxidative and reductive profile of His-LeuDHs.

Substrate	Lys68Ala	Lys68Cys	Lys68Gly	Lys68Met	Lys68Ser	Lys68Thr	Lys68
Reductive Amination[a]	Activity U/mg	U/mg	U/mg	U/mg	U/mg	U/mg	U/mg
(α-keto acid structure)	0.23	1.91	1.85	1.99×10^{-3}	0.64	4.1	48
Oxidative Deamination[b]	Activity mU/mg	mU/mg	mU/mg	mU/mg	mU/mg	mU/mg	mU/mg
(NH_3^+ amino acid structure)	0.07	0.16	0.20	0.08	0.01	0.38	8.9×10^3
(HO NH_3^+ structure)	7.0	8.8	0.6	ND	0.02	8.0	6.0
(NH_3^+ amine structure)	1.78	1.40	1.20	8.28	3.39	0.36	0.11
Reductive Amination	Activity mU/mg	mU/mg	mU/mg	mU/mg	mU/mg	mU/mg	mU/mg
(ketone structure)	0.10/0.14[c]	0.10/0.15	0.09/0.13	0.21/0.28	0.14/0.19	0.03/0.04	0.05/0.1

[a]Amination was carried out at 25°C 1M NH₄Cl/NH₄OH pH 9.4 (for details see Materials and Methods)
[b]Deamination was carried out at 25°C 0.1M glycine pH 10.0
[c]ammonium chloride buffer/ammonium formate buffer

2.3.3. Characterization of LeuDH variants: site-saturation at Lys68

After acquiring all 19 variants of LeuDH, they were screened for amination and deamination activity via GC. Of the tested variants none were detected to aminate MIBK, while six (i.e., Lys68Ala, Cys, Gly, Met, Ser, and Thr) were detected to deaminate 1,3-dimethylbutylamine. These variants were then cloned into an N-terminal His-tag containing vector for purification and further characterization. Once

purified enzyme was obtained, variants were characterized to determine their ability to aminate and deaminate various substrates (Table 2.1). Of the six variants Lys68Met and Lys68Ser seem to be the most promising since deamination of 1,3-dimethylbutylamine was detected in mU/mg range (i.e., 8.3 and 3.4 mU/mg, respectively).

2.3.4. Development of a high throughput assay

In an attempt to further improve the activity of LeuDH toward amination we decided to randomize both Met65 and Lys68 simultaneously. The randomization of these two positions around the area of interest allows for potential synergistic conformational effects arising from unpredictable interactions that cannot be observed by single-site mutagenesis (129) (Figure 2.10). To properly screen this library more than 3000 variants need to be screened.

En route to finding an appropriate screening approach many challenges were encountered. The biggest challenge was finding a screening method selective toward primary amines but not amino acids. Benzofurazan fluorescent compounds, such as 4-chloro-7-nitrobenz-2-oxa-1,3-diazole (NBD-Cl) and 4-fluoro-7-nitrobenz-2-oxa-
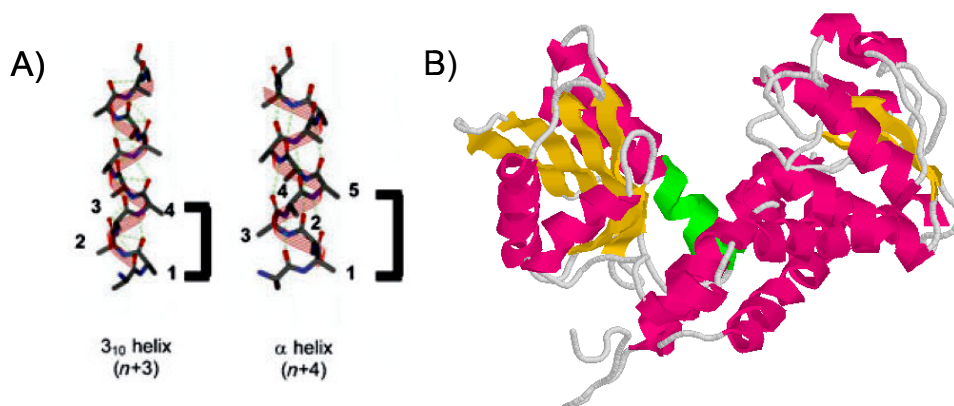
Figure 2.10. A) Orientation of side chains based on secondary structure (129). B) Location of Met65 and Lys68 (shown in green)

1,3-diazole (NBD-F) have been previously used for high throughput experiments (57). NBD-F has the advantage of being ~500 times reactive than NBD-Cl (103), but it is ~800 times more expensive (57). Unfortunately, in an aqueous environment, NBD-F hydrolyzes to form 4-hydroxy-7-nitrobenzo-2-oxa-1,3-diazole, a fluorescent product (103) and ammonia also reacts with NBD-Cl/F rendering this technique inappropriate for LeuDH.

Ultimately, an NAD(P)(H) fluorescent assay was chosen for screening. This assay was reported by Tsotsou et al. (163) based on initial work by Lowry et al. (91) and Kaplan et al. (73). In this work, produced NAD+ was converted to a fluorescent product under an alkaline environment. Sensitivities in the 0.01 μM for the NADP+ fluorophore have been reported (89). The effect of B-PER, cell lysate, volume of acid and base, and buffer was tested to minimize background, improve ease of handling and test sensitivity (see materials and methods for final conditions used). Our first goal was to be able to achieve reaction, pH changes and measurements in a single plate. Previous work carried out reaction and pH changes in distinct microtiter plates (total of three) (163). With this in mind, B-PER and cell lysate (from BL21(DE3)) were tested to quantify potential effects on background (Figure 2.11). 10 μL of B-PER with (~862 μg/mL of BL21(DE3) cell lysate) and without cell lysate was tested. B-PER with cell lysate resulted in a slightly higher background, this is observed in a higher intercept (increased background) and a reduction in the slope values (reduction in sensitivity). pH is also known to influence stability of nicotinamide cofactors (69, 90, 164, 177). Qualitatively, NAD(P)H is more stable in acidic conditions while NAD(P)+ is more stable in alkaline solutions (90, 177). In this experiment we aimed to elucidate the required HCl and NaOH to destroy NADH and produce the alkaline fluorescent NAD+ product. To do this, different volumes of 6.0 N HCl and 10N

20
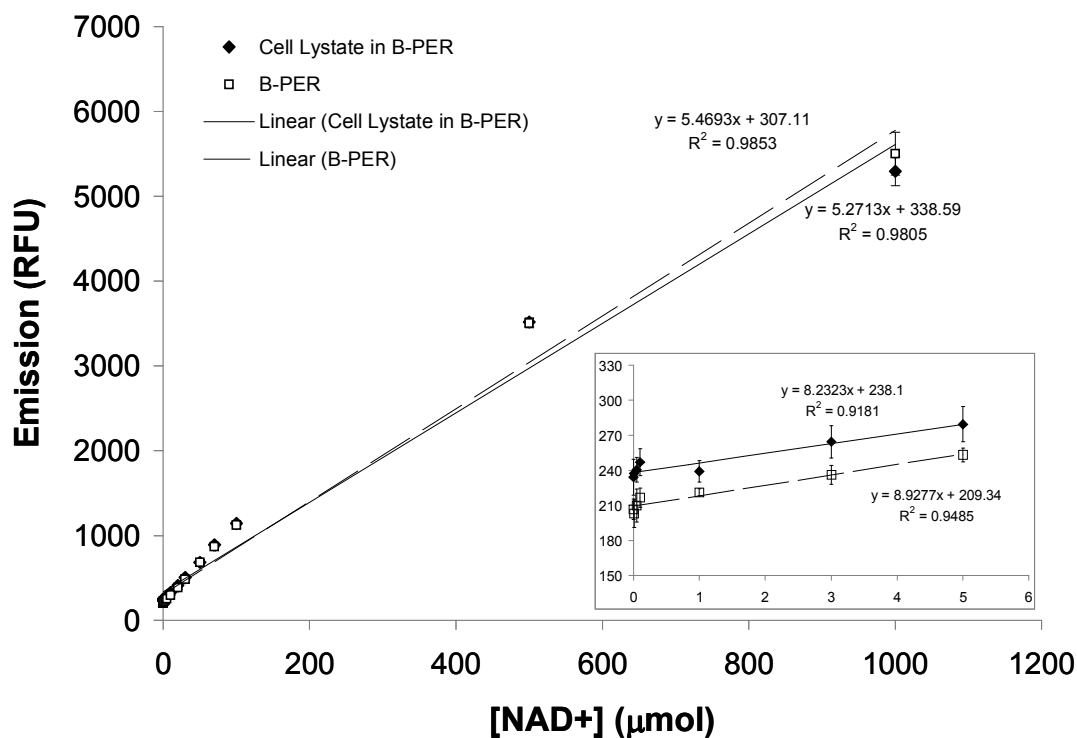
Figure 2.11. Effects of B-PER and cell lysate on fluorescence assay.

NaOH were added and fluorescence measured (Figure 2.12). It is evident that at least 10μL of 6.0 N HCl is necessary for complete destruction of NADH. This is observed with a sharp reduction in fluorescence at volumes greater than 10μL of 6.0 N HCl. A leveling in fluorescence is observed for all volumes (concentrations) of NaOH with values ranging from 91 to 302 relative fluorescent units (RFU). For the final assay 30 μL of 6.0 N HCl and 80 μL of 10.0 N NaOH were used. A higher concentration of 10 N NaOH was used since it has been observed by Kaplan et al. (73) that competing, unwanted reactions can be minimized at higher concentrations of NaOH. Stability of NAD(H) was also studied in 0.5 M NH$_4$Cl/NH$_4$OH pH 9.7 (Figure 2.13). It is evident that incubation overnight in 0.5 M NH$_4$Cl/NH$_4$OH pH 9.7 causes an increase in background and a decrease in sensitivity. No significant changes were observed between 1 hour and 4 hour incubation. For the final fluorescence assay, samples were incubated for at least 2 hours.

Figure 2.12. Effects of HCl and NaOH on fluorescence assay.
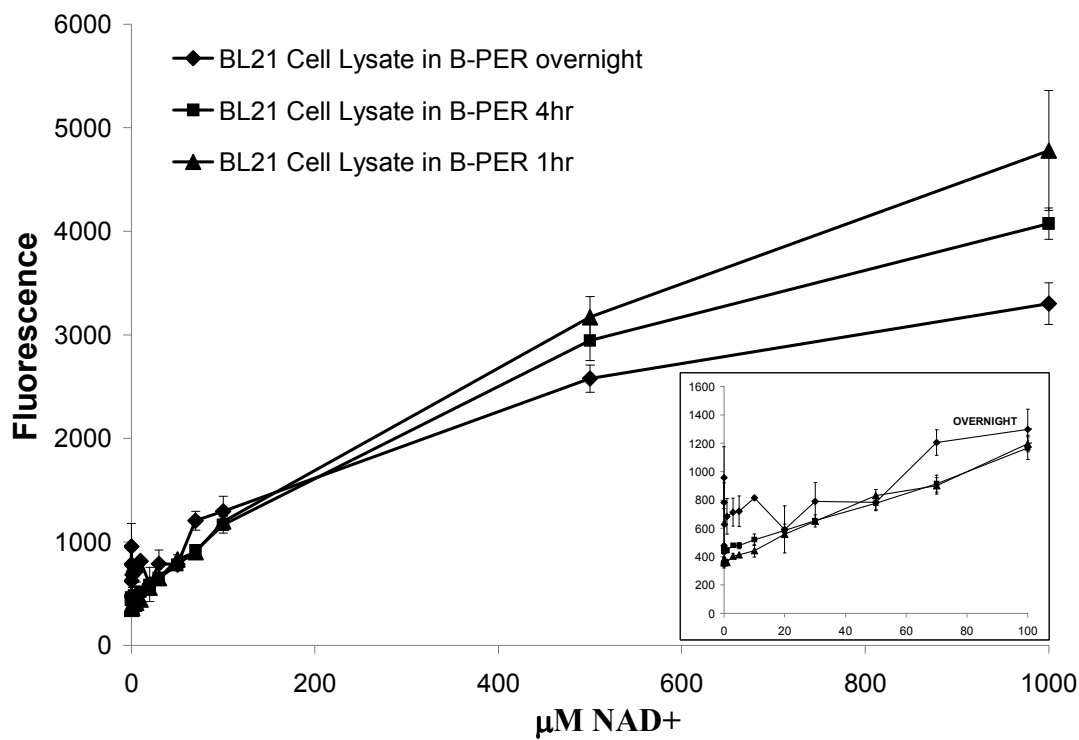
Figure 2.13. Effects of incubation time in reaction buffer on fluorescence assay.
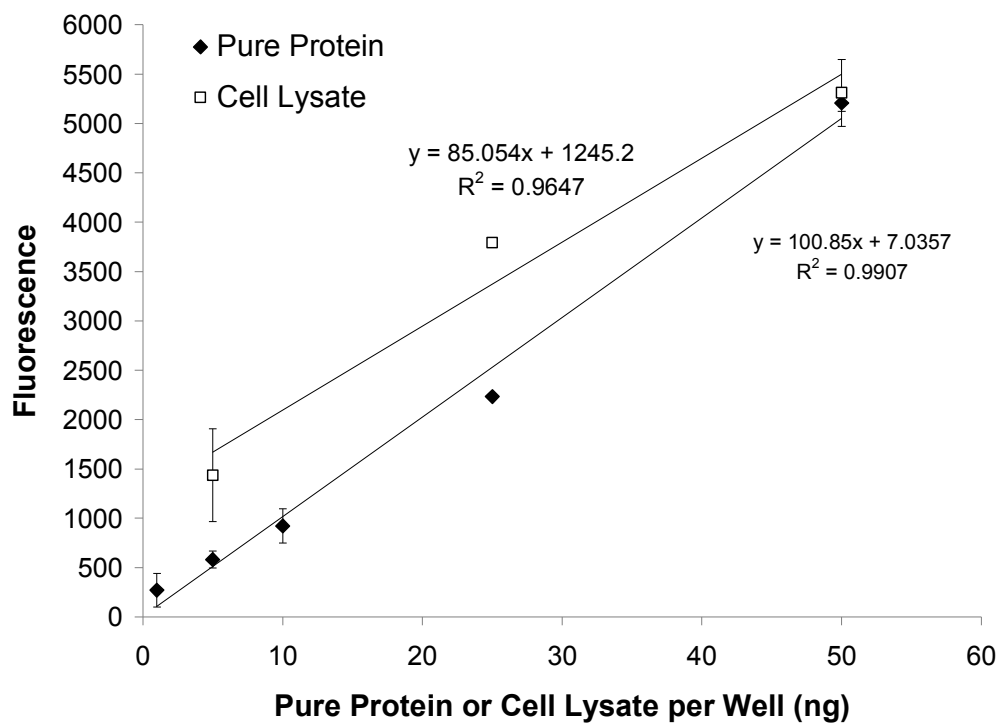
Figure 2.14. Sensitivity of fluorescence assay based on protein loading.

Finally, to test the sensitivity of the assay, different protein loadings of 6xHis-tag, wild-type LeuDH was used in the reductive amination of its natural substrate (4-methyl-2-oxopentanoic acid) (Figure 2.14). This will allow estimation of the specific activity necessary of the novel enzyme for detection. The enzyme stock used had a specific activity of 5.75 U/mg for pure protein and of 1.16 U/mg for cell lysate (high activity is not necessary since we are interested in testing sensitivity). Being conservative and taking 10 ng of pure protein as our detection limit (RFU for 10ng is 922 while for 1ng is 270—this corresponds to ~3.4 times greater than the lowest point recorded) the assay can detect an enzyme than can produce 5.75×10^{-11} moles of NADH/min. In our screening process the average protein loading was ~18 μg of protein per well. Assuming that 35% of the loaded lysate protein is the LeuDH variant of interest the assay should be able to detect an enzyme with a specific activity

of ~ 9.1 mU/mg. A schematic representation of the fluorsecence assay can be seen below (Figure 2.15).

2.3.5. Screening of Met65NNK, Lys68NNK library

Equipped with a high-throughput NAD+ fluorescence assay (see above) (163), 3752 clones were screened (Figure 2.15). Due to some misfire of pins from the
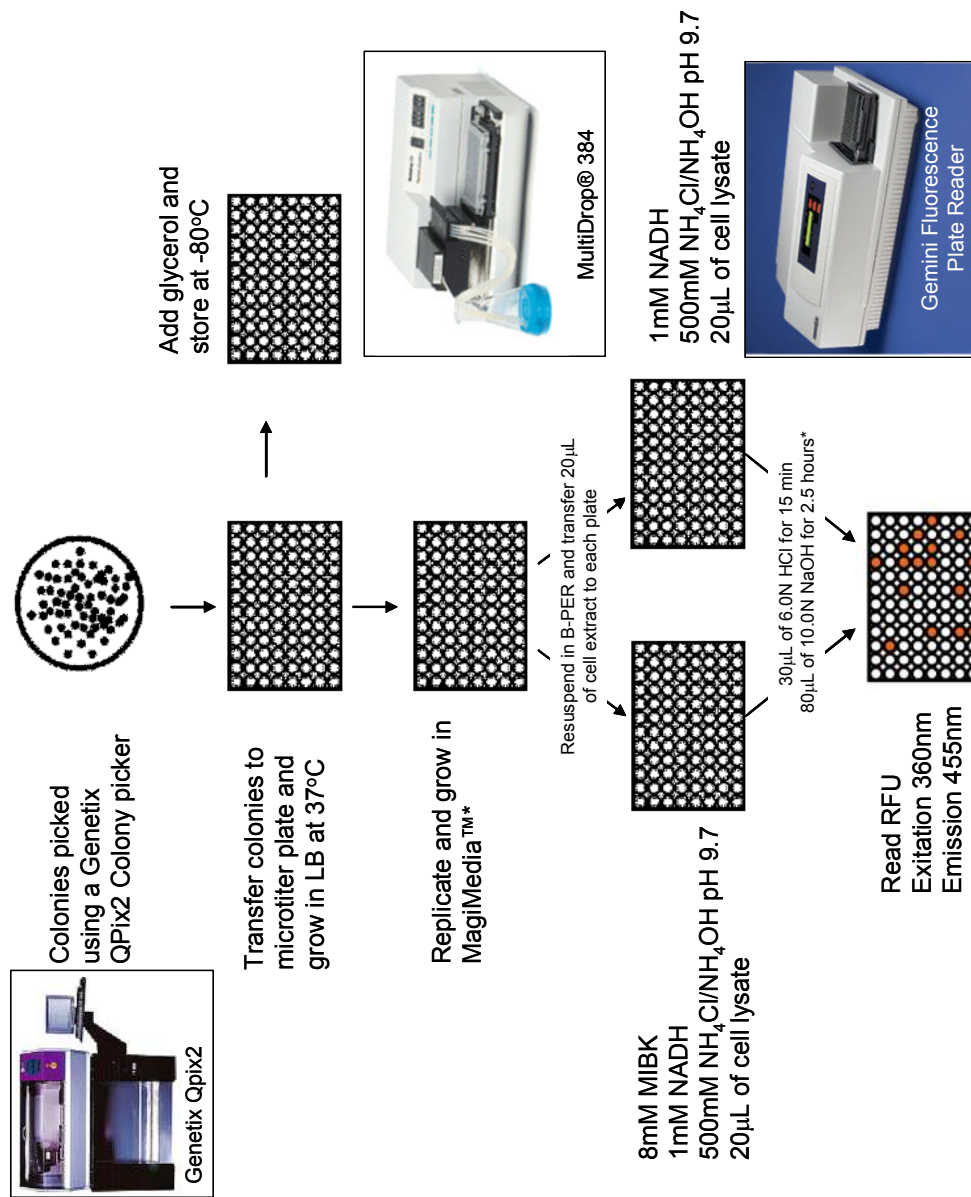
Figure 2.15. Schematic representation of fluorescence NAD+ assay.

LeuDH Library Met65NNK Lys68NNK
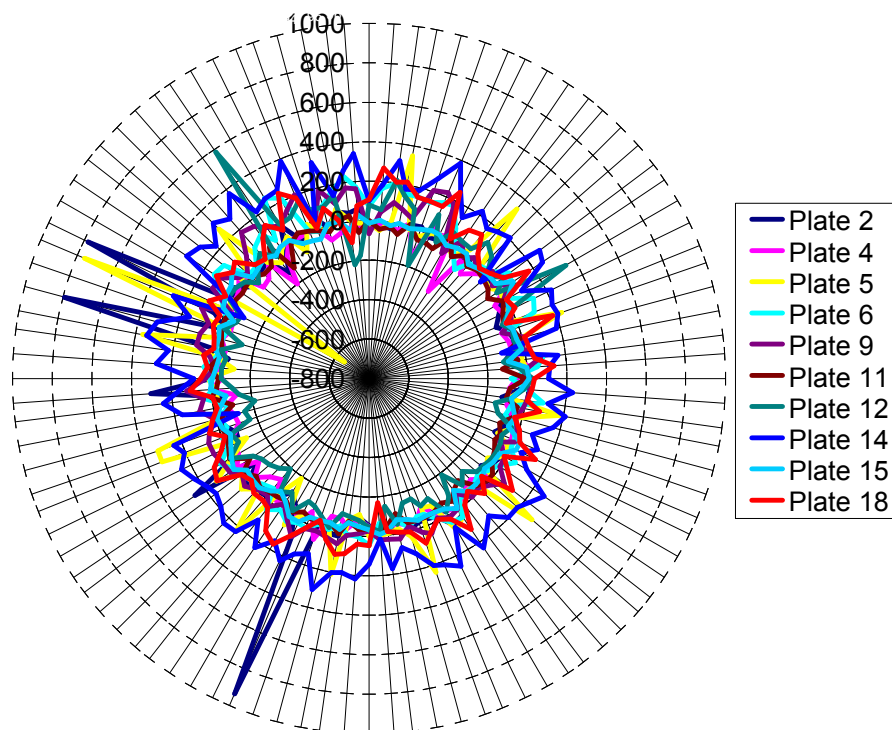
Figure 2.16. Plot 1, screening of LeuDH variants.

colony picker and some partial empty plates, a total of 43 96-well plates were used.

Plates were excited at 360 nm and fluorescence was measured at 455 nm (Figure 2.16,

to see all plots see Appendix, Figure A2.2-A2.5). The radial axis represents RFU

(with background subtracted) while the radial grid represents one well of the 96 in a

plate. For each well a blank was measured (without substrate) and subtracted. The

library mean fluorescence (in RFU) was 858.79 ± 7.7 (± values represent a 95%

confidence interval unless otherwise stated) while for the background the mean

fluorescence was 865.09 ± 7.75 (Figure 2.17). Statistically there is no difference

between the reading with and without MIBK. When each background is subtracted

from its corresponding variant the mean fluorescence is -6.303 ± 3.64. For
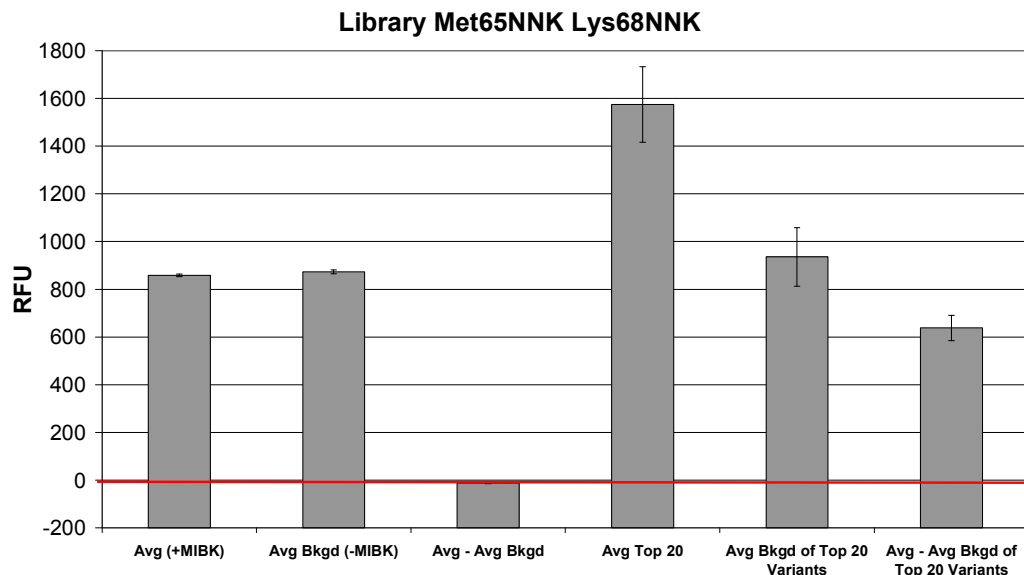
Figure 2.17. Summary of LeuDH variants. Avg = average, Bkgd = background.

thoroughness, the top 20 variants based on fluorescence were sent for sequencing (Table 2.2). These variants had a mean RFU of 1574.5 ± 158.5 with a background of 936.15 ± 123. The mean fluorescence (minus background) was 638.36 ± 53 (Figure 2.17). For complete statistical analysis of library see Appendix (Figure A2.5-A2.10). For the top 6 clones, it is evident that other than Lys68Glu (clone 1) and Met65Asn (clone 3), hydrophobic amino acids are prevalent (Table 2.2).

2.3.6. Screening of promising LeuDH variants via GC

The top 20 variants were then tested via GC to verify product formation. The variants were tested both for amination and deamination of MIBK and 1,3-dimethylbutylamine. For amination, 8 mM MIBK, 1 mM NADH, 500 mM NH₄Cl/NH₄OH pH 9.7, 8.8 mM D-glucose and 0.5 mg/mL of GDH-103 (Biocatalytics now Codexis, Redwood City, CA) was used. The addition of a recycling system (GDH-103) will allow us to overcome equilibrium limitations. For deamination, 8 mM 1,3-dimethylbutylamine, 8 mM NAD+ in 0.1 M glycine pH 10.0

was used. 2 mL reactions were prepared with 100 μL of clarified cell extract.

Reactions were incubated overnight at room temperature. pH was then adjusted to >

11 and extracted with 1 mL of dichloromethane (CH_2Cl_2). Samples were then loaded

to the GC. Unfortunately, no amination or deamination was detected.

Since the K_m of these substrates is unknown, experiments were repeated under

the same conditions but with saturated MIBK (solubility of MIBK in water at 25°C is

170 mM (154)) and saturated 1,3-dimethylbutylamine. Again, no amination or

deamination was detected.

Table 2.2. Sequences and fluorescence (RFU) of the top 20 LeuDH variants.

Clone	Position 65	Position 68	RFU
1	Leu	Glu	935.296
2	Met	Leu	792.708
3	Asn	Phe	778.014
4	Ala	Met	759.867
5	Met	Leu	743.039
6	Phe	Trp	689.327
7	Leu	Cys	674.582
8	Met	Arg	642.346
9	Thr	Glu	629.916
10	Met	Ser	619.663
11	Met	Lys	583.849
12	Bad sequence		578.416
13	Tyr	Cys	574.011
14	Tyr	Glu	555.483
15	Lys	Trp	552.431
16	Leu	Stop	542.891
17	Bad sequence		531.83
18	Phe	Glu	528.529
19	Ile	Lys	528.075
20	Met	Lys	526.963

2.4 Conclusions

In summary, we have carried out site-saturated mutagenesis at position Lys68.

Six variants: 1) Lys68Ala, 2) Lys68Cys, 3) Lys68Gly, 4) Lys68Met, 5) Lys68Ser and

6) Lys68Thr seem to carry out the oxidative deamination of 1,3-dimethylbutylamine.

These six variants were later cloned into a vector with an N-terminal 6xHis-tag for purification. Using UV-absorbance we were able to estimate specific activities and Lys68Met was the best variant at this stage. Lys68Met had a specific activity of 8 mU/mg for the oxidative deamination of 1,3-dimethylbutylamine and 0.2 mU/mg for the reductive amination of MIBK. These results have not conclusively been corroborated via GC. We believe the low activity of the enzymes in addition to potential product loss during extraction leads to no detection of product via chromatography.

In an attempt to further increase the amine dehydrogenase activity of LeuDH, site-saturated mutagenesis was carried out at Met65 and Lys68. 3752 variants were screened via an NAD+ fluorescence assay. The top 6 variants were: 1) Met65Leu Lys68Glu, 2) Met65 Lys68Leu, 3) Met65Asn Lys68Phe, 4) Met65Ala Lys68Met, 5) Met65 Lys68Leu, and 6) Met65Phe Lys68Trp. Again, methionine is prevalent at position 65 and other than glutamate, hydrophobic residues are prevalent at position 68. As before, we believe low activity does not allow for the conclusive corroboration of product via GC or other analytical tools such as NMR. Amine and ketone were not detected via GC above baseline at concentrations < 5 mM and < 0.5 mM, respectively.

The enzymatic reduction of prochiral ketones to chiral amines still remains a desired chemistry in the chemist's tool kit. To demonstrate that amino acid dehydrogenases have the potential to reductively aminate a prochiral ketone would be a significant accomplishment. Amino acid dehydrogenases have been shown to reductively aminate via the same mechanism (9, 20, 140). This would suggest that if an amino acid dehydrogenase is proven to reductively aminate, there is prospect of

converting other enzymes of this family into amine dehydrogenases. This would be a great accomplishment since this family of enzyme accepts a wide range of amino acid side chains and will allow chemists to reductively aminate a wide array of nonproteogenic substrates.

CHAPTER 3

DEVELOPMENT OF A THERMOSTABLE GLUCOSE 1-DEHYDROGENASE

3.1 Introduction

In an effort to produce increasingly complex chemical products, a number of research efforts focus on the development of proteins with novel functionality and improved activity, selectivity, and stability (15, 17, 43, 74, 112, 122, 128). Many of these products are required to be enantiomerically pure for biological activity and often require processing at nonnative conditions of pH, solvent, and temperature. As the proteins have not been optimized for such nonnative conditions during evolution, protein stability often is insufficient. A number of strategies have been employed to apply protein engineering to modify protein structural features to improve their stability. These include improved core packing, (41, 79) turn optimization, (153) increased rigidity by Xaa to Proline or Gly to Xaa substitutions, (44, 100) optimizing surface charges, (157, 179) introduction of disulfide bridges,(96) and improved hydrophobic interactions (82).

Protein engineering methods can generally be classified into one of three categories: rational, (37) combinatorial, (38) and data-driven protein design (14, 26, 119). Combinatorial protein design, often called directed molecular evolution, involves the iteration of diversity generation with subsequent selection or screening to identify improved variants. The advantage of this approach lies in that little knowledge of the protein is required for its application. However, its success relies on the availability of an effective selection or screening tool to sieve through the large

number of variants that are produced. In contrast, rational design requires extensive knowledge of protein structure and function to identify specific mutations to make, and thus generates a manageable number of variants. Lastly, data-driven protein design utilizes available sequences, structures, mechanisms as well as activity, selectivity, and stability data to identify potential amino acids or domains to target for mutation. In data-driven protein design, methods based on sequence alignments are promising; these methods are traditionally applied to homologous protein families to allow a reduced number of candidates.

Steipe et al. (156) were the first to demonstrate the consensus concept to create thermostable immunoglobulin domains. This concept determines the most prevalent amino acid at a given position of an alignment of homologous proteins. The resulting sequence of most prevalent amino acids is termed the *consensus sequence*. Consensus stabilization of proteins is based on the notion that the frequency of sequence elements correlate with stability contribution (156). This fundamental assumption was corroborated by analyzing a library of thermostable cytochrome P450s generated via a consensus approach (88). Lehmann et al. (85-87) and Amin et al. (3) later applied the consensus approach to generate a stable fungal phytase and β-lactamase, respectively. Watanabe et al. (171) utilized a phylogenetic approach to identify ancestral amino acids that ultimately conferred thermal stability onto a 3-isopropylmalate dehydrogenase. Binz et al. (12) used a consensus approach to develop a conserved scaffold (i.e., residues important for maintaining repeat structure) in which potential interacting amino acids where randomized en route to improving interactions between repeating units of the ankyrin protein. This resulted in increase expression, solubility and stability of the ankyrin repeat protein. DiTursi et al. (35) used Bayesian sequence-based algorithms on serine protease sequences to identify a

31

stabilizing motif in subtilisin E, improving its melting temperature by 13°C. In cases of low levels of sequence identity and/or few available amino acid sequences, however, the consensus residue of a specific position cannot be found easily or not at all. We therefore developed the structure-guided consensus approach—which utilizes structural information (in addition to other criteria; see below) to reduce the number of residues in a given protein to be mutated and checked for stabilization—and demonstrated thermostabilization on penicillin G acylase (PGA) in 47% (10 of 21) of investigated single variants (120). To this end we had only 8 sequences, ranging from 32.5 to 87.7% and 28.6-87.6% amino acid identity for the α- and β- chains, respectively. Moreover, PGA is a large, heterodimeric entity with a complicated maturation process to the active protein (23 and 63 kDa in α- and β-subunit, respectively) and so is difficult to access through rational or combinatorial design. Here we extend the structure-guided consensus approach to an important, homotetrameric biocatalyst with cofactor-dependent catalysis: glucose dehydrogenase.

Glucose dehydrogenase (GDH; E.C. 1.1.1.47) catalyzes the regeneration of either NADH or NADPH from the oxidized precursors NAD^+ or $NADP^+$ with concomitant oxidation of inexpensive and readily available glucose (61) (Figure 3.1); its high specific activity, up to 550 U/mg, (116) towards both NAD^+ or $NADP^+$ by far surpasses other approaches (174, 183). Because of the prohibitively high prices of nicotinamide cofactors, in combination with the requirement of cost-effective industrial processes, a robust cofactor regeneration system is a necessity for the production of enantiomerically pure compounds from prochiral precursors. Prior attempts to develop thermostable cofactor-recycling enzymes (6, 68, 93, 107) have been conducted mostly by employing directed molecular evolution.
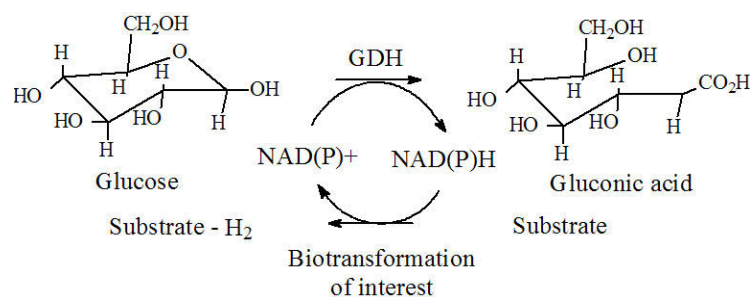
Figure 3.1. Cofactor reduction by glucose 1-dehydrogenase.

3.2. Materials and Methods

3.2.1 gDNA preparation and gene isolation

Bacillus subtilis strain 167 was obtained from the Bacillus Genetic Stock Center (BGSC No. 1A1) as spore dots. Spores were plated and then grown in LB at $37^{\circ}C$. Genomic DNA (gDNA) was obtained using method B described by Mehling et al. (101). GDH was isolated from gDNA using the primers, 5'-ATGTATCCGGATTTAAAAGGAAAAGTCGTCGC-3' and 5'-TTAACCGCGGCCTGCCTGGAATGAAGGATATTG-3'.

3.2.2. Site-directed mutagenesis, cloning and overexpression

Mutations were introduced using primers via overlap extension (60, 63) and then sequenced for confirmation. Overlap extension was favored due to the ability to successively follow progress and its robustness. Primers were designed using the guidelines provided by Stratagene (Cedar Creek, Texas) for its QuikChange® Site-Directed Mutagenesis Kit. Variants were cloned into pET28a (Novagen, Madison, WI) vector carrying an N-terminal His-tag via the *Nde*I and *Hind*III sites and the assembled vector was transformed into *E. coli* BL21(DE3) (Invitrogen, Carlsbad, California). Overnight cultures from a single colony containing the plasmid were used to inoculate 300 mL cultures. Fermentations were grown to an OD_{600} of 0.3-0.5

at 37°C, induced with 0.1 mM IPTG, and incubated for 16 h. Cultures were then pelleted by centrifugation.

3.2.3. Protein purification

Pelleted cells were resuspended in 50mM phosphate pH 8.0 with 300 mM NaCl and 20 mM imidazole. Cells were lyzed via sonication. Clarified cell extracts were purified by immobilized-metal affinity chromatography (IMAC) using a Ni-NTA resin (Qiagen). Clarified cell extracts were loaded onto the resin and washed with 50 mM phosphate pH 8.0 with 300 mM NaCl and 50 mM imidazole. Pure protein was eluted with 50 mM phosphate pH 8.0 with 300 mM NaCl and 250 mM imidazole. A representative SDS-PAGE image can be seen below (Figure 3.2).
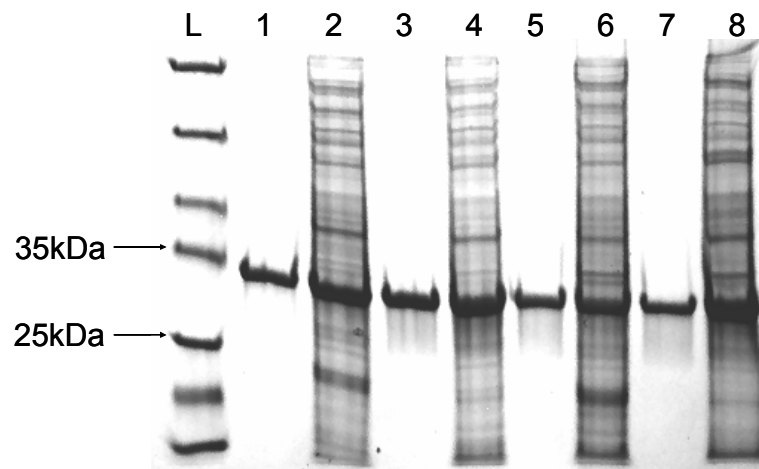
Figure 3.2. SDS-PAGE-Purified Samples. Lane L: Protein markers (i.e., 116, 66.2, 45.0, 35.0, 25, 18.4, and 14.4 kDa), Lane 1: Eluted HisGDH Q252L, Lane 2: Clarified cell extract Q252L, Lane 3: Eluted HisGDH K107E, Lane 4: Clarified cell extract K107E, Lane 5: Eluted HisGDH P45A, Lane 6: Clarified cell extract P45A, Lane 7: Eluted HisGDH wild-type, Lane 8: Clarified cell extract.

3.2.4. Spectrophotometric assay and circular dichroism

Purified, single variants were incubated in elution buffer (50 mM phosphate pH 8.0, 300 mM NaCl and 250 mM imidazole) at 25°C or 65°C and residual activity

34

monitored spectrophotometrically at 25°C by NADH absorbance at 340 nm. Measurements were made in a Beckmann-Coulter DU®800 (Fullerton, CA) spectrophotometer. Triplicate assays were carried out in 50 mM phosphate (NaH$_2$PO$_4$ pH 8.0) with 1 mM NAD+ and 100 mM D-glucose. CD ellipticity at 222 nm was measured using a Jasco J-810 (Easton, MD) with a Peltier-controlled cell holder. Purified samples were dialyzed against 50mM phosphate pH 8.0 with 300 mM NaCl to remove imidazole prior to measuring apparent melting temperature (T$_m$). CD was carried out at a protein concentration between 100-200 µg/mL in a 1 mm path length quartz cuvette.

3.2.5. Calculating Apparent Melting Temperature

Using CD it is possible to monitor effects on secondary structure (e.g., at 222 nm we observe changes in α-helix). A representative protein melting plot represented in ellipticity (in mdeg) at 222 nm against temperature can be seen below (Figure 3.3). Data was nonlinearly fit to a simple sigmoidal model (equation 3.1) using SigmaPlot 8.0. From this data the T$_m$ was calculated as a fitted parameter to equation 3.1.
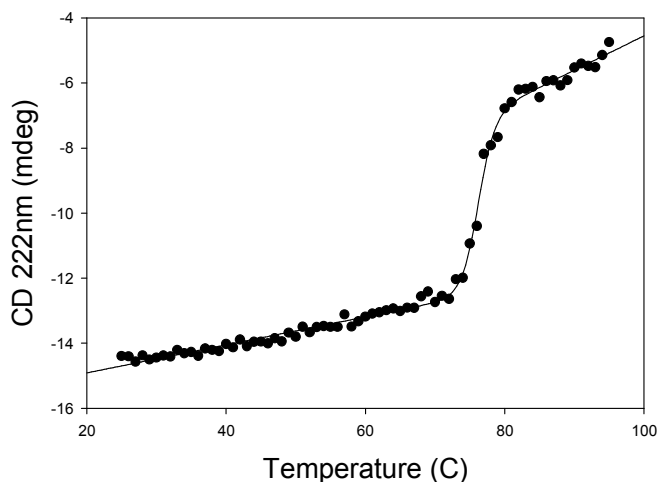
Figure 3.3. Representative CD plot for GDH at varying temperatures. HisGDH E170K.

$$y = A_0 + \frac{A_1 - A_0}{1 + e^{m(T_m - x)}}$$
$$A_0 = B_n + m_n x \qquad (3.1)$$
$$A_1 = B_d + m_d x$$

3.3. Results and Discussion

3.3.1 Identification of substitutions in GDH from *Bacillus subtillis*

In generating a temperature-stable cofactor regenerating system, we decided to use the highly active but thermally labile glucose dehydrogenase (GDH) from *Bacillus subtilis* strain 168 (*B. subtilis*) as a starting template due to its high activity towards NAD^+ and $NADP^+$ (127). In addition, a homolog was previously stabilized by directed molecular evolution by screening 40,000 colonies, (6, 93, 107) allowing us to compare our results directly. There are over 200 available homologous sequences for GDH (we counted 244). Sequences were narrowed down by considering only active glucose-1-dehydrogenases and putative or probable genes, also sequences were selected as to represent the main branches of the phylogenetic tree to reduce bias toward highly similar sequences (see Appendix, Figure A3.1). This resulted in 12 additional GDHs.

The sequences we picked ranged in amino acid identity from 24.8 to 61.1% with respect to the wild-type—a significantly lower and more expansive range than previous applications of the consensus concept by other groups, but still assuring essentially the same three-dimensional structure for all chosen sequences, as GDH belongs to the structurally well-conserved family of short-chain dehydrogenases (70). To pick substitutions, the alignment was carried out using PRALINE, (59, 151, 152) a multiple sequence alignment tool that generates pre-alignment profiles for each sequence in a given set using database searching. The alignment was carried out with

use of the default settings, in which neither the structural-expanded search nor pdb files were utilized for the alignment. These profiles replace the single-sequence input and are progressively aligned, allowing the incorporation of extensive position-specific information from homology-profiles (for output, see Appendix, PRALINE alignment and consensus sequence with corresponding Table A3.1). From this alignment, 31 of the 261 positions were distinct and at least 50% frequent in the consensus sequence. The 50% consensus cut-off can be varied depending on the screening capacity and the stringency of other criteria. The candidates were further reduced by applying the following criteria: 1) the substitution had to be more than 6 Å from the cofactor binding site, to reduce effect of mutagenesis on activity, 2) amino acid secondary structure propensity was considered (for example, if the mutation was found in a helix, a helix-stabilizer was not changed to a helix-destabilizer), and 3) amino acids involved in existing direct hydrogen-bonds or salt-bridges were left unchanged. For this, a homology model based on the crystal structure from *B. megaterium* IWG3 (180) (84% identical to GDH from *B. subtilis*) was used. These criteria should be modified based on available knowledge and data on a specific protein. For example, quaternary structure integrity is critical for activity and stability of GDH (7, 61) as well as many alcohol dehydrogenases (80). Thus, strengthening subunit-subunit interactions is critical for stabilization. As we wanted to evaluate the importance of the interface in stabilization of GDH, we did not include location at the subunit interface as a criterion for inclusion in or exclusion from the list of targeted amino acid substitutions (see Appendix Figure S3.2). Application of criteria 1) to 3) reduced the number of targeted amino acid substitutions from 31 to 24 (9% of the total protein), of which seven are located at the subunit interface. Glutamate at position 170 was mutated into both arginine and lysine because no residue occurred

with more than 50% frequency but more than 60% of the residues were R or K with more than 90% being charged (i.e., five arginines, four lysines, three glutamic acids, and one valine).

In this approach, the most important inputs are the amino acid sequences selected and the consensus cut-off percentage, which, when tuned by the experimentalist, significantly impacts the number of variants to be studied. We decided not to disrupt any preexisting stabilizing interactions based on the homology model and an available crystal structure (180). In addition, amino acid propensities toward different secondary structures have been extensively studied and thus are considered as a reasonable criterion for reducing potential candidates (94, 145). Finally, Morley and Kazlauskas, (105) scrutinized the literature to study the relation between mutation distance from the active site and its effect on enantioselectivity, selectivity, activity, and thermal stability. For the last of these, the authors do not observe beneficial mutations closer than 8 Å from the active site, though in all cases mutations < 5 Å were rarely observed. Thus, our criterion for substitutions to be > 6 Å from the active site is reasonable, if rather conservative.

3.3.2 Single variants

GDH variants were generated by overlap extension and cloned into pET28a, in frame with the N-terminal His-tag. Variants were expressed in *Escherichia coli* (BL21DE3) and purified using immobilized metal affinity chromatography (IMAC). Purified, single variants were incubated in elution buffer at 25°C or 65°C and residual activity monitored spectrophotometrically at 25°C by the decrease in NADH absorbance at 340 nm. Specific activities were measured immediately after purification to minimize deactivation. Half-lives were determined from residual activity with first-order deactivation kinetics (Figure 3.4).
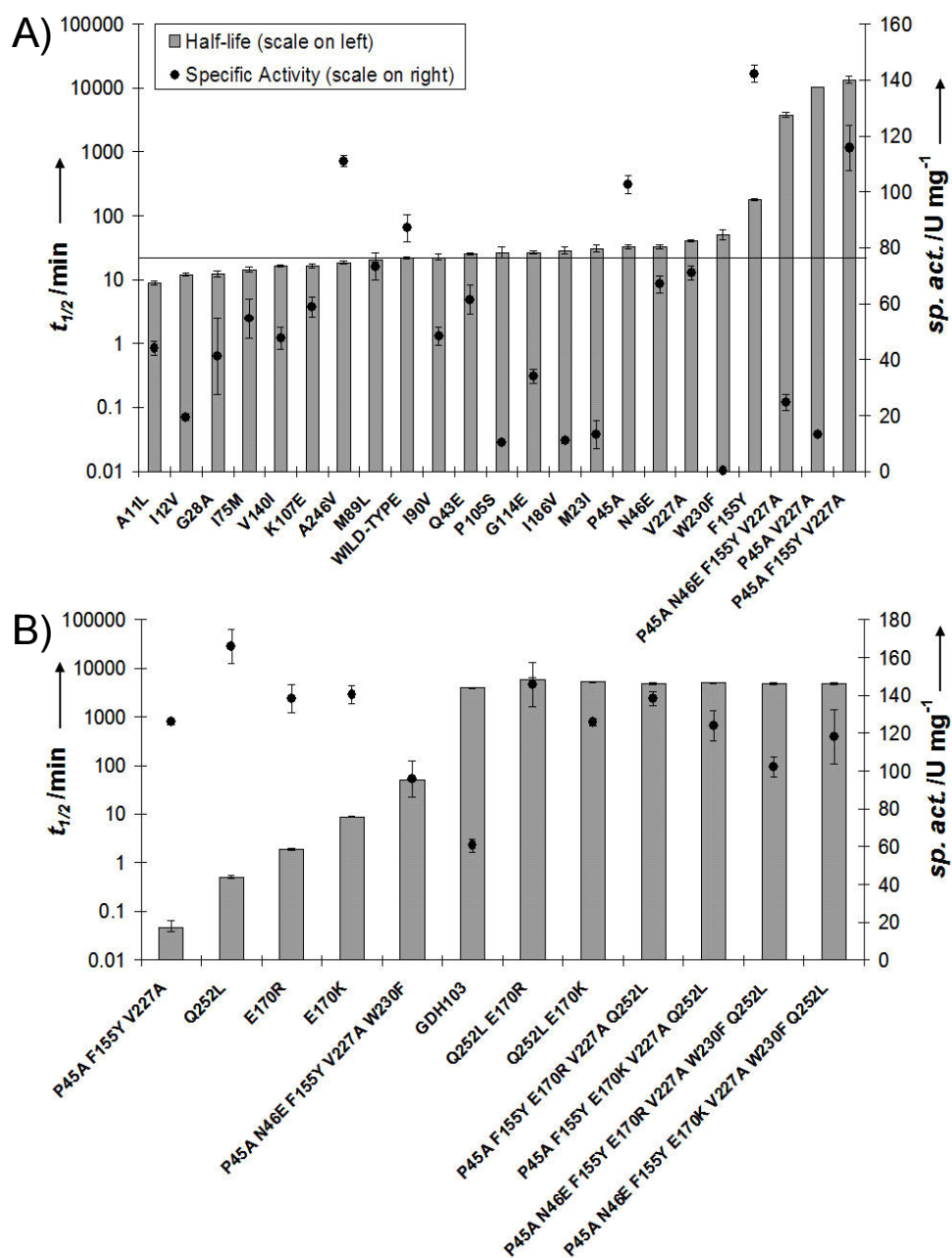
Figure 3.4. Half-lives and specific activity for HisGDHs. A)25°C, B)65°C. Error bars for half-lives represent 95% confidence intervals. Error bars for specific activity represent standard deviations. The most stable variant in the top is the least stable in the bottom one.

When we compare the results for single mutants of all 24 of the selected residues, 11 had higher thermal stability compared to the wildtype, seven had a lower thermal stability, three were unchanged, and three were not active. This is equivalent

to a 46% success rate, comparable with our previous application of this technique in penicillin G acylase (120). In contrast, a combined 40,000 variants generated via chemical mutagenesis (93) and family shuffling (6) were previously screened to identify E170K and Q252L. This is typical of directed molecular evolution; Johannes et al. screened 19,000 variants created by error-prone PCR and identified 12 amino acids that increased the thermal stability of phosphite dehydrogenase (PTDH) (68).

If we had limited our search to amino acids at the subunit-subunit interface, we would have had to characterize seven variants and would have found six with higher thermostability (G114E, F155Y, E170(K or R), V227A, W230F, and Q252L) and one less stable one (A246V). This finding is equivalent to a 86% success rate, and the successful set included the best variants: F155Y, E170(K orR) and Q252L (see Appendix, Figure A3.2).

All active variants had essentially unchanged levels of expression. The three inactive variants—Q31G, K204E and K234D—came as a consequence of aggregation as verified by SDS-page (Figure 3.5). All three substitutions were present at the protein surface, and the instability might have been an effect of decreased rigidity for Q31G and unfavorable charge distributions for K204E and K234D. K204E is in close range of D202 and D208, while K234D is flanked by D4, E223, and E235. Although interaction of charged residues is difficult to predict, it has been shown to be important for protein stability (157, 179) and could thus be considered in the future as a criterion to minimize inactive variants.

On reviewing the selection criteria it is evident that, although the consensus sequence is effective in predicting stabilizing residues, the stability of the variant or its success are not directly related to the frequency of a particular amino acid. The 24 amino acids had a frequency ranging from ~50 to ~ 80%, but the more frequent ones did not seem to confer more stability. For example, the three inactive variants Q31G, K204E, and K234D had a consensuses of 80, 70 and 60% respectively. In addition, the stabilizing mutations that were ultimately chosen for subsequent combination and characterization (see below)—P45A, N46E, F155Y, E170(K or R), V227A, W230F, and Q252L—had a consensus ranging between 50-70% (i.e., 60, 50, 50, 60, 70, 50, and 60%, respectively). From this it is evident that we would not have benefited by being more stringent with the selected 50% cut-off.
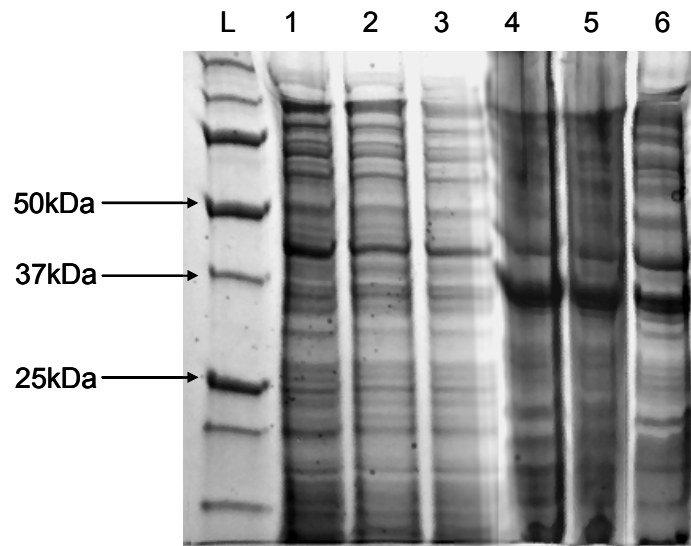
Figure 3.5. SDS-Page 1-Aggregated Samples. Lane L: Protein markers (i.e., 150, 100, 75, 50, 37, 25, 20, 15, and 10 kDa), Lane 1: Clarified cell extract HisGDH Q31G, Lane 2: Clarified cell extract HisGDH K204E, Lane 3: Clarified cell extract HisGDH K234D, Lane 4: Solubilized cell debris HisGDH Q31G, Lane 5: Solubilized cell debris HisGDH K204E, Lane 6: Solubilized cell debris HisGDH K234D.

3.3.3. Combining single mutations

Only single variants that resulted in at least a 50% increase in thermal stability relative to the *B. subtilis* wild-type (i.e., P45A, N46E, F155Y, E170(K or R), V227A, W230F, and Q252L) were combined and further analyzed. However, the effect of two mutations—Q252L and E170(K or R)—dominate the effect of the other individual substitutions. When substitutions are combined, we can gather the following order in increasing stability: P45A/V227A < P45A/F155Y/V227A < P45A/N46E/F155Y/V227A/W230F < Q252L < E170(K or R) < E170(K or R)/Q252L. As can be observed, some mutations, when combined, increase the half-life (i.e., P45A, F155Y, V227A and W230F) while others do not (i.e., N46E). As in the case of single variants, the amino acid changes Q252L and E170(K or R) in combined mutations of *B. subtilis* GDH contribute significantly to stabilization while the others (i.e., P45A, N46E, F155Y, V227A, W230F) make minor contributions, resulting in a plateau of stabilization for Q252L/E170(K or R) variants of GDH that does not seem to be overcome with additional mutations. The role of Q252L and E170(K or R) for GDH stabilization is in agreement with the findings of Baik et al.(6, 7) who discovered the importance of residues Q252L and E170K in the stability of GDH. Assisted by crystal structures from the wild-type and mutant GDHs from *Bacillus megaterium* IWG3, Baik et al. determined that a cooperative effect between Q252L and E170K stabilizes the tetramer structure by strengthening the hydrophobic interactions within the interface(7)— significant, since GDH has been demonstrated to be an obligate tetramer. We surmise E170R has a similar structural effect on GDH. The half-life of the best variant was ~ 3.5 days (inactivation constant ~ 2.2×10^{-6} s^{-1}) at 5°C with specific activities ranging from 100-145 U/mg at 25°C. The variants had a higher thermal stability and specific activity when compared to GDH-103, a

commercially available (Biocatalytics now Codexis, Redwood City, CA) thermostable

GDH. The superior specific activity is a result of better purity achieved through

IMAC chromatography—corroborated with SDS-page (Figure 3.6)—and possibly of

the added His-tag of our enzyme. Specific activity is the clear advantage of GDH

when compared to commercially available NAD^+ and $NADP^+$ formate

dehydrogenases (FDH; 5-20 U/mg at 25^oC-30^oC) and phosphite dehydrogenase (5

U/mg at 25^oC) since inactivation constants as low as 9.7×10^{-8} s^{-1} at 55^oC and 1.37×10^{-6} s^{-1} at 45^oC, respectively, have been reported for FDH and PTDH (68, 162).+
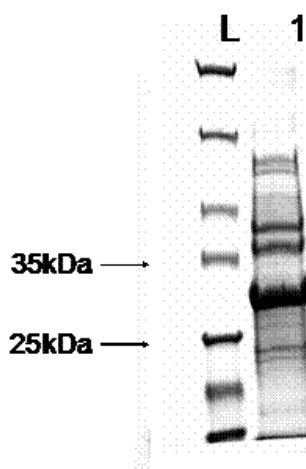
Figure 3.6. SDS-Page GDH 103 (Biocatalytics now Codexis, Redwood City, CA).

3.3.4. Incorporation of F155Y, E170K, and Q252L into homologues

To test the broad applicability of the structure-guided consensus concept

across homologs, the three most significant mutations (F155Y, E170K, and Q252L)

were chosen and appropriate amino acid changes were introduced into two

homologous GDH scaffolds. The two selected candidates (*Bacillus thuringiensis* and

Bacillus licheniformis) were in the same branch as *B. subtilis* and share 81.7% and

67.2% amino acid identity, respectively.

HisGDH from *B. thuringiensis* was as active (specific activity: 93 U/mg vs. 87

U/mg at 25^oC and pH 8.0) and deactivated much more slowly than the *B. subtilis*

wildtype but slightly more rapidly than its improved variants (Table 3.1). Interestingly, the kinetic stability (that is the half-life) of the most stable *B. thuringiensis* variant (E170K/L252) is only slightly less than the equivalent *B. subtilis* variant (E170K/Q252L) at the same temperature and pH value (65°C, pH 8.0). In contrast, HisGDH from *B. licheniformis* was hardly active (50mU/mg at 25°C, pH 8.0) and so we additionally measured the thermodynamic stabilities of all variants through their apparent melting temperatures (T_m) by circular dichroism (CD) at 222 nm (Table 3.1). For comparison, half-lives of promising candidates were determined.

Table 3.1. Half-lives and apparent melting temperatures (T_m) as measured by CD.

	Position			ΔT$_m$(K)[a]	Half-life at 65°C(min)
	155	170	252		
B. subtilis wt	F	E	Q	0	<0.05[b]
	Y	E	Q	1.2	<0.05[c]
	F	**K**	Q	25.9	8.7
	F	E	**L**	17.3	0.5
	F	**K**	**L**	34.7	5323
B. thuringiensis wt	F	E	L	0	8.0
	Y	E	L	4.5	
	F	**K**	L	18.5	3950
	F	E	**Q**	-19.3	
B. licheniformis wt	F	E	K	0	8.7
	Y	E	K	1.9	
	F	**K**	K	-7.1	
	F	E	**L**	8.9	7.1
	F	E	**Q**	-2.7	
	F	**K**	**L**	5.6	4.2
[a] CD: ellipticity at 222 nm (α-helix). The T$_m$ of the wild-type GDH are 50.2^{o}C for *B. subtilis* wt, 72.2^{o}C for *B. thuringiensis* wt and 70.3^{o}C for *B. licheniformis* wt. [b] Half-life at 25°C ~22 min. [c] Half-life at 25°C ~176 min .					

We have demonstrated that in fact the mutations F155Y, E170K, and Q252L confer increased stability on homologs (except in one situation). F155Y stabilizes all three homologs, albeit slightly (1.2 - 4.5 K). The mutation E170K markedly stabilizes *B. subtilis* and *B. thuringiensis* GDH (25.9 and 18.5 K, respectively) but destabilizes *B. licheniformis* GDH by 7.1 K. The cause is verly likely electrostatic repulsion between E170K and wildtype K252, since it is known from the crystal structure that both residues interact with each other. The mutation Q252L stabilized *B. subtilis* GDH by 17.3 K, while conversely, the mutation L252Q from wildtype *B.*

thuringiensis GDH destabilized by 19.3 K. In *B. licheniformis* GDH, the difference in T_m between Q252 and L252 was 11.6 K. Furthermore, Baik et al. (6, 7) has found stabilization of the E170K and Q252L variants in *B. megaterium* GDH—83.5% identical to *B. subtilis* GDH—so the stabilizing effects of Y155 has been shown in three cases, and of K170 and L252 in four cases each. Clearly, the positions 155, 170, and 252 are all important for stability, and residues in these positions apparently can be ordered with respect to stability as: 155: F < Y, 170: E < (K or R) (except in *B. licheniformis*, where K252 needs to be mutated to L to see a net improvement in stability, see above), 252: Q,K < L.

3.4. Conclusions

In summary, we have demonstrated the stabilizing effect of consensus mutations in multiple natural and diverse variants of the same functionality. Additionally, we demonstrated that the *structure-guided consensus concept* is a very successful and widely applicable technique for generating thermostable proteins (86, 120, 156)—further demonstrated in the improvement in the stability of GDH. The stability of the His-tagged *B. subtilis* GDH was improved from ~20 minutes at 25°C to ~3.5 days at 65°C, a 10^6-fold improvement. Using this approach we were able to identify residues critical to stability with considerably less effort than is involved in traditional directed molecular evolution ($O(10)$ vs. $O(10^4)$—screening in the order of 10 vs. 10,000). The number of investigated variants can be tuned by the experimentalist; the two most influential factors are the amino acid identity range of the selected sequences and the chosen consensus level (here 50%). In contrast to directed evolution, where key amino acids are first identified via high-throughput screening (HTS), each variation from the parent sequence has to be incorporated into the parent gene, and the variant purified and characterized. In the case of difficult

and/or expensive substrates, not amenable to HTS, a small number of variants can be characterized directly on the substrate of interest, a strong advantage ("you get what you screen for") (4). The success of the structure-guided consensus concept depends on the diversity of available sequences. The number of consensus candidates can then be reduced with available information, such as 1) stability data from thermostable or labile homologs, (47, 118) 2) crystallographic structures, including B-factor analysis (132) and Ramachandran plots, (50, 144) and 3) information about the deactivation mechanism, such as crucial subunit interactions, which might prefer surface residues. Often, use of all available data and tools optimizes output while minimizing effort. As discussed, if we had limited our search to the subunit-subunit interface, we would have increased our success rate from 46% to 86%. The development of thermostable enzymes is of interest not only to the specialty chemical and pharma industries, but also to the protein engineer, since recent work has shown that evolvability is linked to stability (11, 13). This finding supports the idea that protein engineering projects will benefit from starting with a stable protein scaffold. Thus, straightforward, time-saving techniques—such as the structure-guided consensus concept—are projected to prove invaluable to protein engineering efforts.

CHAPTER 4

CHARACTERIZATION OF GDH VARIANTS

IN SALT SOLUTIONS AND

HOMOGENEOUS AQUEOUS-ORGANIC MEDIA

4.1. Introduction

Enzymes are extraordinary catalysts that provide unsurpassed fidelity and selectivity under ambient and near ambient conditions of pH, temperature and solvent composition. As target compounds in pharma become more complex and processes more environmentally demanding, biocatalysts are increasingly favored over non-biological catalysts (29, 34, 104, 115, 122). However, despite their indisputable favorable qualities, insufficient protein stability in reaction media has hampered biocatalyst implementation in organic synthesis. Recognizing this shortcoming, a number of approaches have been developed to highly stable proteins for biocatalysis applications, either by isolating and expressing novel proteins with exceptional stability (52) or engineering the protein environment through the addition of co-solutes (19, 33, 119), chemical modifications (119), and immobilization of the protein in solid or gel-like matrices (72, 119). Alternatively, protein engineering methods have been used to modify the protein scaffold to improve hydrophobic interactions (82), electrostatic-surface interactions (157, 179), internal packing (41, 79), and increased rigidity (44, 96, 100) to withstand denaturing conditions. Although comprehensive, generalizable guidelines for engineering stable proteins have not been

achieved, these innovative approaches have produced stable proteins with significant success (37, 38, 87) while also making significant contributions to our understanding of protein stability.

The stability of proteins in organic solvents is one particular problem that continuously receives attention due to increasing interest for industrial applications, in particular fine chemicals and pharmaceuticals. Increased biocatalyst tolerance to organic solvents can allow for an increased concentration of denaturing substrates and products (e.g., alcohols and ketones for alcohol dehydrogenases and ketoreductases) to be used. This in turn allows a wider range of operating conditions (i.e., reactant concentrations) to be utilized to optimize reactions by shifting reaction equilibrium, reduce/eliminate hydrolysis or polymerization side reactions, and avoid microbial contamination (36). This results in increased productivity and thus more economically attractive processes. Though much work has gone into understanding and predicting organic solvent effects on enzymatic activity based on water activity (165, 166), log P (83, 84), and solvent polarity (51, 77), most medium optimization for biocatalysis is done on a case by case basis. Khemelnitsky et al. (76) have studied enzyme behavior in the combined presence of salts in organic solvents and progress has been summarized elsewhere (143). In this work, salt and organic solvents effects on enzyme behavior will be studied independently.

Most interesting, it has been noted that an improvement in temperature stability results in a simultaneous increase in stability in the presence of organic solvents (and other denaturants) (31, 113). Cowan et al. (31, 113) showed a positive correlation between thermal stability and tolerance to organic solvents and Hao et al., showed how a thermostable fructose biphosphate aldolase found via directed evolution also showed increased stability in organic solvents (54).

In our previous work (see chapter 3) we improved the thermostability of a glucose 1-dehydrogenase (GDH) from *Bacillus subtilis* (*B. subtilis*) strain 168 via a structure-guided consensus concept (168). This approach makes use of the consensus sequence (156) in addition to protein structure (or homology) to select and sieve through potential amino acid mutations. GDH catalyzes the oxidation of glucose to gluconolactone with the concomitant reduction of NAD(P)+ to NAD(P)H. GDH is an enzyme of great interest since it utilizes an abundant and inexpensive substrate, it accepts both NAD+ and NADP+, has very high activity (>100 U/mg), and is not thermodynamically limited due to the sponataneous hydrolysis of the gluconolactone. Currently, this enzyme is widely used (104, 121) and has great potential in the regeneration of nicotinamide cofactors for ketoreductions (34, 104), Baeyer-Villiger oxidations (102), alkane/alkene hydroxylation (178), carbon-carbon double bond reductions (27, 34, 53, 122) and amino acid dehydrogenases (16, 34). In this work we characterize four variants of GDH from *B. subtilis* with different kinetic and thermodynamic stabilities in different stabilizing and destabilizing salts and a variety of organic cosolvents. We intend to test the potential of the structure-guided consensus method toward developing enzymes tolerant to denaturing salts and organic cosolvents.

4.2. Materials and Methods

4.2.1. Cloning, site-directed mutagenesis via overlap extension and purification

GDH 6xHis-tagged variants were cloned, mutated and purified as described in section 3.2 (168). Purified variants were then dialyzed in 50 mM phosphate pH 8.0 at 4°C to remove imidazole and sodium chloride. The concentration of the purified His-tag GDH variants was measured via Bradford assay using commasie protein reagent

from Pierce (Rockford, IL). For a representative SDS-PAGE of purified samples see Figure 3.2.

4.2.2. Enzyme activity assay

Enzymes were assayed using a Beckmann-Coulter DU®800 spectrophotometer (Fullerton, CA). Enzyme was added to 1mL of assay solution containing 100 mM D-glucose, 1 mM NAD+ in 50 mM sodium phosphate pH 8.0. Enzyme activity was determined by monitoring NADH formation at 340 nm. $k_{d,obs}$ was calculated from first-order deactivation kinetics.

4.2.3. Hofmeister series, kinetic stability

Concentrated, pure His-tag GDH was diluted to 50 μg/mL with the different salts of interest (i.e., NaI, NaNO₃, NaBr, NaCl, NaHCO₂, NaF, NaSO₄, and NaCH₃CO₂) in 50 mM sodium phosphate pH 8.0 at 0.95 and 0.99 water activity. Enzyme-salt mixtures were deactivated by incubating at 65°C and periodically assaying residual activity.

4.2.4. NaCl, GDH temperature dependence

For varying NaCl concentrations, a concentrated stock solution of each GDH variant and a stock of 4 M NaCl in 50 mM phosphate pH 8.0 was mixed with 50 mM phosphate pH 8.0 to obtained the varying salt concentrations studied.

4.2.5. Circular dichroism (CD)

CD ellipticity was measured with a Jasco J-810 (Easton, MD) with a Peltier-controlled multi-cell holder using a quartz cuvette with a 1 mm path length. Samples

were heated at approximately $3^{\circ}C$/min and change in CD ellipticity at 222 nm was monitored. This data was then used to calculate apparent T_m.

4.2.6. Kinetic stability at different temperatures: determination of T_{50}^{60}

Pure, dialyzed GDHs at 50 μg/mL were incubated at the different temperatures for one hour. After incubation the residual activity was measured as described above (section 4.2.2.).

4.2.7. Half-life in organic solvent

For half-life experiments, samples were prepared in 20% v/v organic solvent and incubated at $25^{\circ}C$. Residual activity was measured until approximately 10% enzymatic activity remained. Half-life was determined by first order deactivation kinetics.

4.2.8. Kinetic stability in organic solvent: determination of C_{50}^{0}

A combination of: 1) concentrated NAD+ and D-glucose in pH 8.0 50 mM phosphate, 2) pH 8.0 50 mM phosphate, 3) organic solvent, and 4) pure, dialyzed enzyme were mixed to obtain different aqueous-organic cosolvent compositions. Mixture was assayed immediately upon enzyme addition.

4.2.9. Kinetic stability in organic solvent: determination of C_{50}^{60}

Different proportions of 5 mM phosphate (to minimize buffer precipitation) and organic solvent by % v/v where made and precooled to $4^{\circ}C$. Pure enzyme was then added to result in 50 μg/mL and the desired % v/v of organic solvent. Samples were incubated for one hour at $4^{\circ}C$ after which the residual activity was measured.

4.3. Results and Discussion

Four previously generated GDH variants (168) were selected to study salt and organic solvent effects on thermostable glucose dehydrogenases: tripe (P45A, F155Y, V227A), single (E170K), double (E170K, Q252L) and septet (P45A, N46E, F155Y, E170K, V227A, W230F, Q252L) variants, termed GDH1, GDH2, GDH3, and GDH4, respectively. All GDHs studied contain an N-terminal His-tag to facilitate purification. These variants were selected since they all had enhanced stability when compared to the wild-type, with half-lives of ~0.05, ~9.0 and ~5000 min at 65°C (300 mM NaCl, 250 mM imidazole, 50 mM phosphate pH 8.0) for GDH1, GDH2, and GDH3/GDH4 (GDH3 and GDH4 have similar kinetic stability under these conditions). The wild-type was not selected due to its low stability and thus difficulty to study at high temperatures in the presence of destabilizing salts or in organic cosolvents. These variants, allowed the examination of a broad range of kinetic stability and to study their behavior in the presence of salts and organic cosolvents.

4.3.1 Effects of salt type on glucose dehydrogenase kinetic stability

For characterization in salts we decided to test GDH in sodium salts along the Hofmeister series: I$^-$, NO$_3^-$, Br$^-$, Cl$^-$, HCO$_2^-$, F-, SO$_4^{2-}$, and CH$_3$CO^{2-} (Figure 4.1). Previous work demonstrated that enzyme deactivation in chaotropic solutions correlated exponentially with anion B-viscosity coefficients, a measure of ion

$$SCN^- > I^- > ClO_4^- > NO_3^- > Cl^- \qquad HCO_2^- < F^- < SO_4^{2-} < CH_3CO_2^- < PO_4^{3-}$$

chaotropes	kosmotropes
B-viscosity coefficient < 0	B-viscosity coefficient > 0
(destabilizing)	(stabilizing)
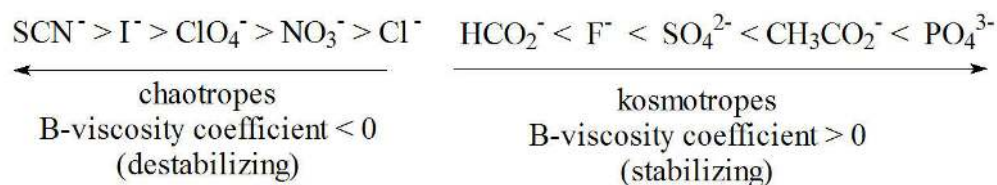
Figure 4.1. Hofmeister series.

hydration (76). Moreover, an asymptotic stability is observed in the kosmotropic regime (19) which suggests a limiting stability obtainable in salt solutions. Residual activities were measured at 65°C and the observable deactivation rate constant ($k_{d,obs}$) calculated based on experimentally verified first-order deactivation kinetics. The logarithm of the deactivation constant, log ($k_{d,obs}$) was plotted against B-viscosity coefficients (Figure 4.2). GDH1 was too unstable at 65°C to be appropriately characterized.

It is evident from the data that the GDH variants do not behave entirely as observed by Broering and Bommarius (19) with the characteristic negative linear correlation in the chaotropic regime with B-viscosity coefficient and constant log ($k_{d,obs}$) in the kosmotropic regime. Though both GDH3 and GDH4 appear to exhibit a lower bound for the value of $k_{d,obs}$, consistent with prior observations(19), only the strongest chaotrope tested (NaI) significantly decreased protein stability. In addition, GDH3 and GDH4 show a slight decrease in stability, up to a twofold increase in $k_{d,obs}$ (or a 10% difference in log ($k_{d,obs}$) with increasing B-viscosity coefficient in the kosmotropic regime; this change might be within experimental error of the results observed previously (19).

The more thermostable variants (GDH3 and GDH4) demonstrate high stability regardless of salt-type at both water activities (a_w) of 0.99 and 0.95. Only NaI solutions demonstrate significant destabilizing effects at 0.95 water activity. A decrease in kinetic stability with increasing chaotropic salt concentration has been previously observed (19).

The kinetic stability of GDH2 depended strongly on salt type. The $k_{d,obs}$ value changed up to 500-fold between salts (I^{-} versus Cl^{-} at 0.95 water activity). The same behavior—though smaller in magnitude is observed for GDH3 and GDH4. From this data we can rank the tolerance of GDHs to deactivation by I^{-} as GDH4 > GDH3 > GDH2.
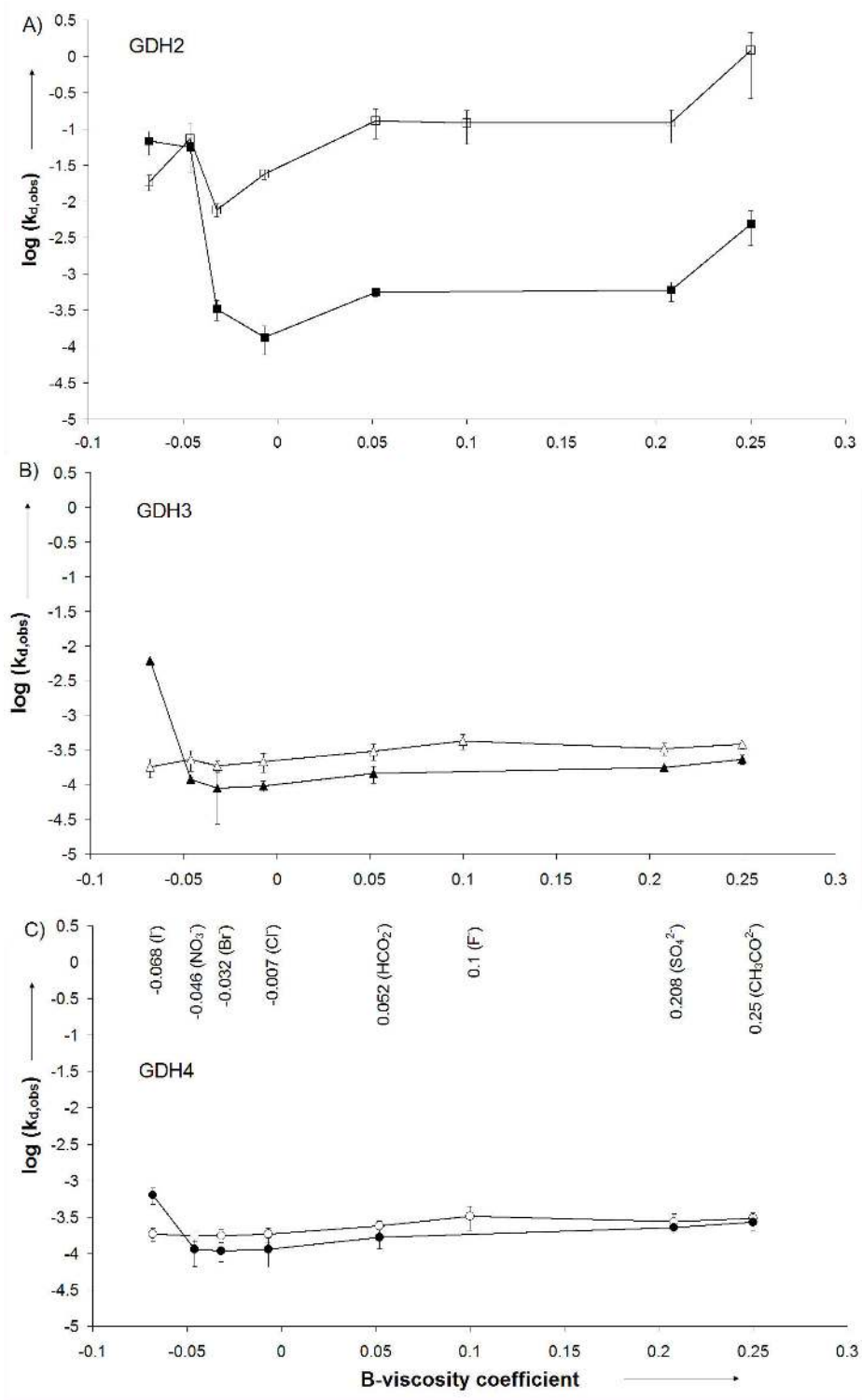
Figure 4.2. Kinetic stability of GDH variants in different salts at a_w=0.95 (solid markers) and a_w=0.99 (open markers) at 65°C and pH 8.0. $k_{d,obs}$ values in 1/min. A) GDH2 (E170K), B) GDH3 (E170K, Q252L), C) GDH4 (P45A, N46E, F155Y, E170K, V227A, W230F, Q252L). Inset values represent the B-viscosity coefficients with the corresponding anion in parenthesis. Fluoride was not tested at a_w=0.95 since it is not soluble at this concentration.

All GDH variants had increased stability at lower a_w—increased salt concentrations—in all salts except in iodide. In addition, this stability dependence on salt concentration is clearly observed for GDH2 and less pronounced for GDH3 and GDH4 with GDH3 showing more dependence in salt concentration than GDH4. All in all, it is evident that those GDH variants with increased kinetic stability (GDH3 and GDH4) demonstrate less susceptibility to variations in stability based on salt types.

4.3.2. Stability of glucose dehydrogenase as a function of salt concentration

For GDH, thermal stability has been previously improved by adding sodium chloride (6, 7, 107, 176). The thermodynamic and kinetic stability of GDH1-GDH4 was tested in solutions with NaCl concentrations ranging from 50 mM to 3 M.

To test thermodynamic stability, apparent melting temperatures (T_m) were recorded at 222 nm via circular dichroism. All variants are stabilized with increasing NaCl concentration (Figure 4.3A). Variants rank in the same order with respect to thermodynamic stability as previously in the case of kinetic stability: GDH4 > GDH3 > GDH2 > GDH1.

If the influence of changing salt concentration on proteins fit a Debye-Hückel model, either log K or T_m, would yield a straight line when plotted against the square root of the ionic strength: $\log K/K_o$ or $T_m/T_{m,o} = -A|z_{protein} \cdot z_{salt}|(I)^{1/2}$ (62). T_m values were normalized with respect to the most labile variant, GDH1 (P45A, F155Y, V227A): $T_{m,o} \sim 46.7^o C$ (319.9K) without salt. The ratios $T_m/T_{m,o}$ were then plotted against the square root of the ionic strength of NaCl (Figure 4.3B). Excellent linear behavior was observed for the three most stable variants (GDH2-4) with $R^2 > 0.99$. The intercepts rank in order of thermal stability (from 1.00 for GDH1 to 1.11 for GDH4) while the slopes of the three most stable variants are the same within
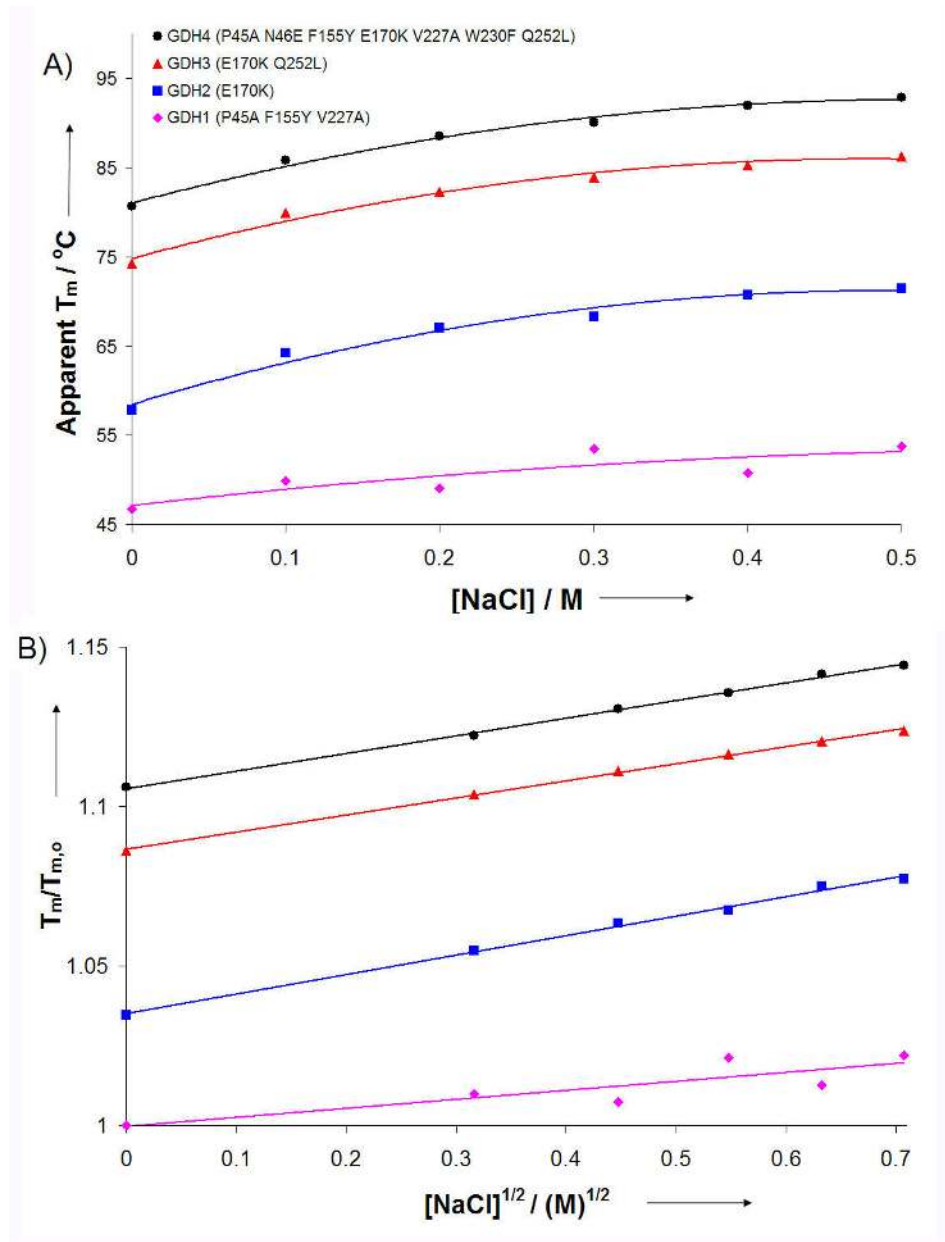
Figure 4.3. Thermodynamic stability of GDH variants in presence of NaCl. A) Apparent melting temperature against NaCl concentration. B) Apparent melting temperature normalized to GDH1 at 0 M NaCl (319.9K) against the square root of the ionic strength. Linear correlations for B: GDH4, $y = 0.055x + 1.1057$; GDH3, $y = 0.0536x + 1.0866$; GDH2, $y = 0.0612x + 1.0351$; GDH1, $y = 0.0282x + 0.9997$.

measurement error. Therefore, thermal stabilization of GDH does not occur via changes of the interaction of the protein surface with the surrounding medium but rather by enhancing cohesive forces in the interior of GDH.
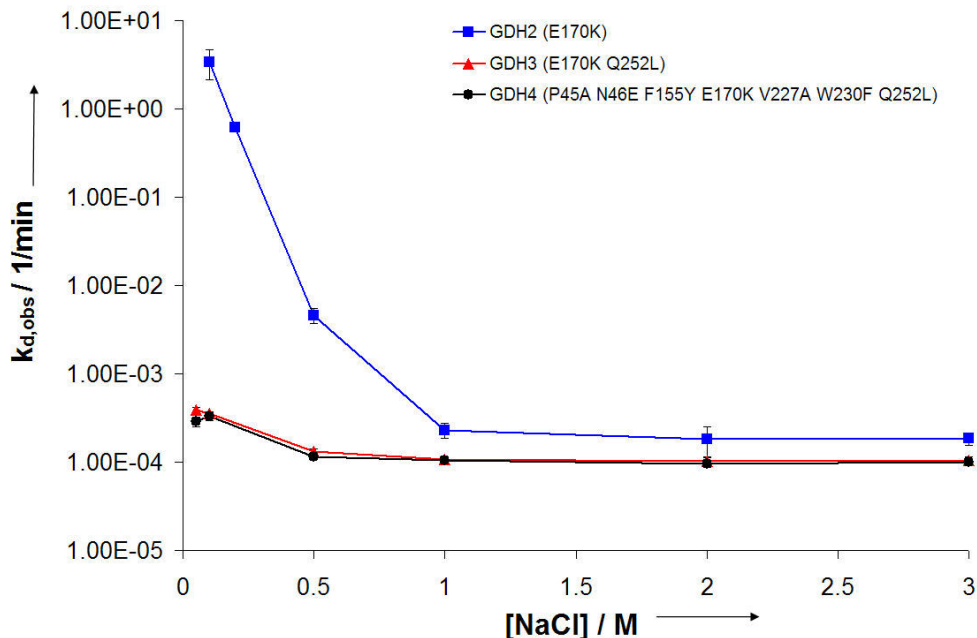
Figure 4.4. Kinetic stability of GDH variants in varying NaCl concentrations.

The test of kinetic stability as a function of salt concentration reveals that GDH2 is strongly stabilized by salt concentrations up to 1 M (Figure 4.4). While this trend has been observed earlier on wiltype GDH from *B. megaterium* (176), a close relative of the protein from *B subtilis* utilized here, although the exact functional relationship is different. The strong salt dependency of GDH2 (E170K) matches finding by Baik et al. (6, 7). The kinetic stability of GDH3 and GDH4 was virtually identical (168) (Figure 4.4), whereas the thermodynamic stability (as expressed by apparent T_m) of GDH4 was considerably higher than that for GDH3 (Figure 4.3). As has been observed for deactivation of GDH upon changes in pH (7), kinetic stability is determined by dissociation of the obligate tetramer into subunits—for pH GDH dissociates into dimers (7). In contrast, melting points as a basis for thermodynamic stability register unfolding events that apparently are more difficult to accomplish for GDH4 than for GDH3.
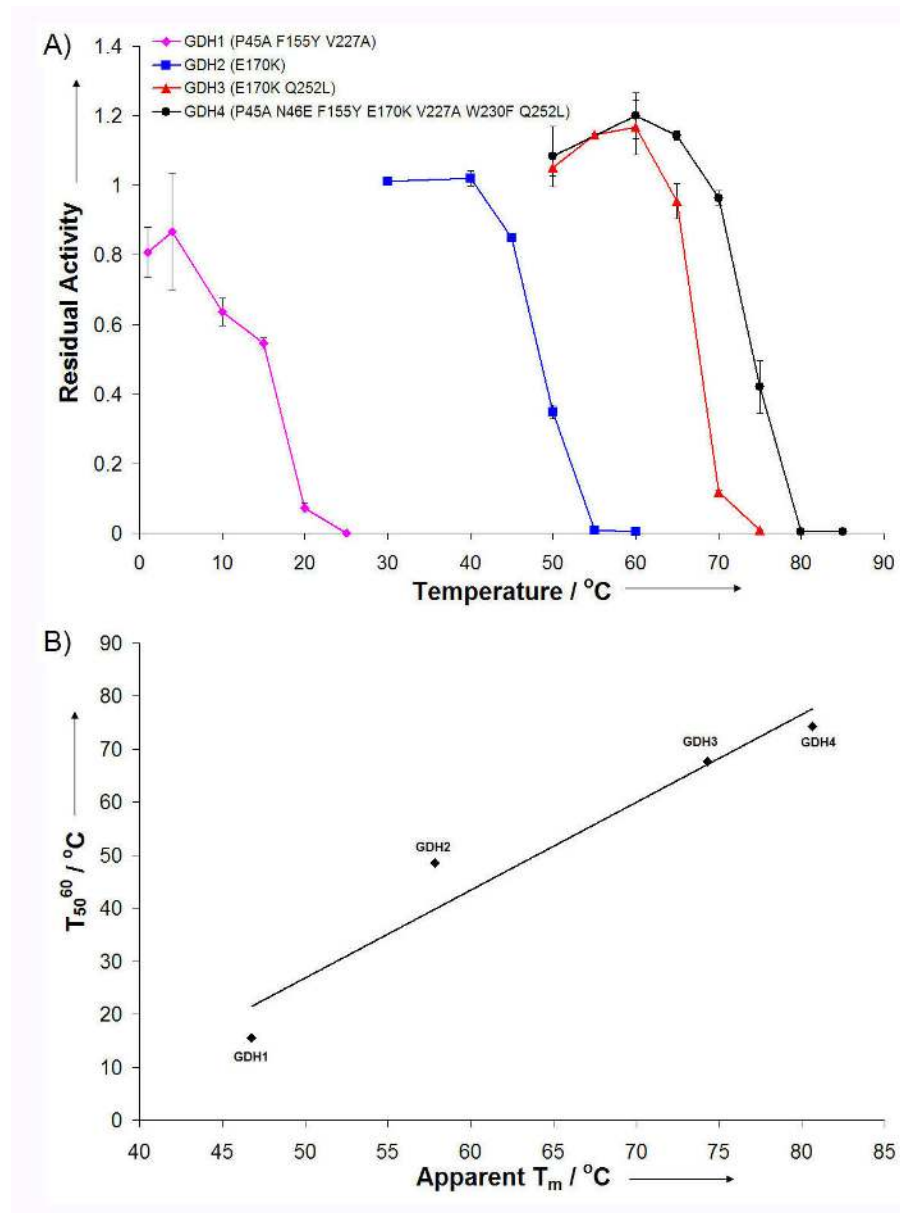
Figure 4.5. A) T_{50}^{60} for GDH variants. GDH variants were incubated for 60 minutes at the different temperatures. B) T_{50}^{60} against apparent melting temperature (T_m).

4.3.3. Organic solvent effects on glucose dehydrogenase

The stability of the GDH variants was studied in water-miscible organic solvent mixtures. The use of monophasic solvent sytems allowed a wide range of aqueous-organic compositions to be considered and avoids deactivation effects due to interfacial denaturation at the aqueous-organic interphase (45, 113). Other groups

Table 4.1. Half-lives (min) for GDHs at 25oC in 20% v/v organic solvent.

Solvent (DC)[a]	GDH1 (P45A F155Y V227A)	GDH2 (E170K)	GDH3 (E170K Q252L)	GDH4 (P45A N46E F155Y E170K V227A W230F Q252L)
ethylene glycol (18.7)	291	-[b]	-	-
1,2-propanediol (38.8)	4621	-	-	-
Ethanol (54.4)	28.5	37.5	4030	27397
acetonitrile (64.3)	<1[c]	<1	1.69	19.0
Acetone (78.2)	3.9	10.0	311	11533
1,4-dioxane (92.1)	<1	15.2	101	182

[a] Denaturing Capacity as calculated by Khemelnitsky et al. (1991). [b] No deactivation was observed after 20 days. [c] Deactivation occurred immediately upon addition of enzyme to aqueous-organic medium.

Table 4.2. $C_{50}^{0[a]}$ in %v/v for GDH variants.

	C_{50}^{0} ethylene glycol	C_{50}^{0} 1,2-propanediol	C_{50}^{0} ethanol	C_{50}^{0} acetone	C_{50}^{0} 1,4-dioxane
GDH1 (P45A F155Y V227A)	33.1	34.8	25.6	25.0	19.4
GDH2 (E170K)	33.0	32.7	37.4	45.7	35.2
GDH3 (E170K Q252L)	36.6	41.5	42.0	45.2	38.1
GDH4 (P45A N46E F155Y E170K V227A W230F Q252L)	32.0	32.4	37.2	44.5	33.7

[a] C_{50}^{0} values were obtained from the linear regressions.

have performed extensive studies of the effect of both miscible and immiscible organic solvents on enzymes (78, 97, 98, 141, 142). Of interest was the ability to rank miscible solvents based on denaturation capacity (DC) (75) and the observed characteristic, sigmoid deactivation behavior (75, 106). This sigmoidal drop-off in activity has also been observed in residual activity experiments to estimate T_{50}^{60} values—where values represent the temperature at which 50% activity (subscript) remains after 60 minutes incubation (superscript)—commonly used for quick screening of enzyme stability (6, 107, 132). To start, T_{50}^{60} values for GDH1-GDH4 were determined (Figure 4.5A). GDH demonstrated sigmoidal drop-off behavior and in addition, the T_{50}^{60}s increased with increasing apparent Tm (Figure 4.5B). Then a range of six organic solvents were selected along the range of DCs (75): ethylene glycol, 1,2-propanediol, ethanol, acetonitrile, acetone, and 1,4-dioxane—arranged in increasing order of DC (Table 4.1). Reactions containing increasing concentration of organic solvent in pH 8.0 phosphate buffer at 25°C were performed. For all solvents except acetonitrile, the initial rate of GDH decreased with increasing organic solvent

concentration in a linear matter with consistent R^2 values > 0.90 (see appendix, Figure A4.1), unlike the deactivation behavior reported by Khmelnitsky et al. (75). For the case of acetonitrile, a minimum initial rate was observed around 40% volume (%v/v) for two GDH variants (see appendix, Figure S4.2). Using the linear fits, C_{50}^0s—the concentration of organic solvent by %v/v where 50% initial activity is retained—were determined for all solvents except acetonitrile (Table 4.2). Although it is evident that the most labile variant (GDH1) has significantly lower stability (as determined by C_{50}^0) in ethanol, acetone and 1,4-dioxane than the others, no particular trend is observed between variant thermostability (T_{50}^{60}), organic solvent type and C_{50}^0s. Furthermore, the approach described by Khmelnitsky et al. (75) is unsuitable for GDH due to the low solubility of glucose, cofactor, and buffer at high organic cosolvent concentrations.

Half-lives where then determined in the presence of organic cosolvents. The difficulty encountered was finding suitable conditions of temperature and organic solvent that would allow appropriate characterization of all variants—the half-lives of the selected variants differ by five orders of magnitude under previously studied conditions (300 mM NaCl, 250 mM imidazole, 50 mM phosphate pH 8.0) (168). GDH variant half-lives determined at 25°C in 20% v/v organic solvent are shown in Table 4.1. Under these conditions no deactivation of the three more stable GDH variants was observed in ethylene glycol and 1,2-propanediol. In the case of 1,4-dioxane and acetonitrile the most labile variants denatured too quickly to measure half-lives effectively. However, for ethanol, acetonitrile, acetone and 1,4-dioxane the more thermostable variants show increased stability (as demonstrated by longer half-lives) at 20% v/v organic solvent.

Next, to minimize any thermal deactivation, protein deactivation in the organic solvent was characterized at 4^oC. This temperature was selected to prevent the sample from freezing (some freezing was observed upon sample perturbation at > 95% v/v for 1,4-dioxane which has a freezing point of 12^oC (154)), and at this temperature no significant loss of activity was observed for all variants after incubating for one hour (Figure 4.5A). GDH was added to increasing organic solvent concentrations, incubated for one hour at 4^oC and the residual activity measured in aqueous buffer (pH 8.0, 50 mM sodium phosphate)—analogous to T_{50}^{60}. The residual activity was then normalized with respect to the activity measured after one hour without organic solvent (Figure 4.6). It is evident that a minimum in activity is observed around 50% v/v for all variants in all solvents studied. This inverted bell-shape has also been observed by Griebenow and Klibanov for subtilisin and lysozyme (48). The authors studied α-helix content via FTIR in THF, 1-propanol, and acetonitrile. They concluded that in aqueous-organic solvent mixtures two effects simultaneously act on protein stability: 1) the propensity of a protein to denature increases with the increase in organic solvent content, 2) as water content decreases, protein conformational mobility decreases and thus its ability to acquire the thermodynamically dictated conformation. We surmise that, analogous to lysozyme and subtilisin (48), GDH has a higher propensity to denature in high organic solvent content but it is kinetically entrapped in a folded and thus active state. For all GDH variants significant activity was observed after one hour incubation in > 80% v/v of organic solvent.

In the regime of low organic solvent content (< 50% v/v), all GDH variants seem to decrease in residual activity with increasing organic solvent content. In this regime, C_{50}^{60} data at 4^oC (analogous to T_{50}^{60})—the concentration in which 50%

activity remained (subscript) after 60 minute incubation (superscript)—were obtained
(Table 4.3). These C_{50}^{60} values were then plotted against T_{50}^{60} and apparent T_m values
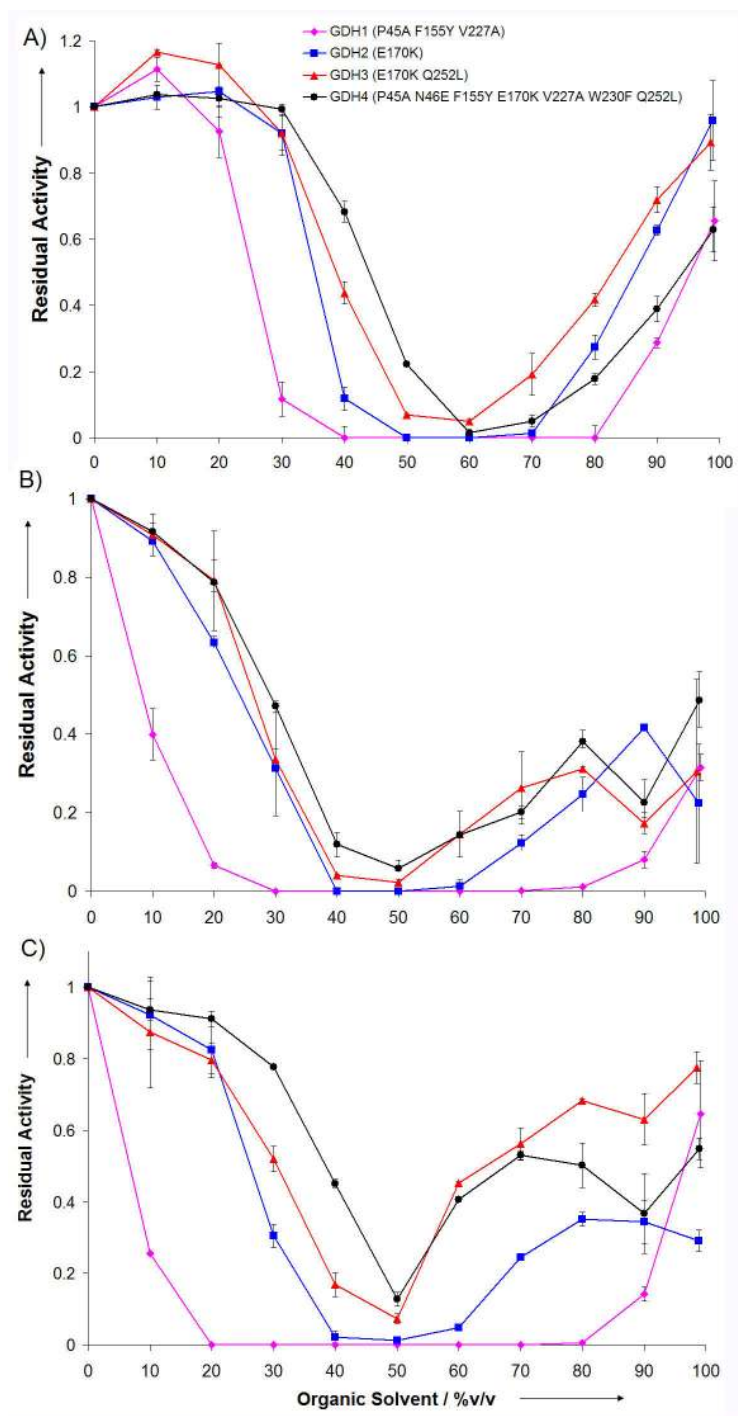
Figure 4.6. Residual activity of GDH variants after 60 minute incubation at 4°C in presence of organic solvents. A) Ethanol. B) Acetone. C) 1,4-Dioxane.

(Figure 4.7). A strong correlation between thermal stability and organic solvent stability exists, as noted by Cowan et al.(31, 113), with correlation coefficients (R^2) > 0.80 and > 0.95 between C_{50}^{60} values and T_m and T_{50}^{60}, respectively. To our knowledge, this work demonstrates for the first time that an enzyme variant constructed via a consensus-based approach resulted in improved organic solvent stability.
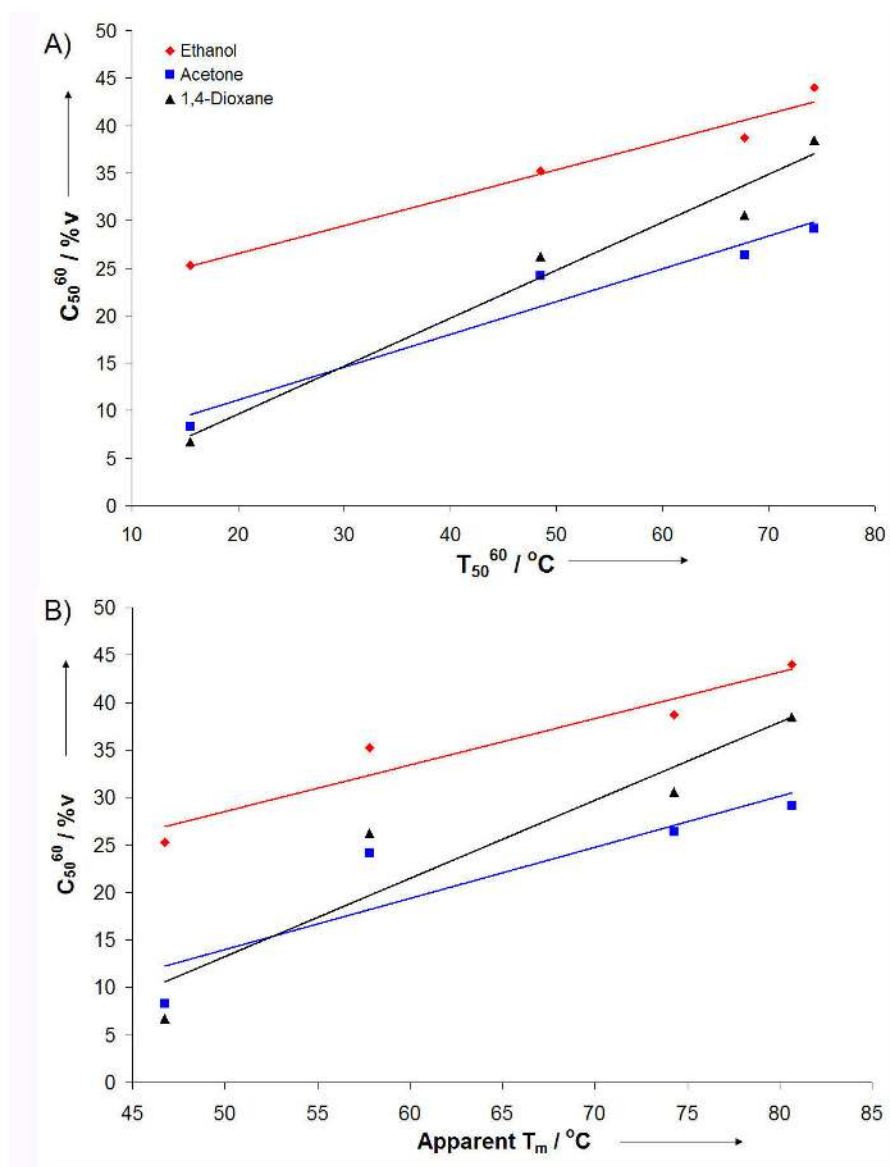
Figure 4.7. Plot of C_{50}^{60} and A) T_{50}^{60} and B) Apparent T_m for GDH variants.

4.4 Conclusions

In summary, we have demonstrated that thermostable GDH variants constructed via the structure-guided consensus method(168) show increased stability in the presence of denaturing salts and organic cosolvents. The two most temperature-stable variants, GDH3 and GDH4, also showed increased stability in the presence of I^-, NO_3^-, and Br^-—three chaotropic anions. As previously observed, NaCl clearly stabilized GDH (6, 7, 107, 176), as demonstrated by increased T_m and $k_{d,obs}$ values. However, even though NaCl increased the T_m (thermodynamic stability) of all variants by $> 7^o$C and the T_m for GDH4 was higher than GDH3 (92.8^oC vs 86.3^oC, at 0.5 M NaCl), this did not translate to higher $k_{d,obs}$ (kinetic stability) in the presence of NaCl.

The most stable variant, GDH4 (P45A, N46E, F155Y, E170K, V227A, W230F, Q252L) showed a > 2500-fold increase in tolerance—as determined by half-lives—to 20% v/v of acetone, acetonitrile and 1,4-dioxane at 25^oC when compared to GDH1 (P45A, F155Y, V227A). In addition, a 1.7-, 3.5-, and 5.7-fold improvement in C_{50}^{60} values at 4^oC were observed for ethanol, acetone and 1,4-dioxane for GDH4 relative to GDH1.

It was shown that in the absence of NaCl, GDH4 is more stable kinetically (T_{50}^{60}) and thermodynamically than GDH3. Under previous experimental conditions (300 mM NaCl, 250 mM imidazole, pH 8.0 50 mM sodium phosphate) this was not observed due to the presence of NaCl (168). Consequently, the incorporation of five additional mutations (N45A, N46E, F155Y, V227A, and W230F) did in fact result in overall increased stability, further proof of the predictive power of consensus-based methods.

While our results do not elucidate the factors governing GDH activity at high temperature, in salt solutions, or organic solvents, we believe the concept of corresponding states with respect to flexibility and stability is applicable(48, 49, 67, 181). It has been shown that enzyme conformational mobility is a key paramater allowing catalysis under given, distinct conditions. Závodsky et al. (181) and Griebenow et al. (49) have demonstrated how enzyme function is maintained—at high temperatures and in the presence of organic solvents, respectively—as long as native-like conformational mobility and stability is preserved. Consequently, thermostable GDH4 is able to preserve its mobility and stability under denaturing conditions due to increased structural stability incurred previously via mutagenesis (168).

The existence of a correlation between temperature and organic solvent has been shown (31, 113), however this is the first time a consensus-based method has been demonstrated to improve stability in the presence of organic solvents and chaotropic salts. The use of consensus sequences is one of many tools available for improving thermal stability of proteins (87) which here has been demonstrated to show promise in improving stability against other denaturants.

CHAPTER 5

RECOMMENDATIONS AND CONCLUSIONS

5.1. Recommendations

5.1.1 Improvement of high throughput screening assay for amines

A key requirement towards the development of a novel amine dehydrogenase is the use of a sensitive and selective assay. In this work, we described and implemented a modified NAD+-dependent fluorescent assay towards amines based on work by Tsotsou et al. (163) (Figure 2.13). In this assay a shift to alkaline pH resulted in a product the exact structure of which was unknown (4, 163). After examining the literature (64, 90, 91), probable pathways are summarized in Figure 5.1. At the instant of NaOH addition there are two competing reactions. At dilute (0.1-1.0 N) NaOH concentrations the nicotinamide-ribose linkage of NAD+ (I) is broken, producing nicotinamide (III) and subsequently adenosine monophosphate (IV) via breakage of the pyrophosphate link (73). This reaction has been shown not to be dependent on NaOH concentration (73). In the presence of strong alkali (5 N NaOH) a strongly fluorescent product is generated without cleavage of the nicotinamide-ribose linkage. J. W. Huff (64) noticed that the unreactive N^1-methylnicotinamide, when placed in an alkaline solution, produced the very reactive α-carbinol. The α-carbinol was found to condense with a variety of alcohols, ketones and aldehydes to yield highly fluorescent products—the ketones and aldehydes products being quite stable. We believe that analogous to N^1-methylnicotinamide—in the presence of NaOH—NAD+ goes through a carbinol intermediate (V) that upon reacting with an alcohol or ketone produces VI, which upon cyclization produces VII (64, 73).

Upon examining the possible pathways of NAD+ in the presence of NaOH, it is apparent that we likely generate a mix of fluorescent products that could

66

Figure 5.1. Probable pathways of the condensation reaction of NAD+.

result in increased variability and thus, loss of sensitivity. During screening the LeuDH variants, (see Chapter 2) a mean background of 865.09 RFU was obtained. This was ~2x higher than previously obtained during our assay development.

To reduce background, a higher concentration of NaOH could be added. This will help in shifting the produced NAD+ toward the desired fluorescent product. Furthermore, the addition of ketones, such as acetophenone (125) and acetone (64) result in the generation of stable fluorescent products with distinct excitation and emission maxima. This mixture of fluorescent products can be optimized towards the generation of a predominant species. The LeuDH screening assay was carried out in the presence of ~ 8 mM MIBK (substrate). The saturation of our screening assay with MIBK (solubility of MIBK in water is ~170 mM (154)) should not only increase the probability of finding a hit by saturating the enzyme but also as consequence result in a predominant alkali fluorescent product with defined spectrofluorometic characteristics (Figure 5.2).
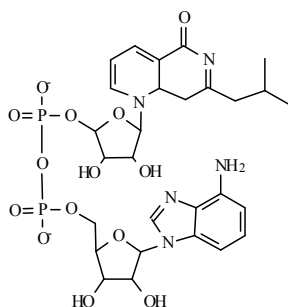
Figure 5.2. Probable alkali fluorescent NAD+ product in presence of MIBK.

5.1.2. Further protein engineering of LeuDH

Further protein engineering is required to an amine dehydrogenase of sufficient functionality+. However, deciding which protein engineering tool to use from the many available (5) is not trivial. Protein engineers use a variety of tools to adapt enzymes to novel substrates: error-prone PCR (23), DNA-shuffling (32), CASTing (129, 131), and site-directed mutagenesis, among others. For selecting the appropriate tool, previous protein knowledge (i.e., sequence, X-ray crystal structure,
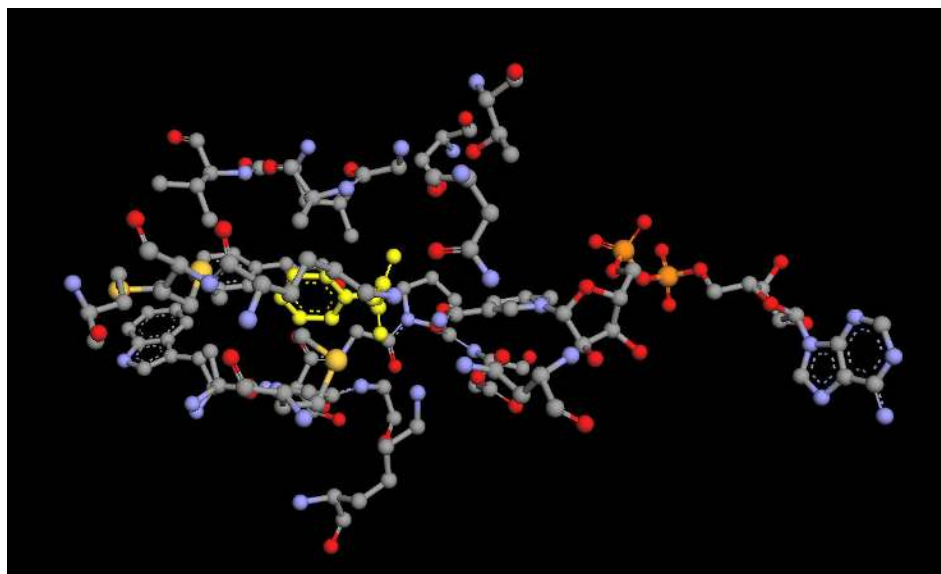
Figure 5.3. Amino acids ≤ 5.5 Å from phenylpyruvate (shown in yellow) in the PheDH structure 1BW9 (167).

mechanism) as well as the availability of a high throughput screening or selection assay will ultimately dictate the most suitable approach. If much is known on the enzyme of interest a rational approach can be taken. In this case, an amino acid is chosen for site-directed mutagenesis, where one amino acid is replaced by another. This approach requires extensive protein knowledge. A quasi-rational approach can be implemented where one or multiple amino acids are selected for focused randomization. In this case, the selected amino acid/s is/are mutated to selected or all other amino acids. Such an example of this technique is Combinatiorial Active-Site Saturation Test (CASTing) (129, 131, 132) or traditional site-saturated mutagenesis. These techniques have the advantage of a reduced library size while at the same time allows the search for non-obvious advantageous mutations.

Alternatively, directed molecular evolution can be used (5). This technique requires little protein knowledge *a priori*. With this approach, Darwinian evolution is carried out in a test tube. The most popular method used for directed evolution is error-prone PCR (ep-PCR). With this technique, as the name implies, the polymerase is stimulated to make mistakes by introduction of manganese ions and skewed deoxyribonucelotides concentrations during gene amplification via PCR. The genes

69

amplified via ep-PCR are prone to erroneous substitutions of nucleotides, resulting in amino acid substitutions during translation. The advantage of this technique is that little protein knowledge is required for its implementation and that a larger sequence space can be investigated. The drawback of this technique is that a robust, high throughput system is essential. Ultimately, "you get what you screened for"(4).

It is evident from our results that further mutagenesis of LeuDH is necessary for obtaining an amine dehydrogenase of sufficient functionality. Using X-ray crystal structures from PheDH and sequence alignments tools, a total of 22 amino acids in the vicinity ($\leq$ 5.5 Å) of the substrate were identified (Figure 5.3; Table 5.1). Using sequence alingment tools (59, 151, 152) the corresponding amino acids were determined in LeuDH and other amino acid dehydrogenases (AADHs) (Table 5.1). For sequences and hosts of AADHs used see Appendix, PRALINE alignment and consensus sequence for AADHs. Lys68, Asp115, and Lys80 are implicated in the catalytic mechanism (8, 138, 140). It is evident from Table 5.1 that Lys68, Asp115 and Lys80 are conserved in all AADHs and this is not surprising since all amino acids are produced from the corresponding α-keto acid and Lys68 stabilizes the carboxyl group, Lys80 stabilizes the carbonyl group and Asp115 is believed to stabilize the imino intermediate (20, 167).

Studies by Baker et al (8) has identified five glycine residues (41, 42, 77, 78, and 290) that contribute to the shape of the LeuDH active site, three of which are in close proximity to the substrate. Furthermore, Sekimoto et al (139), mutated a tetrapeptide sequence (Gly-Gly-(Gly/Ala)-Lys80) upstream of Lys80. They concluded that the conserved glycyl residues are important for fine-tuning the positioning of the ϵ-amino group of Lys80—Gly77 and Gly78 variants had markedly decreased thermal stabilities (139). Leu40, Gly41, Ala113, Val291 and Val294 from LeuDH have been recognized via structural comparisons with GluDH to interact with the aliphatic side chain of L-leucine (8, 9, 18).

Table 5.1. Amino acids ≤ 5.5 Å from phenylpyruvate. Column 1, amino acids in the vicinity of the substrate (PheDH structure 1BW9). Column 2, corresponding amino acids in LeuDH from *Bacillus sphaericus* structure 1LEH. Column 3, corresponding amino acid in LeuDH from *(Geo)Bacillus stearothermophilus*. Column 4, recurring amino acids in LeuDHs. Column 5, corresponding amino acids in amino acid dehydrogenases—a single entry means amino acid is conserved.

1BW9 (PheDH from Rhodococcus sp. M4)	1LEH (LeuDH from Bacillus sphaericus)	LeuDH (Geo)Bacillus stearothermophilus	LeuDHs	AADHs
Ala38	Leu40	Leu40	Leu	LeuDH (Leu), PheDH (Leu), GluDH (K), AlaDH (Gly)
Gly39	Gly41	Gly41	Gly	Gly
Gly40	Gly42	Gly42	Gly	Gly
Met63	Met65	Met65	Met	AlaDH (Ala), GluDH (Met, Gln)
Lys66	Lys68	Lys68	Lys	AlaDH (Ile, Val, Arg)
Met67	Asn69	Asn69	Asn	PheDH (Cys), AlaDH (Cys, Ala), GluDH (Cys, Ala, Asn)
Met74	Leu76	Leu76	Leu	PheDH (Phe), GluDH (Tyr, Met, Ile)
Lys78	Lys80	Lys80	Lys	AlaDH (Met, Leu)
Trp114	Ile111	Ile111	Ile, Tyr	PheDH (Tyr, Trp), GluDH (Ile, Val), AlaDH (Leu, Val)
Thr 115	Thr112	Thr112	Thr, Ile	PheDH (Thr), GluDH (Pro), AlaDH (Thr, Ala)
Gly116	Ala113	Ala113	Ala, Thr	PheDH (Gly), GluDH (Ala), AlaDH(Gln, Lys, Glu)
Pro117	Glu114	Glu114	Glu, Ala	PheDH (Thr, Pro), GluDH (Pro, Gly), AlaDH (Ala)
Asp118	Asp115	Asp115	Asp	AlaDH (Ala, Leu)
Phe137	Thr134	Thr134	Thr	PheDH (Ala, Val, Asn, Phe), GluDH (Thr), AlaDH (Ala)
Ser149	Pro146	Pro146	Pro	PheDH (Ser, Thr), Glu (Arg), AlaDH (Glu)
Thr153	Val149	Ala149	Val, Ala	PheDH (Pro), GluDH (Ala, Thr), AlaDH (His, Lys, Gln)
Asn262	Asp261	Asn261	Asn	AlaDH (Asp)
Asn288	Asn287	Asn287	Asn	GluDH (Ser, His), AlaDH (His)
Gly291	Gly290	Gly290	Gly	AlaDH (Ala, gly)
Ala292	Val291	Val291	Val	PheDH (Leu, Ala), AlaDH (Val, Ile)
Leu295	Val294	Val294	Val	PheDH (Val, Leu), GluDH (Ser), AlaDH (Met)
Val296	Ala295	Ala295	Ala	PheDH (Ala, Val), GluDH (Tyr, Gly, Ala), AlaDH (Pro)

With this information we recommend that future protein engineering efforts focus on Leu40, Gly41, Gly42, Ala113, Glu114, Asp115, Val291 and Val294. For convenience, these amino acids will be divided into three groups: Leu40-Gly42 (red or group A), Ala113-Asp115 (blue or group C), and Val291 and Val294 (cyan or group D) (Figure 5.4). Mutagenesis of Leu40-Gly42 will help optimize the shape of
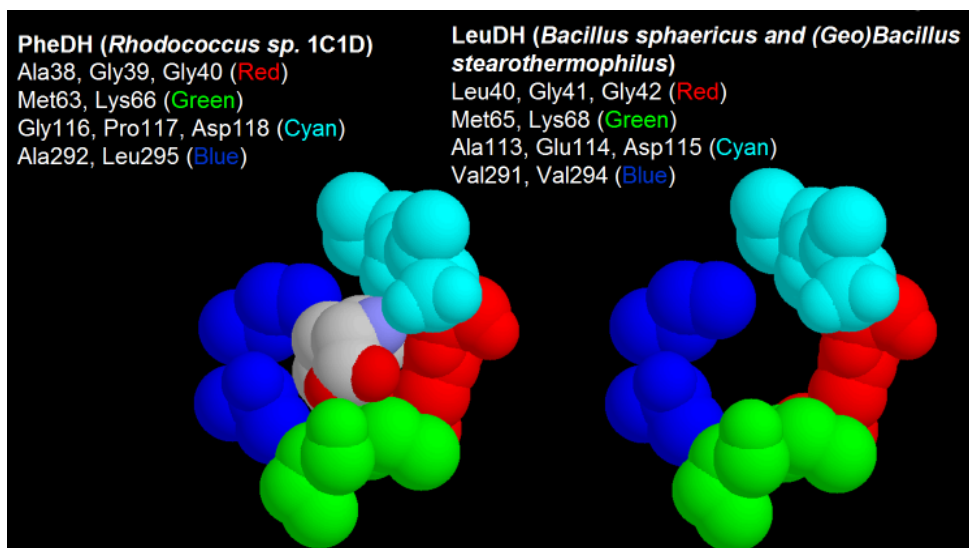
Figure 5.4. Selected amino acids for further mutagenesis. Left, amino acid positions in PheDH and 1C1D active site with bound L-phenylalanine. Right, amino acid positions in LeuDH and 1C1D active site without L-phenylalanine.

the active site (Gly41 and Gly42) and enhance protein interactions with the aliphatic side-chain of L-leucine (Leu40) for improved binding of the substrate(8). The mutagenesis of Ala113-Asp115 will allow for optimization of the protein interactions with the (unstable) imine intermediate and with a catalytic water molecule (20). Finally, mutagenesis of Val291 and Val294 will further improve binding by optimizing protein-substrate (substrate side-chain) interactions (8, 9, 18). The selected amino acids, with bound phenylalanine are shown below (Figure 5.4).

For future mutagenesis we recommend CASTing, also known as iterative site-saturation mutagenesis (ISM) (129-131). This approach will allow us to mutate the selected amino acids by groups. By focusing on these "hot sites", the sequence space is significantly reduced. A schematic representation of ISM is shown below (Figure 5.5). Various "routes" to improved activity can be used, but an exhaustive study is not necessary. Mutagenesis of group B (green) has already been performed (see Chapter 2). If mutagenesis is carried out using NNK (32 codons for 20 amino acids) degenerate codons at the remaining positions a total of at least 199,392 clones need to be screened, or > 2000, 96-well plates. To reduce the library size NDT (12 codons for 12 amino acids) or NNT (16 codons for 15 amino acids) degenerate primers should be
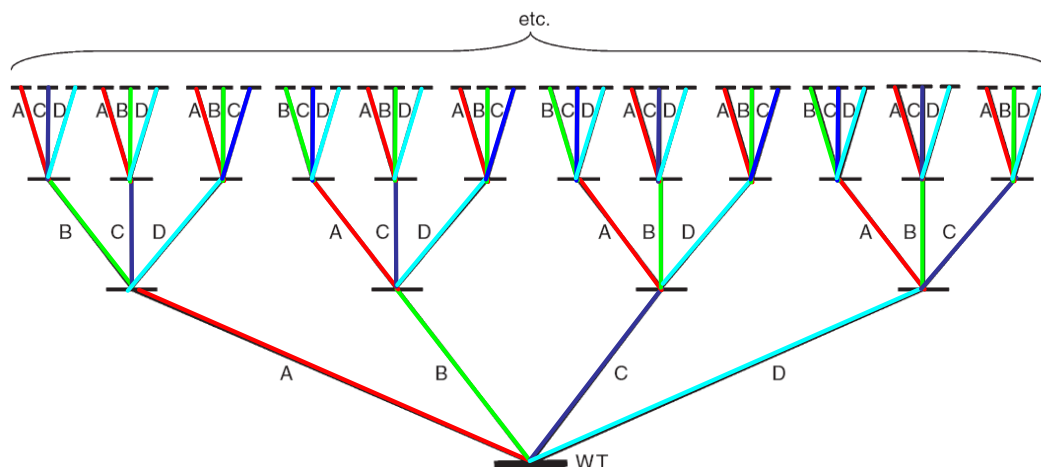
Figure 5.5. Schematic representation of ISM. The four sites selected are randomized in series. The best variant from one group is selected for further mutagenesis at the different groups. Adapted from Reetz 2007 (130).

used. This will reduce the library size by at least ~7.8x, resulting in 25,303 clones or ~263, 96-well plates.

5.1.3. Other scaffolds for engineering towards an amine dehydrogenase

The best enzyme candidates for engineering towards an amine dehydrogenase should have: known chemical mechanisms, available X-ray structures, high activity towards natural substrate, and slight activity towards desired substrate. The most apparent alternate enzyme candidates would be other amino acid dehydrogenases. One of the most characterized is PheDH from *Rhodococcus sp.* M4 (20, 21, 167). X-ray crystal structures with substrates and analogs (167) and detailed chemical mechanism (167) are available and specific activities as high as 285 U/mg (for PheDH from *Bacillus sphaericus*) have been recorded (111). Furthermore, PheDH is
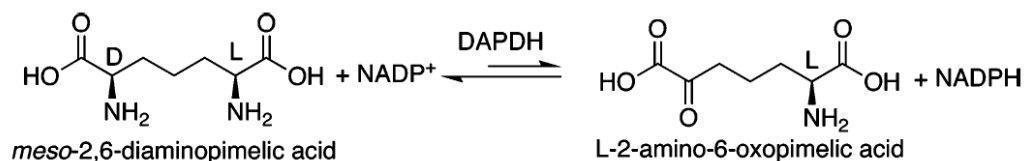
Figure 5.6. *In vivo* reaction performed by DAPDH. From Vheda-Peters 2006(169).

73
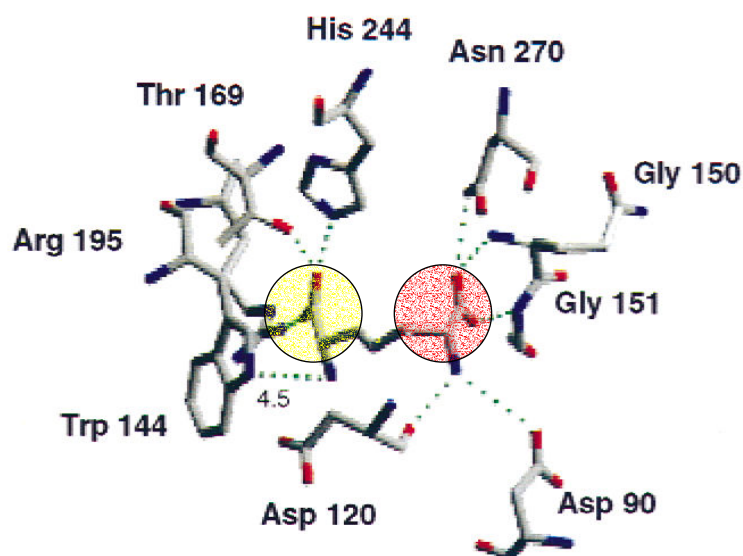
Figure 5.7. *meso*-DAP bound at the active site. L-amino center shown in yellow, D-amino center shown in red. Adapted from Cirilli 2000 (28).

known to accept a wide array of α-keto acids and amino acids (21, 111). The similarity in chemical mechanism and protein structure with LeuDH would allow for a straightforward transition and application of acquired knowledge (e.g., beneficial mutations and screening assay) to PheDH.

Another interesting candidate for protein engineering is *meso*-2,6-diaminopimelate dehydrogenase (DAPDH) (28, 135). This enzyme has the unique trait of reductively aminating α-keto acids into the corresponding D-amino acids (Figure 5.6). The enzyme complex with bound meso-aminopimelic acid and other analogs has been resolved (28, 135) (Figure 5.7). As can be observed in Figure 5.7, this enzyme has very defined residues that stabilize the D and L amino acid groups in the substrate, *meso*-2,6-diaminopimelic acid. Recently, Vedha-Peters et al. (169) where able to engineer the active site to accept a variety of substrates while retaining enantioselectivity. The authors used a combination of rational design (mutagenesis of the L-amino acid-interacting residues) and directed evolution. This enzyme has the potential of allowing a complete new set of chiral amines to be produced.

5.1.4. Further stabilization of glucose 1-dehydrogenase: introduction of inter-subunit disulfide bonds

GDH is an obligate tetramer and deactivation follows after loss in quaternary structure (subunit dissociation). Upon increase in pH, GDH is known to dissociate into dimers resulting in enzyme inactivation (6, 7). We believe temperature inactivation proceeds via the same route. It was demonstrated by Baik et al. (7) that inactivation of the enzyme carrying either the E170K or Q252L mutations occurred via dissociation from a tetramer to a dimer without significant changes in the CD spectra. This suggests that even though dissociation takes place, secondary structure is retained. This is in agreement with our results from Chapter 3 and 4. Five of the total 7 amino acids substitutions combined were at the interface, thus for improved stability the subunit-subunit interface needs to be reinforced (168). Furthermore, it has been shown that in the presence of NaCl, GDH4 (P45A N46E F155Y E170K V227A W230F Q252L) has a high apparent melting temperature as demonstrated via CD but this did not translate to an increase in kinetic stability (determined via half-life) when compared to other variants. This suggests that the enzyme dissociates prior to unfolding, resulting in loss of catalytic activity.

Introduction of disulfide bonds at the subunit-subunit interface should improve kinetic stability by covalently linking the subunits together and preventing dissociation. Disulfide bonds can make significant contributions to protein stability primarily by decreasing conformational chain entropy of the denatured protein (30, 42, 114, 161, 173, 182). The introduction of disulfide bonds for increased protein stability has successfully been applied to thermolysin-like proteases (95, 96), T4 lysozyme (99), subtilisin E (158), ribonuclease H (71) and xylanase (170).
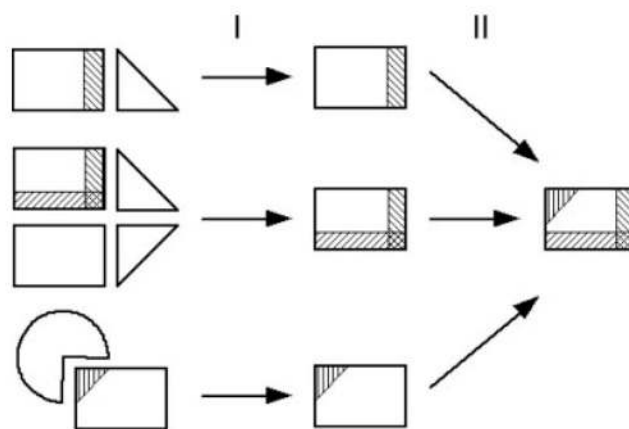
Figure 5.8. Schematic representation for determination of protein-protein interaction sites. From Porollo 2007(123).

First, amino acids at the subunit-subunit interface of GDH were determined. This was done using a program that integrates enhanced relative solvent accessibility (RSA) with high resolution data—SPPIDER (123). SPPIDDER uses machine learning to combine protein interaction data, sequence-based data and structural features for enhanced recognition of protein-protein interaction sites (123). RSA is defined as a ratio of the solvent exposed surface area (SA) of *i*-th amino acid residue in a given structure and some maximum value of exposed surface area (MSA) for the same amino acid (Equation 5.1).

$$RSA_i = \frac{SA}{MSA}(100) \qquad\qquad (5.1)$$

Porollo et al. (123) defined MSA as the exposed area obtained from an extended conformation of a tripeptide with the residue of interest as the central residue. In the case of GDH, the program compares the relative solvent accessibility of all amino acids between a subunit in the tetramer (SA) and the free monomer (MSA). A schematic representation of the procedure used to map interaction sites from multiple complexes is shown in Figure 5.8. This program is available from
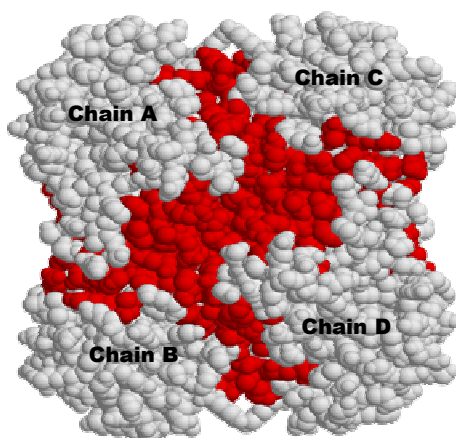
Chain	Residue numbers (Preserving PDB numeration)
A	69 101 102 104 106 109 110 113 114 117 118 121 122 125 129 132 148 149 151 153 156 159 160 163 164 166 167 168 170 171 172 174 175 176
B	69 101 102 104 106 109 110 113 114 117 118 121 122 125 129 132 148 149 151 153 156 159 160 163 164 166 167 168 170 171 172 174 175 176

Chain	Residue numbers (Preserving PDB numeration)
A	147 149 150 151 152 153 154 155 199 208 209 212 250 252 253 254 255 256 257 258 259 260 261
C	147 149 150 151 152 153 154 155 199 208 209 212 250 252 253 254 255 256 257 258 259 260 261

Chain	Residue numbers (Preserving PDB numeration)
A	1 2 26 170 174 177 178 180 182 214 215 217 219 223 226 227 230 235 237 238 239 240 241 242 244 245 246 247 248 249 252 256 257
D	1 2 26 170 174 177 178 180 182 214 215 217 223 226 227 230 235 237 238 239 240 241 242 244 245 246 247 248 249 252 256 257

Figure 5.9. Interfacial amino acids of GDH. Interfacial amino acids are shown in red (top) and the corresponding number is shown to the right (bottom).

(123, 124). Using the tetramer structure of GDH from *Bacillus megaterium* IWG3 (180) and the solvent accessibility based protein-protein program (123, 124) residues at the interface were determined (Figure 5.9). *Bacillus megaterium* IWG3 and *Bacillus subtilis* share 83.5% amino acid identity and are 88.9% similar. Equipped with the interfacial amino acids, the next step is identifying among this reduced set, potential pairs for disulphide bond formation.

SSBOND—a computer program—was used to identify potential sites for engineering a disulfide bond (55). SSBOND is available from (55). This algorithm starts by generating Cβ positions from the N, Cα and C atom coordinates provided from a three-dimensional structure (for a schematic with the nomenclature of the dihedral angles in a disulfide bond see Appendix, Figure A5.1). An initial screen selects residue pairs based on Cβ–Cβ distances. For this limited set of pairs, Sγ positions are generated which satisfy a predefined (55) set of ideal values for all other bond lengths and bond angles. Finally, the new conformations are subjected to energy minimization to remove large deviations from ideal geometry and their final energies and pairs are reported in the output.

GDH PDB files based on the crystal structure from *Bacillus megaterium*
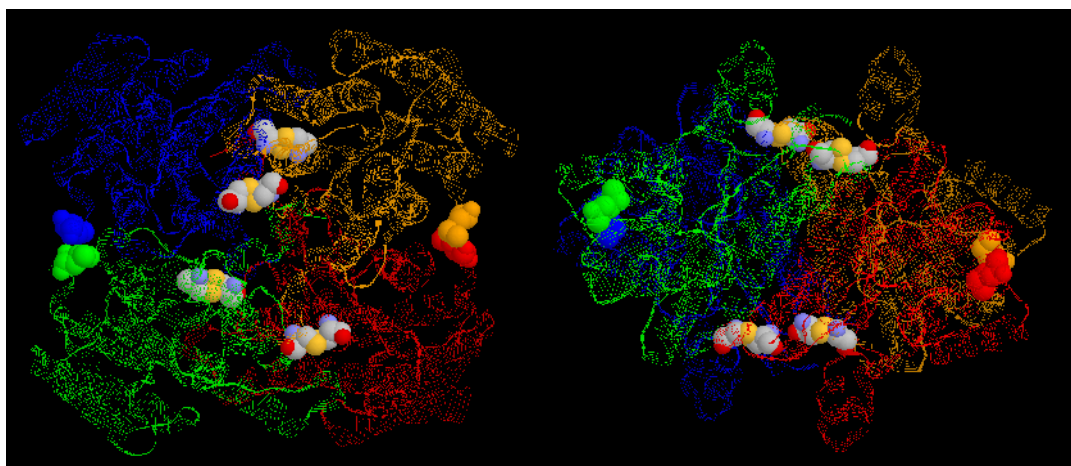
Figure 5.10. Interfacial disulfide bond identified via SSBOND. Chain A (blue), chain B (orange), chain C (red), chain D (green). N terminal amino acids shown in spacefill and colored by chain. P153C and E173C shown in spacefill and colored CPK.

IWG3 were created, with only interfacial amino acids. One file was created for each protein interface: AB, AC, and AD. These PDB structures were loaded into the program (SSBOND) and the output was screened for inter-subunit disulphide bonds. Only one pair was identified, Pro153-Glu175. Mutagenesis of these two positions to cysteines would result in the creation of four disulphide bonds, two between chain A and B and two between chain D and C (Figure 5.10). Unfortunately, no potential pairs were identified between the AD, AC, BC and BD interface. The N-terminus, where the 6xHis-tag is located, lies in the AD and BC interface (Figure 5.10).

A Pro153Cys and Glu175Cys variant was constructed starting from GDH4. The final GDH variant had a total of nine mutations: P45A, N46E, P153C, F155Y, E170K, E175C, V227A, W230F, and Q252L. The 6xHis-tag purified variant had a specific activity of ~ 0.7 U/mg and a half-life ~20 min at 65°C in elution buffer (300 mM NaCl, 250 mM imidazole in 50 mM phosphate pH 8.0). These experiments were carried out without oxidation of the disulfide bonds. It is evident that these two mutations (P153C and E175C) destabilized the protein and reduced the activity (GDH4 has a specific activity ~120 U/mg and a half-life of ~3.5 days under the same conditions). Sequence alignment shows that Pro153 is conserved in 9 out of 13 GDHs aligned while Glu175 is conserved in all 13 (see Appendix, PRALINE alignment and consensus sequence). Furthermore, Pro153 and Glu175 are located in a turn and α-helix, respectively. Pro and Glu are known to stabilize these types of secondary structures (94). Further studies without an N-terminal 6xHis-tag should be performed to test its influence on stability.

5.1.5. Applications of the structure-guided consensus concept: cellulases

With prices of crude-oil rising, the need for alternative carbon sources for fuels and materials is evermore important. Lignocellulosic matter as agricultural, industrial and forest residuals account for the majority of biomass in the world (81). Lignocellulosic material is primarily composed of 35-50% cellulose, 25-35% hemicellulose, 10-25% lignin, and pectins (81, 92). The hydrolysis of these sugar

polymers for generating value-added products is a great challenge due to: the complexity of the enzymes necessary for hydrolysis, the lack of an organism that produces the necessary enzymes in sufficient quantities, the crystallinity and heterogeneity of the cellulose sources, and feedback product inhibition of the enzymes or repression of metabolic pathways (81, 92, 117). In this discussion we will focus in the improvement of the enzymes required for hydrolysis—specifically cellulases.

Cellulases are divided into three major groups: endoglucanases, cellobiohydrolases (exoglucanases), and β-glucosidases (81). An impediment hindering cellulases is the recalcitrance of cellulosic biomass. This is primarily due to the crystallinity of the substrate, forming fibrils with such tight packing that makes it difficult for even water to penetrate (92). As a result, various mechanical and chemical pretreatment methods have been developed for pretreatment prior to subsequent hydrolysis and fermentation (56, 66, 172). However, the presence of residuals from the pretreatment often results in enzyme deactivation and inhibition. Thus, the development of robust, stable cellulases would allow for improved performance and increase in sugar yields, all vital to reducing costs (117).

Engineering cellulases is a difficult task due to the complexity of the enzyme, the dynamic enzyme-substrate interaction, the heterogeneous, insoluble substrate and the overall limited understanding of the system(126). A general trait of most cellulases is a modular structure often including both catalytic and carbohydrate-binding modules united by an interdomain linker (92, 126). The binding module is involved in attachment of the cellulase to the cellulose surface bringing the catalytic domain in close proximity to the substrate (92). The bi-functional linker is believed to provide spatial separation via a glycosilated, rigid part whereas a flexible, hinge-like section allows autonomous domain function (126). On the other hand, there are no definitive relationships between cellulase activities on soluble substrates and those on insoluble substrates. Consequently, soluble substrates should not be used for screening improved cellulases (117).

Although some successes have been reported (137), this lack of enzyme-substrate understanding and its complexity makes rational design extremely difficult. Most work via rational design has focused on elucidating the chemical mechanism of the enzymes (126). In contrast, directed molecular evolution allows for mutagenesis without a complete understanding of the enzyme structure or function. However, the modular nature of the enzyme permits for much of the enzyme variants to be inactive and for much of the effort to be focused in undesired areas. Furthermore, the insolubility of the substrate makes high-throughput screening difficult. We believe that structure-guided consensus (120, 168) is a more attractive approach towards engineering stable cellulases. This approach would allow for the binding, catalytic and linker domain to be optimized separately. Furthermore, since this approach generates a reduced number of variants a more thorough enzyme-substrate characterization could be carried out.

5.2. Conclusions

All in all, this work describes progress towards the development of a novel dehydrogenase for synthesis of chiral amines starting from a leucine dehydrogenase (Chapter 2). Using site-saturated mutagenesis two key residues were mutated en route to enhanced binding and catalytic activity towards an unnatural substrate. A NAD+-dependent fluorescent high throughput assay was optimized for screening.

Then, the traditional consensus method was modified by utilizing protein structural information and applied towards improving the overall stability of a glucose 1-dehydrogenase. This approach not only significantly enhanced thermal stability (Chapter 3) but also incurred stability towards other denaturants—specifically organic solvents and salt solutions (Chapter 5).

Future steps towards the development of an amine dehydrogenase should focus on those amino acids that surround the active site. Other protein scaffolds to consider for the development of a novel amine dehydrogenase should be other amino acid dehydrogenases, in particular phenylalanine dehydrogenase, due to the wealth of

knowledge available. Another candidate to consider is *meso*-2,6-diaminopimelate dehydrogenase. This enzyme is of particular interest since it has been demonstrated to synthesize a wide array of D-amino acids, a chemistry not common in nature (nature is based on L-amino acids).

Further stabilization efforts on glucose 1-dehydrogenase should certainly focus at the subunit-subunit interface. Our results suggest that this enzyme dissociates prior to unfolding, implying that further significant stabilization will be achieved by improving the interface and not by stabilizing intra-subunit interactions. An attempt, at further stabilizing glucose 1-dehydrogenase by insertion of inter-subunit disulphide bonds resulted in loss of activity and stability. Sequence alignments and inspection of secondary structure showed that these residues are highly conserved among glucose dehydrogenases and that mutagenesis at these positions to cysteines did not favor the present secondary structures.

Application of the structure-guided consensus method will prove useful to complex proteins whose substrates are not amenable to high throughput screening. A good example would be cellulases. The modular nature of cellulases and the crystallinity, low solubility and heterogeneity of their substrates makes this an ideal test case for the structure-guided consensus.

APPENDIX: Additional Figures
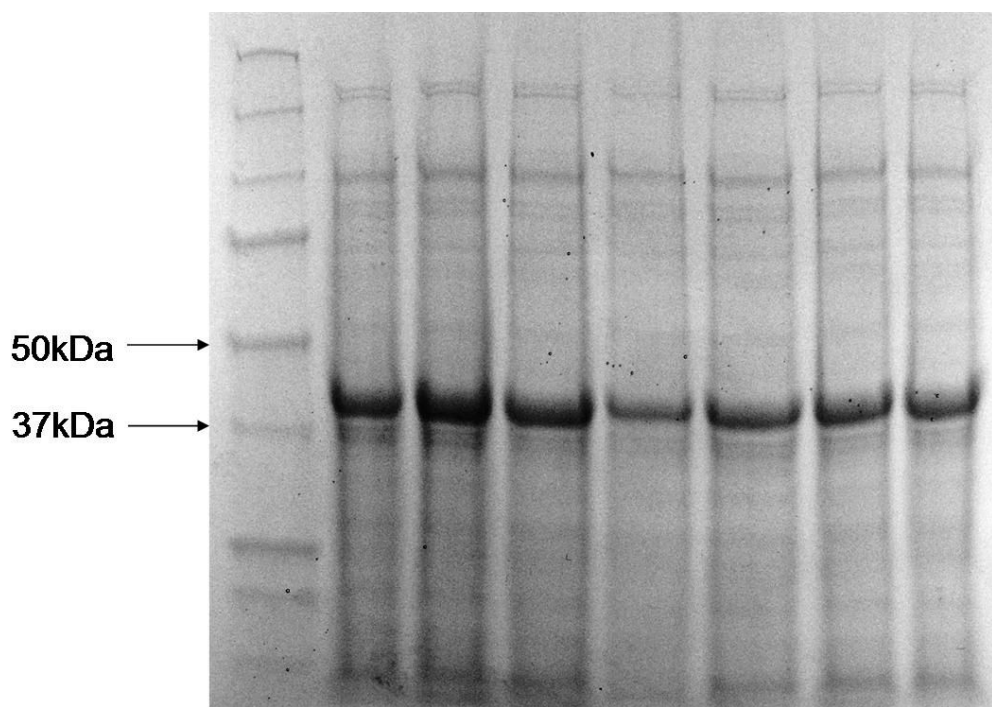
Figure A2.1. Representatibe SDS-PAGE of overexpression of LeuDH variants grown in 96-well plates with MagicMedia™. Random wells were selected for testing overexpression. Left lane, protein markers.

LeuDH Library Met65NNK Lys68NNK
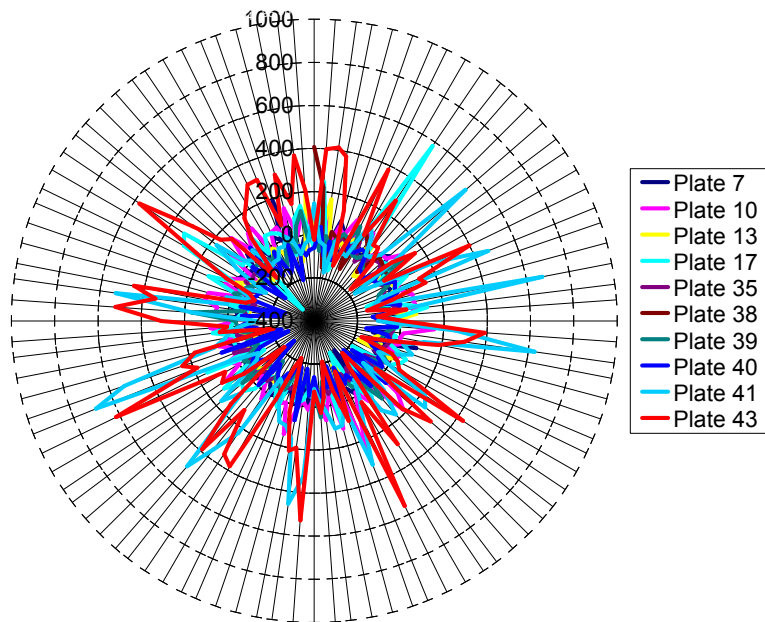
Figure A2.2. Plot 2, screening of LeuDH variants.

LeuDH Library Met65NNK Lys68NNK
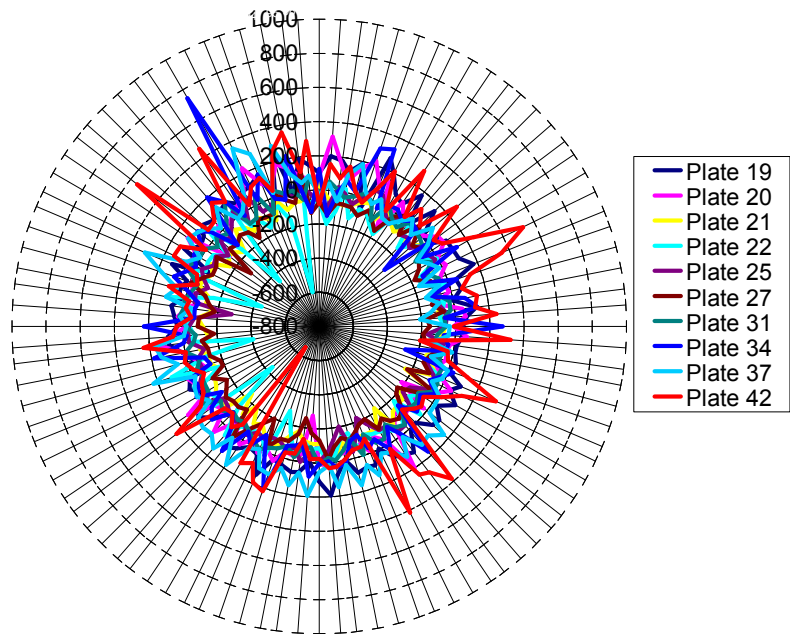
Figure A2.3. Plot 3, screening of LeuDH variants.

LeuDH Library Met65NNK Lys68NNK
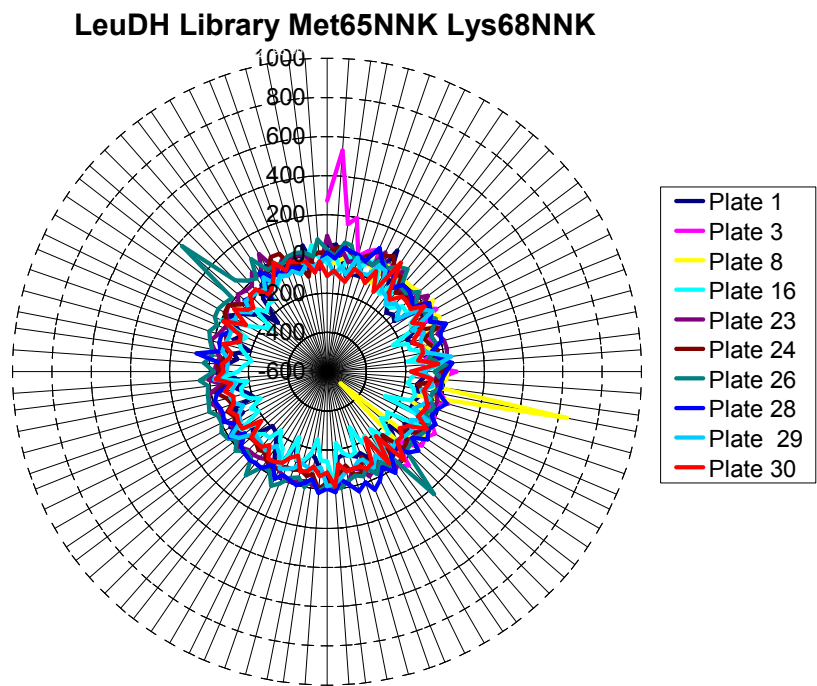
Figure A2.4. Plot 4, screening of LeuDH variants.

LeuDH Library Met65NNK Lys68NNK
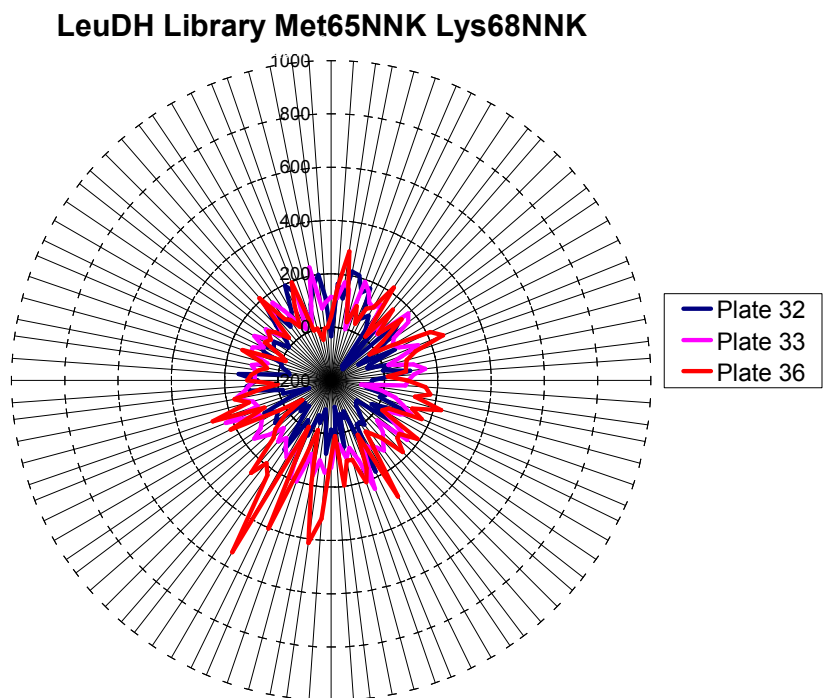
Figure A2.5. Plot 5, screening of LeuDH variants.
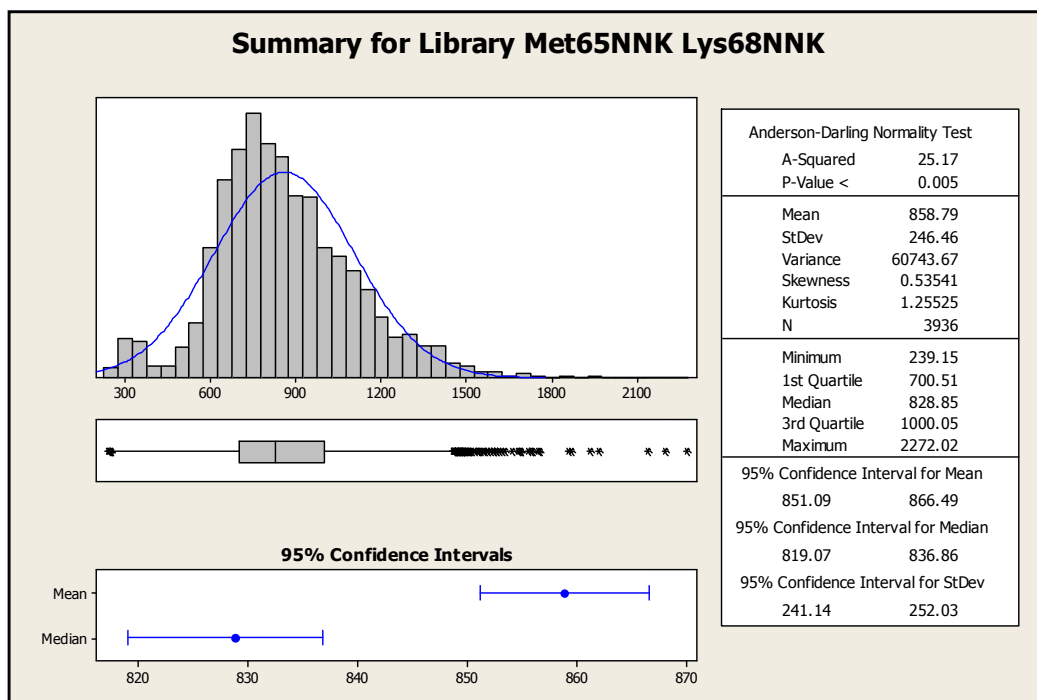
Figure A2.6. Statistical analysis of Met65NNK Lys68NNK library.
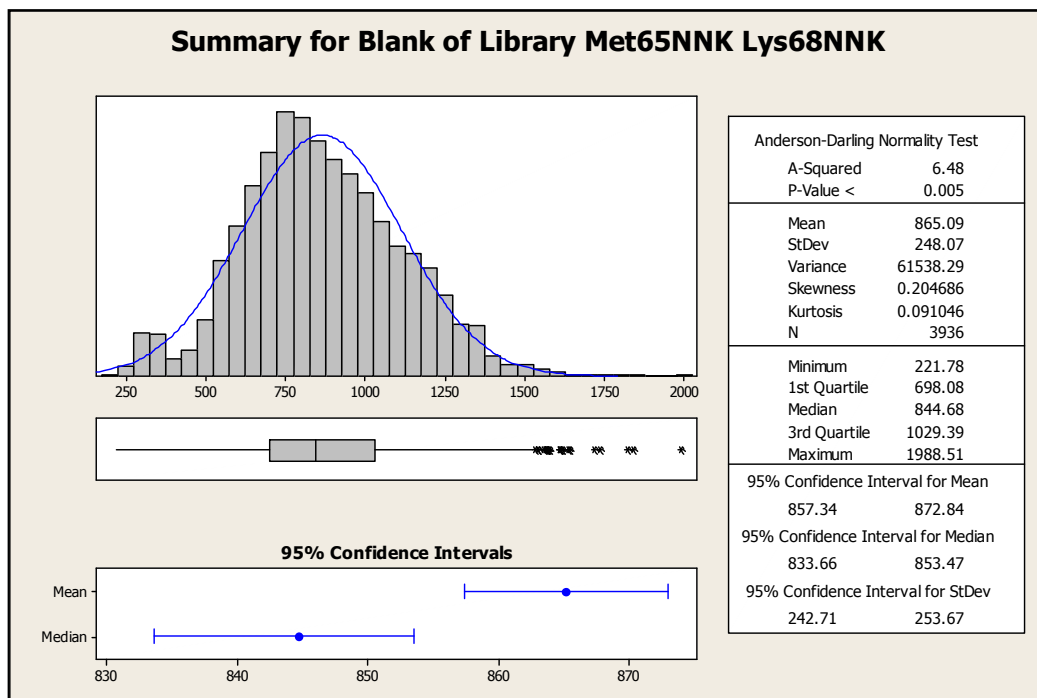
Figure A2.7. Statistical analysis of the background for Met65NNK Lys68NNK library.
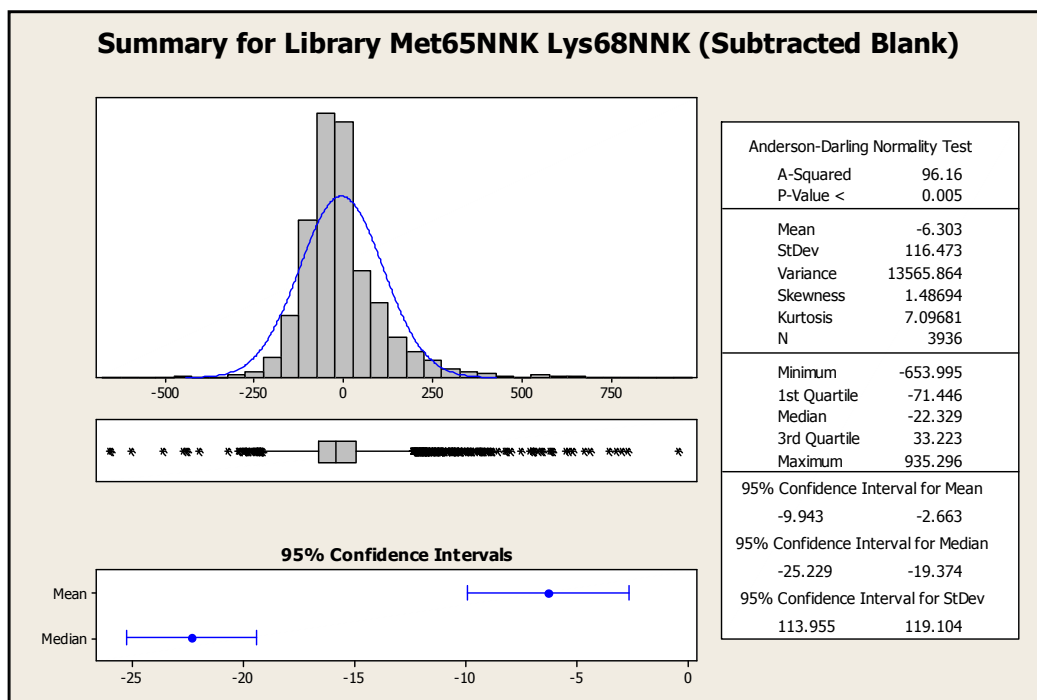
Figure A2.8. Statistical analysis of Met65NNK Lys68NNK library with substracted background.

Figure A2.9. Statistical analysis of top 20 LeuDH variants.
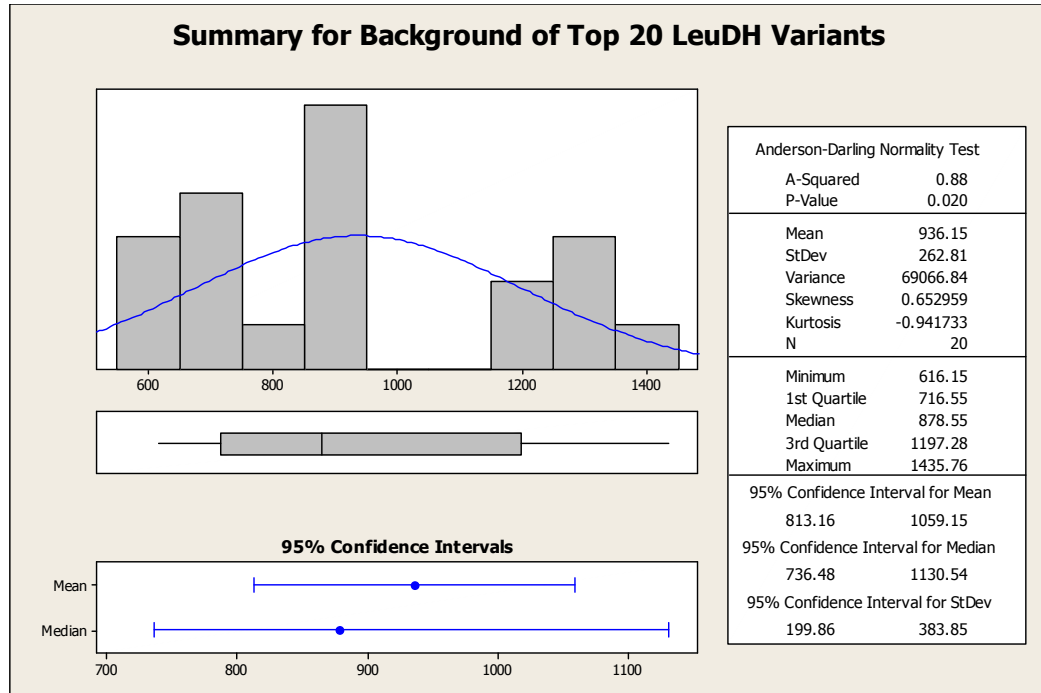
Figure A2.10. Statistical analysis of the background of the top 20 LeuDH variants.
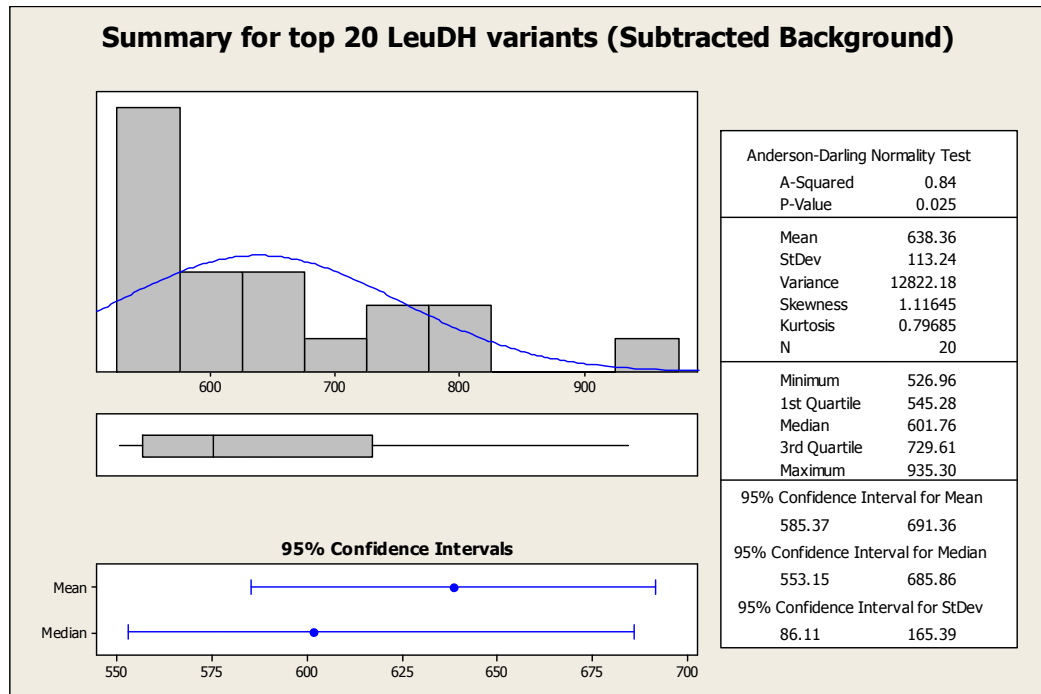
Figure A2.11. Statistical analysis of the top 20 LeuDH variants with subtracted background.
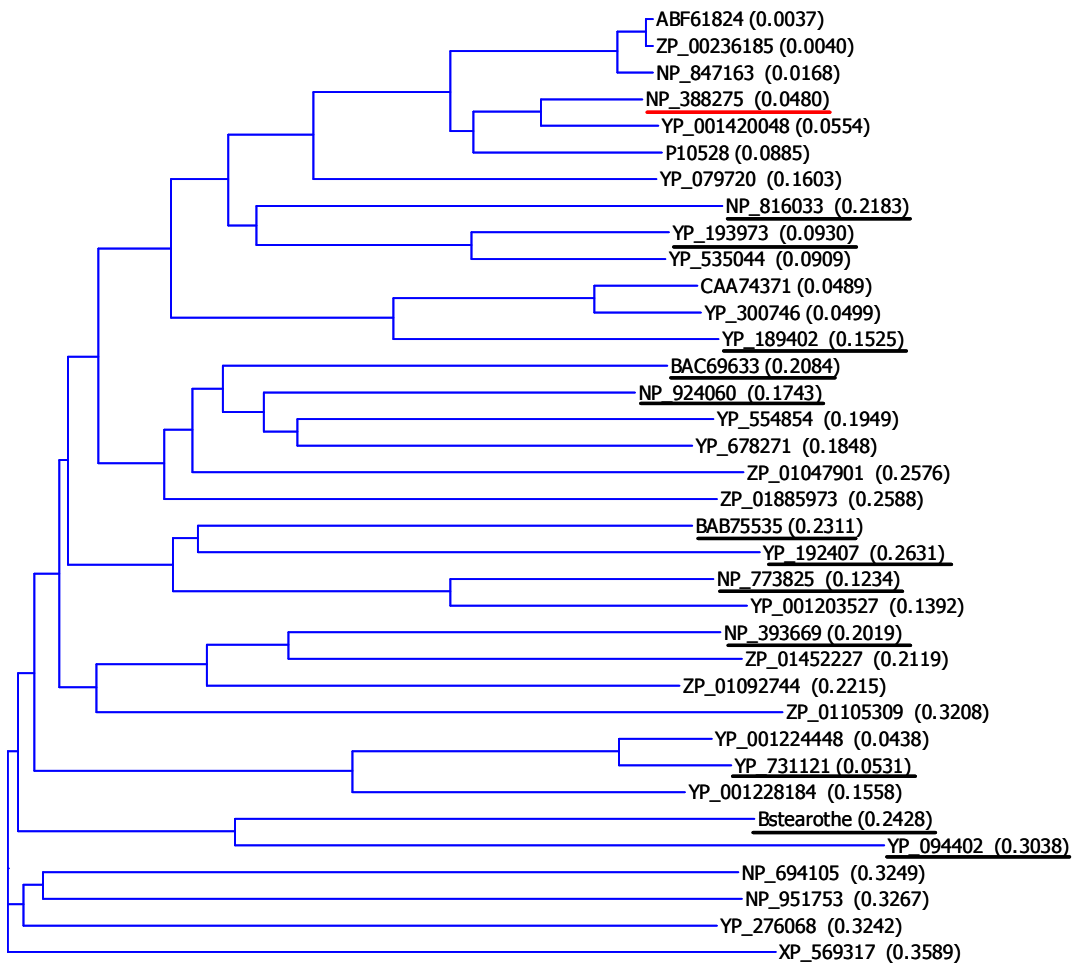
Figure A3.1. Representative phylogenetic tree constructed by the neighboring joining method (133). Underlined, sequences used for alignment (wild-type *Bacillus subtilis* strain 168 in red).

PRALINE alignment and consensus sequence for GDH

```
NP_388275B        ------------MYPDLKGKVVAITGAASGLGKAMAIRFGKEQAKVVINYYS--NKQDPN
YP_193973L        -----------MVYSELKNRVAVVTGGSKGIGTAISERFGKEGMKVVINYHS--DEKGAQ
NP_816033E        -----------MYGELHGKVAVVTGAATGLGLFITLRYILEGMNVVADYVG----ELPK
YP_189402S        -----------MFEELENKVVLITGAATGIGKSIAENFGKAKAKVVINYRSDRHHDEIE
YP_731121S        ----MTQVPHALASLCLQNKVIVVTGGNSGIGKAIVEVVGSLGAKVVIDYRS--HPERTE
NP_924060G        ---------MGYSDTLLKGQKAIVTGASAGIGEAIARHLAAAGAAVAVNYRS--DPDSAK
BAC69633St        METTGSDGRQPLAPQLLKGQKALVTGANSGIGKATAIGLGRVGADVVVNYVA--GRDAAE
NP_393669T        ----MDENRGSRIAINLTGKIALVTGSSSGLGAAIAKYMARAGAKVAVHYRS--GKDRAD
NP_773825B        MPQT---DRSIPFPTRLVGQYALVTGASQGIGRAVAVRLAQEGATVAINYFD--HPGKAE
BAB75535No        -------------MNGLKGKNTLITGASSGIGQAIAIRLAQEGCNIAINYRK--SPSGAE
YP_192407G        MPAPYKD--------RFAGKKVLVTGASQGIGEATALRFAEEGAQVALNGRK--EDKLIA
Bstearothe        ----------------MSGKVAVVTGSSRGIGKAIALRLAKEGYDIVVNYAR--NKTAAE
YP_094402L        -----------MSSLVFSGKVGLITGGARGIGKATALKLAQSGCDIAIVYYN--SSDEAQ
consensus/90%     ................htsphsllTGuspGlGtuhs.hhut.thplslsYht..t.p..p
consensus/80%     ................hpsphsllTGuupGlGpAhu.thuptGhplslsYhp..t.ptsp
consensus/70%     ................LpG+hsllTGuupGIGpAhAhphuptGspVslsYhs..p.ptsp
consensus/60%     ............h..pLpGKlulVTGAusGIGcAIAl+huppGAcVslNYtu..spctAc
consensus/50%     ............h.tpLcGKVALVTGASoGIG+AIAlRlAcEGAcVVINY+u..sc-pA-



NP_388275B        E--------VKEEVIKAGGEA-------VVVQGDVTKEEDVKNIVQTAIKEFGTLDIMIN
YP_193973L        E--------AADAVKKNGGDA-------VIVQADIGSEEGAQKLIDAAVNNFGGLDIWVN
NP_816033E        E--------FEDVQAKHADRV-------KFVKADVSNEEDIKALAETALKEFGHVDIWVN
YP_189402S        E--------IKQTVAKFGGQT-------LVVQGDVSIEEDIKRMIETTINHFGTLDIIIN
YP_731121S        E--------LIEEVGELGGQS-------IGVQADVSKLVDLQHLIDTAVQTFGKIDVMVN
NP_924060G        K--------IVEAIQAAGGEA-------VAIQADVSKEDAVKAMFAQALETFGTLDILVN
BAC69633St        E--------VVREIESFGVRA-------YAHEADVSQEDQVADMVSRMVKEFGTIDVMVA
NP_393669T        A--------IVDEIKNDGGFA-------MAFYGDVSKKEDVQKLFSEIDSKLGTVDILVN
NP_773825B        E--------TLALARTASGDRGHGRLDHVIVQADIGSEQDIAAMFETILARWKRLDCLVN
BAB75535No        ETEEMALQKACKNVEICGVKS-------LLVQGDVSQEEDVVEMVNTVIEEFGSLDILIN
YP_192407G        V--------REKLPKVSGGEH-------PIATGDISKEDDVKRLVAESIKAMGGLDVLVC
Bstearothe        E--------TAREIEALGRKA-------LVVKANVGDVEKIHAMFAQIDEVFGRVDVLVN
YP_094402L        S--------LVNEITGMGRKA-------VALQADVADCQSVKEMFSQFREHFKHLNFLVS
consensus/90%     t........h.p..t..uhpt.......hhhpuDlup.ptltthhtph.pthttlDhhls
consensus/80%     p........hhp.htt.Gsps.......hhhpuDlupppsltthhtphhppaGplDlhls
consensus/70%     c........hhctltthGscs.......hhlpuDVupc-slpthhpphlppFGplDlhVN
consensus/60%     E........hhcplpptGGcu.......lsVQADVSpE-Dlpchhspslccfgsld llvn
consensus/50%     E........lscplcssGGcA.......llVQADVScEEDV+cMlssslccFGsLDILVN



NP_388275B        NAGLE-NPVPSHEMPLKDWDKVIGTNLTGAFLGSREAIKYFVEN-------DIKGNVINM
YP_193973L        NAGME-NQVATKDMSLEDWNRVINVNLTGVFLGTKMALRYFTDH-------NKKGNIINM
NP_816033E        NAGVE-ASFPTIDMPLKEWQRVIDVNLNGVFLGSREALRIFRDQ-------KIKGSIINM
YP_189402S        NAGFE-NSIPTHEMSIDDWQKVIDINLTGAFVGSREAINQFLKE-------NKKGTIINI
YP_731121S        NAGIE-TRTSILDTSPDDFDKVMNVNLRGVFFATQYAAKQMISQ-------GSGGRIINI
NP_924060G        NAGRQ-NDAPFTEMTAEQWRSVIDVNLTGPFLCAQEAARLFLKQGVREGVSRAAGKIIFI
BAC69633St        NAGLQ-RDAPVIDMTMAQWQKVLDVNLTGQFLCAREAAKEFMRRGVVPEVSRSAGKIICM
NP_393669T        NAGIDGKREL VGEDDPDDWEKVIEVNLMGPYYCAREAVKRMKPK--------KSGVINI
NP_773825B        NAGFQ-RESPSDALDVETYRRIIDVNLNGAVLCAQKALAHFVAR-------GGGGSIINC
BAB75535No        NAGIQ-TECPSHEITAEDFDRVIGVNLRGSYLCARETIKHLLTQ-------NRSGVIINI
YP_192407G        NAGYQ-IPSPSEDIKLEDFEGVMAVNVTGVMLPCREVIRYWLEN-------GIKGTIIVN
Bstearothe        NAASG-VLRPAMELEETHWNWTMNINSKALLFCAQEAAKRMERV--------GGGKIVSI
YP_094402L        NAASG-VLKPALKMSTKHWRWCLETNALALNHLVTEGLALMPKN----------SRVIAL
consensus/90%     NAu.t....s..ch..ppaphshthNh.u.hhssp.sht.h.tp.........tG.lIsh
consensus/80%     NAGhp....ssh-hs.cpaptlhslNlpG.hhsspcuhp.hhpp.........tGpIIsh
consensus/70%     NAGhp.t.hPsh-hshccappVlslNLpGshhsspEAh+hhhpp.......thtGpIIsh
consensus/60%     NAGhp.spsPst-hsh--Wc+VIsVNLsGsaLsu+EAl+hahcp.......tttGsIINh
consensus/50%     NAGhp.spsPop-Moh-DWc+VI-VNLsGsFLsuREAl+pFlcp.......stuGsIINl
```

90

PRALINE alignment and consensus sequence (cont.)

```
NP_388275B      SSVHEVIPWPLFVHYAASKGGIKLMTETLALEYAPKGIRVNNIGPGAINTPINAEKFADP
YP_193973L      SSVHEQIPWPTFAHYAASKGGVKLFTETVAMEYAKQNIRVNAIGPGAINTPINAKKFADP
NP_816033E      SSVHQRIPWPTFAHYAASKGATEMFTKTIALEYAEYGIRANCIAPGAINTPINAEKFSDP
YP_189402S      SSVHDTIPWPNYVHYAASKGGLKLMMETMSMEYAQYGIRINNISPGAIVTEHTEEKFSDP
YP_731121S      SSVHEDWPMPNNTPYCVAKGGVRMLTRTAGVELAGKGVSIVNVGPGAVATPINDSTMNDP
NP_924060G      SSVHEVIPWAGRVNYAASKGGIEQLMKSIAQELAPSKIRVNSIAPGAIKTDINRESWEKP
BAC69633St      SSVHQIIPWAGHVNYAASKGGVLMMMATLAQELAPHRIRVNAVAPGAIRTPINRSAWDTP
NP_393669T      TSVHEYVPWSGYTAYSSAKAGLSMFTKALAQELSDYNIRVVAIAPGAIKTPINKDVWGNP
NP_773825B      SSVHQIIPKPGYLAYSISKGGLANLTRTLALEFARHGIRVNAVGPGAIDTPINAAWTGDP
BAB75535No      SSVHEIIPRPMYVSYSISKGGMENMTKTLALEYAHRGIRVNSVAPGATITPINEAWTDDP
YP_192407G      SSVHQIIPKPHYLGYSASKGAVGNIVRTLALEYATRGIRVNAVAPGAIVTPINMSWIDDP
Bstearothe      SSLGSIRYLENYTAVGVSKAALEALTRYLAVELAPKNIAVNAVSGGAVDTDALKHFPNRE
YP_094402L      SSLGASRAIPNYAFIGASKAALEALVRSLSLELASYGITVNTVSAGVVDTDALKHFPNRE
consensus/90%   SSltp.hshs.hh.huhuKuuht.hhcshu.EhA..tIplsslusGAl.Tsh.tth.sp.
consensus/80%   SSVHp.hPhstastYusSKuult.hhcolu.EhA.hsIplNsluPGAl.Tshstphhscp
consensus/70%   SSVHphlPhssastYusSKGulp.hscolAhEhA.hsIRVNsluPGAltTsINtphhscP
consensus/60%   SSVHphIPhPsassYuASKGGlchhT+TLAlEhAspGIRVNuluPGAIsTPINcchhsDP
consensus/50%   SSVH-lIPWPsYlsYuASKGGlchhT+TLAlEaAs+GIRVNuVuPGAIsTPINcctasDP


NP_388275B      KQKADVESMIPMGYIGEPEEIAAVAAWLASKEASYVTGITLFADGGMTQ-YPSFQAGRG-
YP_193973L      EQKATTTSMIPMGRIGDPEEVAAAAAWLASDESSYVTGITMFVDGGMTL-YPSFQGGRG-
NP_816033E      EQLKQTTSMVPMGIIGKPNQVAAAAAWLASDESSYVTGTTLFVDGGMSL-YPSFQHGAG-
YP_189402S      TTREETIKMIPAREIGNAQDVANAVLFLSSDLASYIHGTTLYVDGGMMN-YPAFMGGKG-
YP_731121S      ELLAKLNAAIPIGRMAQPAEIASVVAFLAGNGASYMTATSVFADGGIMMSSPGL------
NP_924060G      EAEAELLKRIPAGRVGESDDIGKAAVWLASDESDYVNGTTLFIDGGMTL-YPEFADGG--
BAC69633St      EAEADLLRLIPYRRVGDPDDIANAVAALASDLFDYVVGTTIYVDGGMTL-FPGFATGG--
NP_393669T      ESLKDLLNKIAMPRLGEVDDIGQAAVFLASDLASYITGTTLLVDGGMAL-YPDFKHGG--
NP_773825B      EKRGVVTSHIPLGRVGTPEEIAAVFAFLASDDASYITGQTIYACGGLTM-FPEFRENWAS
BAB75535No      EKKAVVESHIPMRRAGTSEEMAAAVAFLASDEAAYITGQTLFVDGGLSL-YADFREPWSA
YP_192407G      EQYKAVSSHIPMKRPGESREIADAITFLAAEDSTYITGQTLYVDGGLTL-YGDFENNWSS
Bstearothe      ELLADAAANTPAGRPVKPDDIVNAVLFLLSDAAEMIRGQTIIVDGGRSLLL---------
YP_094402L      QLLDEYQAHALADRALTPEDVANAIYLLCLPEAAMINAHTLYVDAGYSQVG---------
consensus/90%   p..t.h.thhshth.spsp-lutshhhLsup.ushlpu.TlhsDGGh...h.........
consensus/80%   c.htth.thlPhtchupsp-lusshhaLuuc.usYlsGpTlasDGGhs..hsth......
consensus/70%   Ethtph.shIPhtRhGpsc-lAsAssaLAS-tusYlsGpTlalDGGhoh.astF.tsh..
consensus/60%   EphuclputIPhsRlGcP--IAsAssaLASD-AoYlTGpTLaVDGGhoL.YPsFttsh..
consensus/50%   EpcA-lpupIPhGRlGcP--IAsAlAFLASD-ASYITGpTLaVDGGMoL.YPsFppGtu.
```

91

Table A3.1. Legend of symbols for consensus alingment[a]		
Group	**Symbol**	**Members**
Alcohol	o	S, T
Aliphatic	l	I, L, V
Any	.	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
Aromatic	a	F, H, W, Y
Charged	c	D, E, H, K, R
Hydrophobic	h	A, C, F, G, H, I, K, L, M, R, T, V, W, Y
Negative	-	D, E
Polar	p	C, D, E, H, K, N, Q, R, S, T
Positive	+	H, K, R
Small	s	A, C, D, G, N, P, S, T, V
Tiny	u	A, G, S
Turnlike	t	A, C, D, E, G, H, K, N, Q, R, S, T
[a] http://coot.embl.de/Alignment/consensus.html		
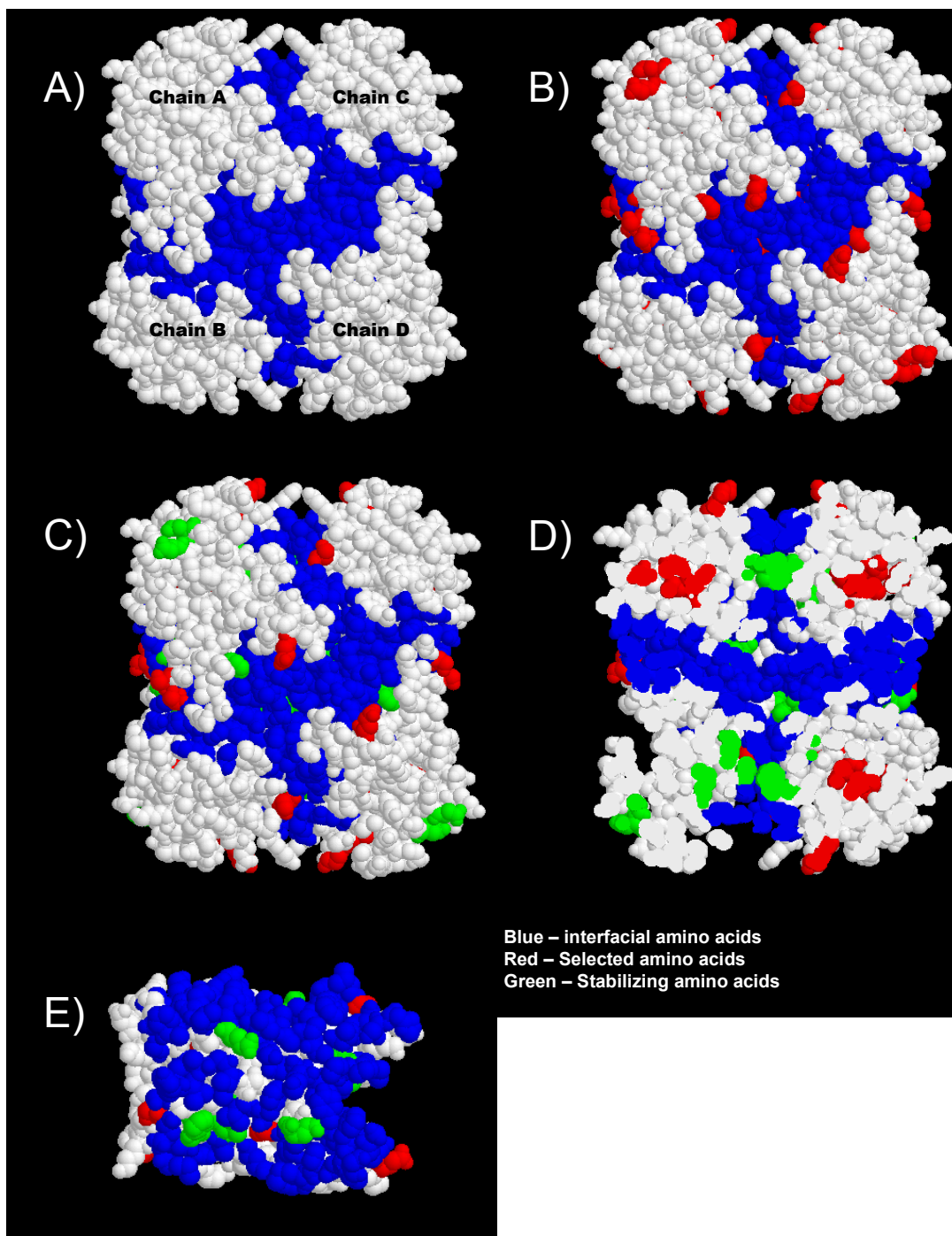
Figure A3.2. Amino acids selected for mutations in GDH. A) GDH structure (white) with interfacial amino acids (blue). B) GDH structure with amino acids selected for substitutions (red). C) GDH structure with stabilizing mutations (green). D) Slab of GDH structure showing buried, surface and interfacial amino acids mutated. E) GDH monomer interface with stabilizing amino acids (green).
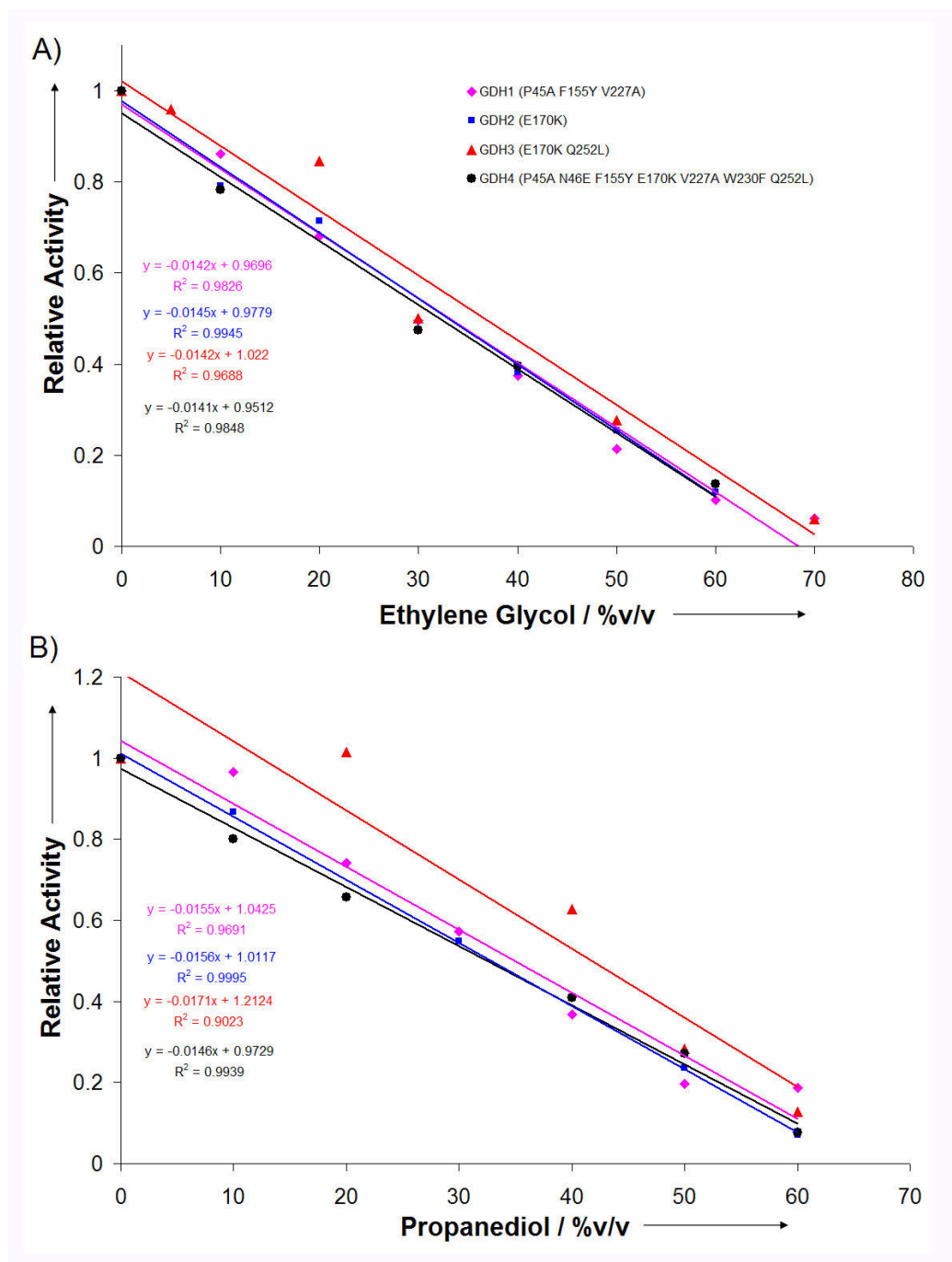
Figure A4.1. Representative linear behavior observed for initial activity experiments in the presence of organic solvent. A) Ethylene glycol. B) 1,2-Propanediol.
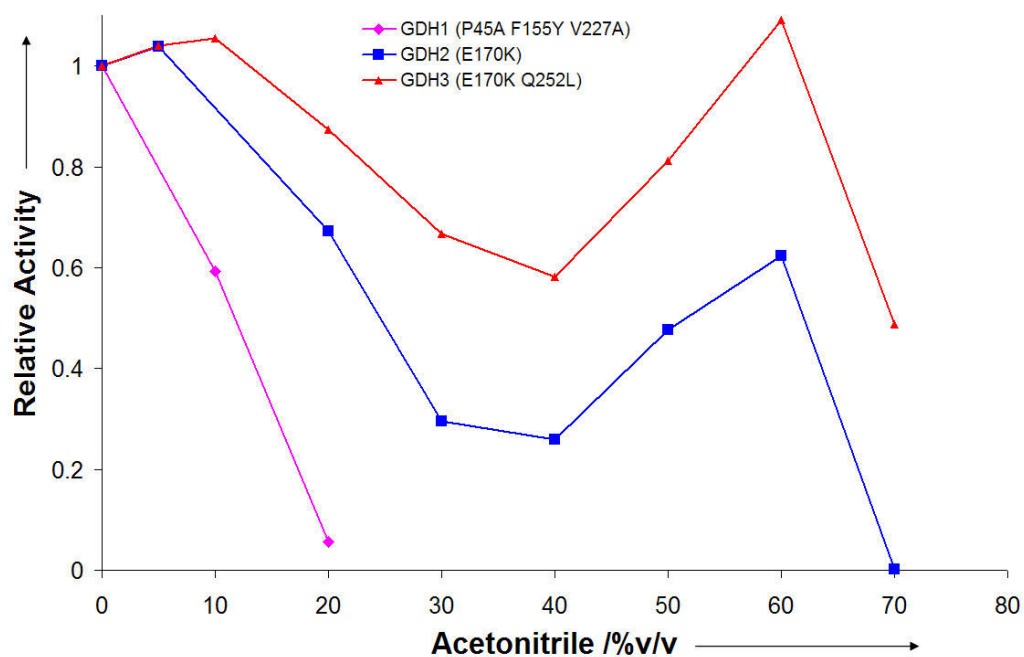
Figure A4.2. Behavior observed for initial activity experiments in the presence of acetonitrile.

PRALINE alignment and consensus sequence for AADHs

```
BAD11020       --------------------------------------------------M-E--------
BAC81829       --------------------------------------------------M-E--------
BAA00524       --------------------------------------------------MRD--------
Q59224         ----------------------------------------MSLVEKTSIIKD--------
BAA19221       -------------------------------------MILVTLEQTLQDDKASV--------
P23307         ---------------------------------------MAKQLEKSSKIGN--------
1BW9_A         --------------------------------------------------MS--------
NP_391659      -----------------------MSAKQVSKDEEKEALNLFLSTQTIIKEALRKLGYP
AAA87979       -----------MDQTYSLESFLNHVQKRDPNQTEFAQAVREVMTTLWPFLEQNPKY-RQ
YP_092525      MHTLEKMEQTNFQTARDYVTQAYETVQKRNFYESEFLQAVKEIFDSLVPVLARHPKY-IE
NP_242488      -------------------MEDKDKRSGKTTDVLASTQTVIKKALDKLG------YP
P17556         --------------------------------------------------MK
CAB04775       --------------------------------------------------MI
P17557         --------------------------------------------------MK
BAA16790       --------------------------------------------------ME
consensus/90%  ............................................................
consensus/80%  ............................................................
consensus/70%  ............................................................
consensus/60%  .....................................................h.t.......
consensus/50%  .................................................hhp......h.

BAD11020       IFKYMEKYDYEQLVFC-------QDEASGLKAVIAIHDTTLGPALGGARMWTYNAEEEAI
BAC81829       LFKYMETYDYEQVLFC-------QDKESGLKAIIAIHDTTLGPALGGTRMWMYNSEEEAL
BAA00524       VFEMMDRYGHEQVIFC-------RHPQTGLKAIIALHNTTAGPALGGCRMIPYASTDEAL
Q59224         FTLFEKMSEHEQVVFC-------NDPATGLRAIIAIHDTTLGPALGGCRMQPYNSVEEAL
BAA19221       LDKMV---EHEQILFC-------HDKATGLQAIIAVHDTTMGPALGGCRMAPYKTMDLAL
P23307         EDVFQKIANHEQIVFC-------NDPVSGLQAIIAIHDTTLGPALGGTRMYPYKNVDEAL
1BW9_A         IDSALNW-DGEMTVTR-------FDSMTGAHFVIRLDSTQLGPAAGGTRAAQYSNLADAL
NP_391659      GDMYELMKEPQRMLTVRIPVKMDNGSVKVFTGYRSQHNDAVGPTKGGVRFHP----EVNE
AAA87979       MSLLERLVEPERVIQFRVVWVDDRNQIQVNRAWRVQFSSAIGPYKGGMRFHP----SVNL
YP_092525      HRILERIAEPERMITFRVPWVDDEGNIRVNRGFRVQFNSAIGPYKGGIRFHP----SVNA
NP_242488      DEMYELMKEPIRMLTVRIPVRMDDGSTKIFTGYRAQHNDAVGPTKGGVRFHP----NVTE
P17556         IGIPKEIKNNENRVAM--------TPAGVVSLTHAGHERLAIETGGGI--------GSSF
CAB04775       IGVPKEIKNNENRVAL--------TPGGVSQLISNGHR-VLVETGAGL--------GSGF
P17557         IGIPKEIKNNENRVAI--------TPAGVMTLVKAGHE-VYVETEGGA--------GSGF
BAA16790       IGVPKEIKDQEFRVGL--------TPSSVRALLSQGHQ-VFVEEGAGV--------GSGF
consensus/90%  ......h.p.p.hlhh.........s..ts.phh...ap.hhs.hhuGh........t.s.
consensus/80%  hth..phhp.Ephlhh.........s.hpshphhhs.Hppshs.shGGh........t.sh
consensus/70%  hthh.chhp.Ephlhh........sshsshphhhshHssshGPshGGsRh......s.uh
consensus/60%  lshhccht-tEphlth.......psssssshpulhAhHsoslGPstGGsRhhs....ssuh
consensus/50%  lslhccht-sEphVss.......psssolhpulpAtHcoslGPstGGsRhhP....ssul

BAD11020       EDALRLARGMTYKNAAAGLNLGGGKTVIIGDPFADKN-----EDMFRALGRFIQGL----
BAC81829       EDALRLARGMTYKNAAAGLNLGGGKTVIIGDPRKDKN-----EAMFRAFGRFIQGL----
BAA00524       EDVLRLSKGMTYKCSLADVDFGGGKMVIIGDPKKDKS-----PELFRVIGRFVGGL----
Q59224         EDALRLSKGMTYKCAASDVDFGGGKAVIIGDPQKDKS-----PELFRAFGQFVDSL----
BAA19221       KDVLRLSKGMTYKCAAADVDFGGGKSVIIGDPLKDKT-----PEKFRAFGQFIESL----
P23307         EDVLRLSEGMTYKCAAADIDFGGGKAVIIGDPEKDKS-----PALFRAFGQFVESL----
1BW9_A         TDAGKLAGAMTLKMAVSNLPMGGGKSVIALPAPRHSIDPSTWARILRIHAENIDKL----
NP_391659      EEVKALSIWMTLKCGIANLPYGGGKGGIICDPRTMSF-----GELERLSRGYVRAISQIV
AAA87979       SILKFLGFEQTFKNALTTLPMGGGKGGSDFDPKGKSE-----GEVMRFCQALMTELYRHL
YP_092525      SIIKFLGFEQIFKNSLTGLPIGGGKGGADFDPKGKSD-----REIMSFTQSFMNELYRHI
NP_242488      KEVKALSIWMSLKAGIVDLPYGGGKGGIVCDPRDMSF-----RELERLSRGYVRAISQIV
P17556         TDAEYVAAGAAYRCIGKEAW-AQEMILKVKEPVASEY------DYFYEGQILFTYL--HL|
CAB04775       ENEAYESAGAEIIADPKQVW-DAEMVMKVKEPLPEEY------VYFRKGLVLFTYL--HL
P17557         SDSEYEKAGAADRCRTWRDAWTAEMVLKVKEPLAREF------RYFRPGLILFTYL--HL
BAA16790       PDGAYAKAGAELVATAKEAW-NRELVVKVKEPLPEEY------EYLTLPKLLFTYL--HL
consensus/90%  p...h.th..thhsth.th..stth.hh.h-P...p........th.p....hhp.l....
consensus/80%  p-shhluhthsh+sshtplshuuthssshlt-Phtcp.......chhRhht.hhp.L....
consensus/70%  hthh.chhp.Ephlhh........sshsshphhhshHssshGPshGGsRh......s.uh
consensus/60%  cDsttLutGMTaKsussslshGGGKslIltDPhscc.......-hFRhhttalstL....
consensus/50%  cDsttLutGMTaKsusuclshGGGKuVIlsDPps-cp......ElFRhstpalsuL..hl
```

PRALINE alignment and consensus sequence for AADHs (cont.)

```
BAD11020      -NGRYITAEDVGTTVDDMDL--IHQET---------DYVTGI-SPAFGSSGNPSPVTAYG
BAC81829      -NGRYITAEDVGTTVADMDI--IYQET---------DYVTGI-SPEFGSSGNPSPATAYG
BAA00524      -NGRFYTGTDMGTNPEDFVH--AARES---------KSFAGL-PKSYGGKGDTSIPTALG
Q59224        -GGRFYTGTDMGTNMEDFIH--AMKET---------NCIVGV-PEAYGGGGDSSIPTAMG
BAA19221      -NGRFYTGTDMGTTLEDFVH--AMKET---------NYIVGK-PVEYGGGGDSSIPTALG
P23307        -NGRFYTGTDMGTTMDDFVH--AQKET---------NFINGI-PEQYGGSGDSSIPTAQG
1BW9_A        -SGNYWTGPDVNTNSADMDT--LNDTT---------EFVFGR-SLERGGAGSSAFTTAVG
NP_391659     GPTKDIPAPDVYTNSQIMAW--MMDEYSRLREFDSPGFITGK-PLVLGGSQGRETATAQG
AAA87979      GADTDVPAGDIGVGGREVGF--MAGMMKKLSNNT-ACVFTGK-GLSFGGSLIRPEATGYG
YP_092525     GPDTDIPAGDIGVGAREVGF--MFGQYKKIRGRYDAGVLTGK-GLEYGGSLTRKEATGYG
NP_242488     GPTKDIPAPDVFTNSQIMAW--MLDEYSRIREFDSPGFITGK-PIVLGGSHGRESATAKG
P17556        APRAELTQALIDKKVVGIAYETVQLANGSLPLLTPMSEVAGKMATQIGAQYLEKNHGGKG
CAB04775      AAEPELAQALKDKGVTAIAYETVS-EGRTLPLLTPMSEVAGRMAAQIGAQFLEKPKGGKG
P17557        AAAERVTKAVVEQKVVGIAYETVQLANGSLPLLTPMSEVAGRMSVQVGAQFLEKPHGGKG
BAA16790      AAERTLTEALIKSGITAIAYETVELADGQLPLLAPMSRIAGRLAVQMGAHYLEKQQGGRG
consensus/90% .stt.hst..h..t...hsh..h..t..........t.hsGh.s..hGuth..t..su.G
consensus/80% .ssp.hstsshtss.tthsa..h..t..........s.lsGh.s.thGuph.pp..suhG
consensus/70% .ssp.hsusDhssssspshsa..h.tps.........shlsGh.shphGuph.pp.sTuhG
consensus/60% .ss+hlTusDlsTssp-hsa..htpEs.........shlsG+.uhphGGutspp.sTAhG
consensus/50% usu+plTusDlGTsss-hua..stpEstpl......shlsG+.ulpaGGutscs.sTAhG

BAD11020      VYRGMKAAAKEAFGSDSLEGLAVSVQGLGNVAKALCKKLNTEGAKLVVTDVNKAAVSAAV
BAC81829      VYRGMKAAAKEAFGSDSLEGKVVAVQGVGNVAYHLCRHLHEEGAKLIVTDINKEAVARAV
BAA00524      VFHGMRATARFLWGTDQLKGRVVAIQGVGKVGERLLQLLVEVGAYCKIADIDSVRCEQLK
Q59224        VLYGIKATNKMLFGKDDLGGVTYAIQGLGKVGYKVAEGLLEEGAHLFVTDINEQTLEAIQ
BAA19221      VFYGIKATNQNLFGDDKVEGRKYSIQGLGKVGYKVAEHIINEGGNVIVTDINEQAIADIQ
P23307        VIYALKATNQYLFGSDSLSGKTYAIQGLGKVGYKVAEQLLKAGADLFVTDIHENVLNSIK
1BW9_A        VFEAMKATVAHR-GLGSLDGLTVLVQGLGAVGGSLASLAAEAGAQLLVADTDTERVAHAV
NP_391659     VTICIEEAVKKK--GIKLQNARIIIQGFGNAGSFLAKFMHDAGAKVIGISDANGGLYNPD
AAA87979      LVYFTEAMLKRH--GMGFEGMRVSVSGSGNVAQYAIEKAMEFGARVITASDSSGTVVDES
YP_092525     LVYFVEEMLKDQ--GMRFENSTVVVSGSGNVALYAMEKAAQFGAKVVACSDSDGYVYDEK
NP_242488     VTICIREAAKKK--GIDLQDARVVIQGFGNAGSFLAKFMHDAGAKVIGISDAYGALHDPD
P17556        ILLG---------GVSGVHARKVTVIGGGIAGTNAAKIAVGMGADVTVIDLSPERLRQLE
CAB04775      ILLA---------GVPGVSRGKVTIIGGGVVGTNAAKMAVGLGADVTIIDLNADRLRQLD
P17557        ILLG---------GVPGVRRGKVTIIGGGTAGTNAAKIGVGLGADVTILDINAERLRELD
BAA16790      VLLG---------GVPGVKAGQVTILGGGVVGTEAAKMAIGLGAMVTILDINVDRLNQLG
consensus/90% lhhh...........thpthhh.l.G.G.su..hhphhht.GA.lhhhs.s.thl.t..
consensus/80% lhhs...........s.tlpshpls l.GhGtsu..hschhht.GAplhhhs.stttltp.t
consensus/70% lhhuhcth.t...Gssslpstp VslpGhGpVGhthschhhphGAclhlhDhstttltphp
consensus/60% Vhhuhctssp.h.GssulcutpVsIQGhGsVGpplAchhhphGAcvhlsDlssstlpphs
consensus/50% VlhGh+Ass+ph.GssuLcGtpVsIQGsGsVGpplA+hhlctGAcVllsDlssssLpslc

BAD11020      A--------------EE---G--ADAVAP-------NAIYGVTCDIFAPCALGAVLNDF
BAC81829      E--------------EF---G--AKAVDP-------NDIYGVECDIFAPCALGGIINDQ
BAA00524      E--------------KY---GDKVQLVDV-------NRIHKESCDIFSPCAKGGVVNDD
Q59224        E--------------KAKTTSGSVTVVAS-------DEIYSQEADVFVPCAFGGVVNDE
BAA19221      K--------------LG---GSAVRVVSS-------EEIYSQQADVFVPCAFGGVINDD
P23307        Q--------------KSEELGGSVTIVKS-------DDIYSVQADIFVPCAMGGIINDK
1BW9_A        A--------------L------GHTAVAL------EDVLSTPCDVFAPCAMGGVITTE
NP_391659     GLDIPY---LLDKRDSFGMVTNLFTDVITN-------EELLEKDCDILVPAAISNQITAK
AAA87979      GFTKEKLARLIEIKASRDGRVADYAKEF-GLVYLEG-QQPWSLPVDIALPCATQNELDVD
YP_092525     GICLET---VKRLKEDGNGRIREYVSEHPEAHYFEGCTGIWSIPCDIALPCATQNEIDEE
NP_242488     GLDIDY---LLDRRDSFGTVTKLFKNTISN-------EELLELDCDILVPAAIENQITEK
P17556        DMF-----------GRDVQTLMSNPYNIAESV-----KHSDLVVGAVLIPGAKAPKLVSE
CAB04775      DIF-----------GHQIKTLISNPVNIADAV-----AEADLLICAVLIPGAKAPTLVTE
P17557        DLF-----------GDHVTTLMSNSYHIAECV-----RESDLVVGAVLIPGAKAK-LVTE
BAA16790      ELF-----------GSRVDLRYSNASQIEDLL-----PHTDLLIGAVLITGKRAPVLVSR
consensus/90% t...............................h.........tt.....sslhhPsAhts.lstp
consensus/80% t...............h..sp.ls........pch..h.sslhhPsAhus.lssc
consensus/70% t..............t...h..sp.lss.......pchhth.sDlhlPsAhus.lssc
consensus/60% s................csthhtstssslss.......p-lhplssDlhlPCAhushlss-
consensus/50% ch...........tpcsthhssssslus.......c-laulsCDIhlPCAhuslls--
```

97

PRALINE alignment and consensus sequence for AADHs (cont.)

```
BAD11020       TIPQL---KAKVIAGSADNQLKDPRHGKYLHELGIVYAPDYVINAGGVINVADEL-----
BAC81829       TIPQL---KAKVIAGSANNQLKEPRHGDIIHEMGIVYAPDYVINAGGVINVADEL-----
BAA00524       TIDEF---RCLAIVGSANNQLVEDRHGALLQKRSICYAPDYLVNAGGLIQVADEL-----
Q59224         TMKQF---KVKAIAGSANNQLLTEDHGRHLADKGILYAPDYIVNSGGLIQVADEL-----
BAA19221       TLKVL---KVRGISGSANNQLAESRHGELLREKGILYAPDYIVNGGGLIQVADEL-----
P23307         TIPKL---KVKAVVGSANNQLKDLRHANVLNEKGILYAPDYIVNAGGLIQVADEL-----
1BW9_A         VARTL---DCSVVAGAANNVIADEAASDILHARGILYAPDFVANAGGAIHLVGREV----
NP_391659      NAHNI---QASIVVERANGP-TTIDATKILNERGVLLVPDILASAGGVTVSYFEWVQNNQ
AAA87979       AAHQLIANGVKAVAEGANMP-TTIEATELFQQAGVLFAPGKAANAGGVATSGLEMAQNAA
YP_092525      AAEVLISNGVKAVGEGANMP-SEEGAVKRFLDAGVLFGPAKAANAGGVAVSALEMAQNSA
NP_242488      NAHNI---KAQIVVEAANGP-TTIEATEILTNRDILLVPDVLASAGGVTVSYFEWVQNNQ
P17556         EMIQSMQPGSVVVDIAIDQGGIFATSDRVTTHDDPTYVKHGVVHY-AVANMPGAVPRTST
CAB04775       EMVKQMKPGSVIVDVAIDQGGIVETVDHITTHDQPTYEKHGVVHY-AVANMPGAVPRTST
P17557         EMVRSMTPGSVLVDIAIDQGGIFETTDRVTTHDDPTYVKHGVVHY-AVANMPGAVPRTST
BAA16790       QEVEQMLPGAVIMDVAIDQGGCVETL-RVTSHSQPSYIEAEVVHV-GIPNMPGATPWTAT
consensus/90%  th......ts.hls.uhs...h..thsphhtp.t.hhh.thhhph.ulh....th.....
consensus/80%  ph.p....tshhls.uhst..h..ptsclhpcts.has.chlsps.Glhp.s.th.....
consensus/70%  sh.ph...tshhlstuAsp.hhp.ptsclhpcpslhYsPchlssuGGlhphssch.....
consensus/60%  shhpl...psplVssuANsshtstctscllpccGllYuPDhllNAGGVhsssssEls....
consensus/50%  shcpl...cuplVssuANsshsspcss+lLsccGILYAPDhlVNAGGVsssusEls.sst

BAD11020       -YGYNRTRAMKRVDGIYDSI-EKIFAISKR--DGVPSYVAADRMAEERIAKVAKARSQFL
BAC81829       -YGYNRERAMKKIEQIYDNI-EKVFAIAKR--DNIPTYVAADRMAEERIETMRKARSQFL
BAA00524       -EGFHEERVLAKTEAIYDMV-LDIFHRAKN--ENITTCEAADRIVMERLKKLTDIRRILL
Q59224         -YEVNKERVLAKTKHIYDAI-LEVYQQAEL--DQITTMEAANRMCEQRMAARGRRNSFFT
BAA19221       -YGTNPARVLAKTENIYTSL-LEVFHQAEQ--DHMTTATAADRMCEKRIADAKNRNSFFT
P23307         -YGPNKERVLLKTKEIYRSL-LEIFNQAAL--DCITTVEAANRKCQKTIEGQQTRNSFFS
1BW9_A         -LGWSESVVHERAVAIGDTL-NQVFEISDN--DGVTPDEAARTLAGRRAREASTTTATA-
NP_391659      GYYWSEEEVAEKLRSVMVSSFETIYQTAAT--HKVDMRLAAYMTGIRKSAEASRFRGWV-
AAA87979       RLGWKAEKVDARLHHIMLDIHHACVEHGGE-GEQTNYVQGANIAGFVKVADAMLAQGVI-
YP_092525      RLHWTAEETDAKLRAIMADIHKRSVEAASEYGRPGNLLDGCNIAGFIKVADAMIAQGVV-
NP_242488      GYYWTEEEVEGKLEAVMVHAFDNIYKLAKG--RKVDMRLAAYMVGVRKMAEASRFRGWV-
P17556         -IALTNNTIPYALQ-IANKGYKQACIDNPALKKGVNALEGH--ITYKAVAEAQGLPYVNV
CAB04775       -IALTNVTVPYALQ-IANKGAVKALADNTALRAGLNTANGH--VTYEAVARDLGYEYVPA
P17557         -FALTNVTIPYALQ-IANKGYRAGCLDNPALLKGINTLDGH--IVYEAVAAAHNMPYTDV
BAA16790       -QALNNSTLRYVLK-LANLG-EQAWENDLPLAKGVNVQAGK--LVQGAVKTV----FPDL
consensus/90%  ..hhpt.ph..thp.lh..h..thh..s.....cths...ut..hs..thtt...h.....
consensus/80%  .hthspppl.hthp.Ihsph..phht.stt..ctlshh.ut..hs.pthtph.thp.h..
consensus/70%  .huhspppl.h+hctIhsph..phhttutt..ctlsshtus.hhs.cpltpstthpthh.
consensus/60%  .huhsppcVht+lctIhssl.cplappupt..cslsshpAAphhshc+lAcspphpuhhh
consensus/50%  .auascEcVhtKLctIhssl.cclappAss..-slsohpAAshhshc+lAcApptpuhsh

BAD11020       QDQRNILNGR---
BAC81829       QNGHHILSRRRAR
BAA00524       EDPRNSARR----
Q59224         SSVKPKWDIRN--
BAA19221       QSNRPKWNFHQ--
P23307         RGRRPKWNIKE--
1BW9_A         -------------
NP_391659      -------------
AAA87979       -------------
YP_092525      -------------
NP_242488      -------------
P17556         DELIQ--------
CAB04775       EKALQDESSVAGA
P17557         HSLL------HG-
BAA16790       -------------
consensus/90%  .............
consensus/80%  .............
consensus/70%  .............
consensus/60%  pt.h.........
consensus/50%  pp.h.........
```
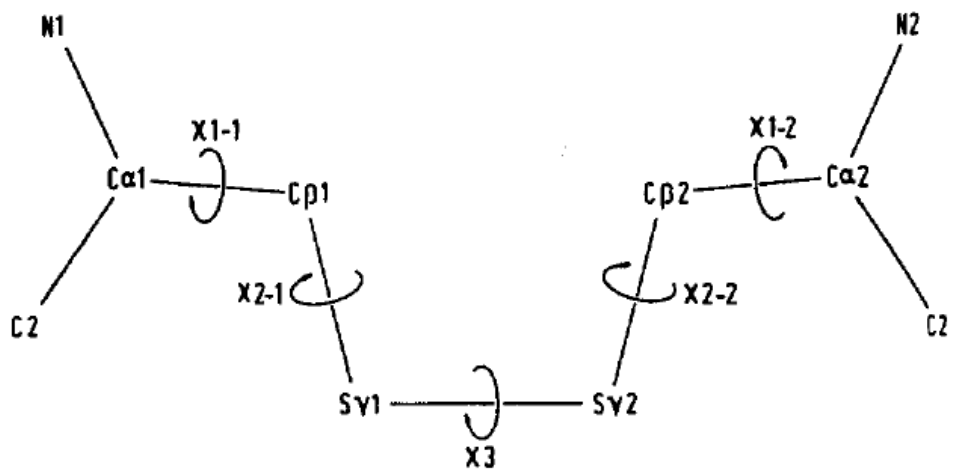
Figure A5.1. Nomenclature of the dihedral angles in a disulfide bond. From Hazes 1988 (55).

REFERENCES

1. **Alexandre, F. R., D. P. Pantaleone, P. P. Taylor, I. G. Fotheringham, D. J. Ager, and N. J. Turner.** 2002. Amine-boranes: effective reducing agents for the deracemisation of DL-amino acids using L-amino acid oxidase from Proteus myxofaciens. Tetrahedron Letters **43:**707-710.

2. **Alexeeva, M., R. Carr, and N. J. Turner.** 2003. Directed evolution of enzymes: new biocatalysts for asymmetric synthesis. Organic & Biomolecular Chemistry **1:**4133-4137.

3. **Amin, N., A. D. Liu, S. Ramer, W. Aehle, D. Meijer, M. Metin, S. Wong, P. Gualfetti, and V. Schellenberger.** 2004. Construction of stabilized proteins by combinatorial consensus mutagenesis. Protein Engineering, Design & Selection **17:**787-793.

4. **Arnold, F. H., and G. Georgiou.** 2003. Directed Enzyme Evolution: Screening and Selection Methods 1ed. Humana Press, Totowa, New Jersey.

5. **Arnold, F. H., and G. Georgiou.** 2003. Directed evolution library creation : methods and protocols. Humana Press, Totowa, N.J.

6. **Baik, S. H., T. Ide, H. Yoshida, O. Kagami, and S. Harayama.** 2003. Significantly enhanced stability of glucose dehydrogenase by directed evolution. Applied Microbiology and Biotechnology **61:**329-335.

7. **Baik, S. H., F. Michel, N. Aghajari, R. Haser, and S. Harayama.** 2005. Cooperative effect of two surface amino acid mutations (Q252L and E170K) in glucose dehydrogenase from Bacillus megaterium IWG3 on stabilization of its oligomeric state. Applied and Environmental Microbiology **71:**3285-3293.

8. **Baker, P. J., A. P. Turnbull, S. E. Sedelnikova, T. J. Stillman, and D. W. Rice.** 1995. A Role for Quaternary Structure in the Substrate-Specificity of Leucine Dehydrogenase. Structure **3:**693-705.

9. **Baker, P. J., M. L. Waugh, X. G. Wang, T. J. Stillman, A. P. Turnbull, P. C. Engel, and D. W. Rice.** 1997. Determinants of substrate specificity in the superfamily of amino acid dehydrogenases. Biochemistry **36:**16109-16115.

10. **Balkenhohl, F., K. Ditrich, B. Hauer, and W. Ladner.** 1997. Optically active amines via lipase-catalyzed methoxyacetylation. Journal Fur Praktische Chemie-Chemiker-Zeitung **339:**381-384.

11. **Besenmatter, W., P. Kast, and D. Hilvert.** 2007. Relative tolerance of mesostable and thermostable protein homologs to extensive mutation. Proteins: Structure, Function, and Bioinformatics **66:**500-506.

12. **Binz, H. K., M. T. Stumpp, P. Forrer, P. Amstutz, and A. Pluckthun.** 2003. Designing repeat proteins: Well-expressed, soluble and stable

proteins from combinatorial libraries of consensus ankyrin repeat proteins. Journal of Molecular Biology **332:**489-503.

13. **Bloom, J. D., S. T. Labthavikul, C. R. Otey, and F. H. Arnold.** 2006. Protein stability promotes evolvability. Proceedings of the National Academy of Sciences of the United States of America **103:**5869-5874.

14. **Bommarius, A. S., J. M. Broering, J. F. Chaparro-Riggers, and K. M. Polizzi.** 2006. High-throughput screening for enhanced protein stability. Current Opinion in Biotechnology **17:**606-610.

15. **Bommarius, A. S., and K. M. Polizzi.** 2006. Novel biocatalysts: Recent developments. Chemical Engineering Science **61:**1004-1016.

16. **Bommarius, A. S., M. Schwarm, and K. Drauz.** 1998. Biocatalysis to amino acid-based chiral pharmaceuticals - examples and perspectives. Journal of Molecular Catalysis B-Enzymatic **5:**1-11.

17. **Bornscheuer, U. T.** 2005. Trends and challenges in enzyme technology, p. 181-203, Biotechnology for the Future, vol. 100.

18. **Britton, K. L., P. J. Baker, P. C. Engel, D. W. Rice, and T. J. Stillman.** 1993. Evolution of Substrate Diversity in the Superfamily of Amino-Acid Dehydrogenases - Prospects for Rational Chiral Synthesis. Journal of Molecular Biology **234:**938-945.

19. **Broering, J. M., and A. S. Bommarius.** 2005. Evaluation of Hofmeister effects on the kinetic stability of proteins. J Phys Chem B **109:**20612-9.

20. **Brunhuber, N. M., J. B. Thoden, J. S. Blanchard, and J. L. Vanhooke.** 2000. Rhodococcus L-phenylalanine dehydrogenase: kinetics, mechanism, and structural basis for catalytic specificity. Biochemistry **39:**9174-87.

21. **Brunhuber, N. M. W., and J. S. Blanchard.** 1994. The Biochemistry and Enzymology of Amino-Acid Dehydrogenases. Critical Reviews in Biochemistry and Molecular Biology **29:**415-467.

22. **Burton, S. G., D. A. Cowan, and J. M. Woodley.** 2002. The search for the ideal biocatalyst. Nat Biotechnol **20:**37-45.

23. **Cadwell, R. C., and G. F. Joyce.** 1992. Randomization of genes by PCR mutagenesis. PCR Methods Appl **2:**28-33.

24. **Carr, R., M. Alexeeva, M. J. Dawson, V. Gotor-Fernandez, C. E. Humphrey, and N. J. Turner.** 2005. Directed evolution of an amine oxidase for the preparative deracemisation of cyclic secondary amines. Chembiochem **6:**637-639.

25. **Carr, R., M. Alexeeva, A. Enright, T. S. C. Eve, M. J. Dawson, and N. J. Turner.** 2003. Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. Angewandte Chemie-International Edition **42:**4807-4810.

26. **Chaparro-Riggers, J. F., K. M. Polizzi, and A. S. Bommarius.** 2007. Better library design: data-driven protein engineering. Biotechnology journal **2:**180-191.

27. **Chaparro-Riggers, J. F., T. A. Rogers, E. Vázquez-Figueroa, and A. S. Bommarius.** 2007. Comparison of Three Enoate Reductases and their Potential Use for Biotransformations Advanced Synthesis & Catalysis **349:**1521-1531.

28. **Cirilli, M., G. Scapin, A. Sutherland, J. C. Vederas, and J. S. Blanchard.** 2000. The three-dimensional structure of the ternary complex of Corynebacterium glutamicum diaminopimelate dehydrogenase-NADPH-L-2-amino-6-methylene-pimelate Protein Science **9:**2034-2037.

29. **Constable, D. J. C., P. J. Dunn, J. D. Hayler, G. R. Humphrey, J. L. Leazer, R. J. Linderman, K. Lorenz, J. Manley, B. A. Pearlman, A. Wells, A. Zaks, and T. Y. Zhang.** 2007. Key green chemistry research areas - a perspective from pharmaceutical manufacturers. Green Chemistry **9:**411-420.

30. **Cooper, A., S. J. Eyles, S. E. Radford, and C. M. Dobson.** 1992. Thermodynamic consequences of the removal of a disulphide bridge from hen lysozyme. J Mol Biol **225:**939-43.

31. **Cowan, D. A.** 1997. Thermophilic proteins: Stability and function in aqueous and organic solvents. Comparative Biochemistry and Physiology a-Physiology **118:**429-438.

32. **Crameri, A., S. A. Raillard, E. Bermudez, and W. P. Stemmer.** 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature **391:**288-91.

33. **Davis-Searles, P. R., A. J. Saunders, D. A. Erie, D. J. Winzor, and G. J. Pielak.** 2001. Interpreting the effects of small uncharged solutes on protein-folding equilibria. Annual Review of Biophysics and Biomolecular Structure **30:**271-306.

34. **De Wildeman, S. M. A., T. Sonke, H. E. Schoemaker, and O. May.** 2007. Biocatalytic reductions: From lab curiosity to "first choice". Accounts of Chemical Research **40:**1260-1266.

35. **DiTursi, M. K., S. J. Kwon, P. J. Reeder, and J. S. Dordick.** 2006. Bioinformatics-driven, rational engineering of protein thermostability. Protein Engineering, Design & Selection **19:**517-524.

36. **Dordick, J. S.** 1989. Enzymatic Catalysis in Monophasic Organic-Solvents. Enzyme and Microbial Technology **11:**194-211.

37. **Eijsink, V. G. H., A. Bjork, S. Gaseidnes, R. Sirevag, B. Synstad, B. van den Burg, and G. Vriend.** 2004. Rational engineering of enzyme stability. Journal of Biotechnology **113:**105-120.

38. **Eijsink, V. G. H., S. Gaseidnes, T. V. Borchert, and B. van den Burg.** 2005. Directed evolution of enzyme stability. Biomolecular Engineering **22:**21-30.

39. **Faber, K.** 2004. Biotransformations in organic chemistry : a textbook, 5th rev. & corr. ed. Springer-Verlag, Berlin ; New York.

40. **Faber, K.** 2001. Non-sequential processes for the transformation of a racemate into a single stereoisomeric product: Proposal for stereochemical classification. Chemistry-a European Journal **7:**5004-+.

41. **Filikov, A. V., R. J. Hayes, P. Z. Luo, D. M. Stark, C. Chan, A. Kundu, and B. I. Dahiyat.** 2002. Computational stabilization of human growth hormone. Protein Science **11:**1452-1461.

42. **Flory, P. J.** 1956. Theory of Elastic Mechanisms in Fibrous Proteins. Journal of the American Chemical Society **78:**5222-5234.

43. **Fox, R. J., S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J. C. Whitman, R. A. Sheldon, and G. W. Huisman.** 2007. Improving catalytic function by ProSAR-driven enzyme evolution. Nature Biotechnology **25:**338-344.

44. **Gaseidnes, S., B. Synstad, X. H. Jia, H. Kjellesvik, G. Vriend, and V. G. H. Eijsink.** 2003. Stabilization of a chitinase from Serratia marcescens by Gly -> Ala and Xxx -> Pro mutations. Protein Engineering **16:**841-846.

45. **Ghatorae, A. S., M. J. Guerra, G. Bell, and P. J. Halling.** 1994. Immiscible Organic-Solvent Inactivation of Urease, Chymotrypsin, Lipase, and Ribonuclease - Separation of Dissolved Solvent and Interfacial Effects. Biotechnology and Bioengineering **44:**1355-1361.

46. **Ghose, A. K., V. N. Viswanadhan, and J. J. Wendoloski.** 1999. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem **1:**55-68.

47. **Gomes, J., and W. Steiner.** 2004. The biocatalytic potential of extremophiles and extremozymes. Food Technology and Biotechnology **42:**223-235.

48. **Griebenow, K., and A. M. Klibanov.** 1996. On protein denaturation in aqueous-organic mixtures but not in pure organic solvents. Journal of the American Chemical Society **118:**11695-11700.

49. **Griebenow, K., M. Vidal, C. Baez, A. M. Santos, and G. Barletta.** 2001. Nativelike enzyme properties are important for optimum activity in neat organic solvents. Journal of the American Chemical Society **123:**5380-5381.

50. **Gunasekaran, K., C. Ramakrishnan, and P. Balaram.** 1996. Disallowed Ramachandran conformations of amino acid residues in protein structures. Journal of Molecular Biology **264:**191-198.

51. **Gupta, M. N., R. Batra, R. Tyagi, and A. Sharma.** 1997. Polarity index: The guiding solvent parameter for enzyme stability in aqueous-organic cosolvent mixtures. Biotechnology Progress **13:**284-288.

52. **Haki, G. D., and S. K. Rakshit.** 2003. Developments in industrially important thermostable enzymes: a review. Bioresource Technology **89:**17-34.

53. **Hall, M., C. Stueckler, W. Kroutil, P. Macheroux, and K. Faber.** 2007. Asymmetric bioreduction of activated alkenes using cloned 12-oxophytodienoate reductase isoenzymes OPR-1 and OPR-3 from Lycopersicon esculentum (Tomato): A striking change of stereoselectivity. Angewandte Chemie-International Edition **46:**3934-3937.

54. **Hao, J. J., and A. Berry.** 2004. A thermostable variant of fructose bisphosphate aldolase constructed by directed evolution also shows increased stability in organic solvents. Protein Engineering Design & Selection **17:**689-697.

55. **Hazes, B., and B. W. Dijkstra.** 1988. Model building of disulfide bonds in proteins with known three-dimensional structure. Protein Eng **2:**119-25.

56. **Hendriks, A. T., and G. Zeeman.** 2008. Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour Technol.

57. **Henke, E., and U. T. Bornscheuer.** 2003. Fluorophoric assay for the high-throughput determination of amidase activity. Analytical Chemistry **75:**255-260.

58. **Henkel, T., R. M. Brunne, H. Muller, and F. Reichel.** 1999. Statistical investigation into the structural complementarity of natural products and synthetic compounds. Angew Chem Int Ed Engl **38:**643-647.

59. **Heringa, J.** 1999. Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. Computers & Chemistry **23:**341-364.

60. **Higuchi, R., B. Krummel, and R. K. Saiki.** 1988. A General-Method of Invitro Preparation and Specific Mutagenesis of DNA Fragments - Study of Protein and DNA Interactions. Nucleic Acids Research **16:**7351-7367.

61. **Hilt, W., G. Pfleiderer, and P. Fortnagel.** 1991. Glucose-Dehydrogenase from Bacillus-Subtilis Expressed in Escherichia-Coli .1. Purification, Characterization and Comparison with Glucose-Dehydrogenase from Bacillus-Megaterium. Biochimica Et Biophysica Acta **1076:**298-304.

62. **Hine, J. S.** 1962. Physical organic chemistry, 2d ed. McGraw-Hill, New York,.

63. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989. Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain-Reaction. Gene **77:**51-59.

64. **Huff, J. W.** 1947. The fluorescent condensation product of N1-methylnicotinamide and acetone .1. Synthesis and properties. Journal of Biological Chemistry **167:**151-156.

65. **Itoh, N., C. Yachi, and T. Kudome.** 2000. Determining a novel NAD(+)-dependent amine dehydrogenase with a broad substrate range from Streptomyces virginiae IFO 12827: purification and characterization. Journal of Molecular Catalysis B Enzymatic **10:**281-290.

66. **Jacobsen, S. E., and C. E. Wyman.** 2000. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes. Appl Biochem Biotechnol **84-86:**81-96.

67. **Jaenicke, R., and P. Zavodszky.** 1990. Proteins under extreme physical conditions. FEBS Lett **268:**344-9.

68. **Johannes, T. W., R. D. Woodyer, and H. M. Zhao.** 2005. Directed evolution of a thermostable phosphite dehydrogenase for NAD(P)H regeneration. Applied and Environmental Microbiology **71:**5728-5734.

69. **Johnson, S. L., and P. T. Tuazon.** 1977. Acid-catalyzed hydration of reduced nicotinamide adenine dinucleotide and its analogues. Biochemistry **16:**1175-83.

70. **Jornvall, H., B. Persson, M. Krook, S. Atrian, R. Gonzalezduarte, J. Jeffery, and D. Ghosh.** 1995. Short-Chain Dehydrogenases Reductases (Sdr). Biochemistry **34:**6003-6013.

71. **Kanaya, S., C. Katsuda, S. Kimura, T. Nakai, E. Kitakuni, H. Nakamura, K. Katayanagi, K. Morikawa, and M. Ikehara.** 1991. Stabilization of Escherichia coli ribonuclease H by introduction of an artificial disulfide bond. J Biol Chem **266:**6038-44.

72. **Kandimalla, V. B., V. S. Tripathi, and H. X. Ju.** 2006. Immobilization of biomolecules in sol-gels: Biological and analytical applications. Critical Reviews in Analytical Chemistry **36:**73-106.

73. **Kaplan, N. O., S. P. Colowick, and C. C. Barnes.** 1951. Effect of alkali on diphosphopyridine nucleotide. J Biol Chem **191:**461-72.

74. **Kazlauskas, R. J.** 2005. Enhancing catalytic promiscuity for biocatalysis. Current Opinion in Chemical Biology **9:**195-201.

75. **Khmelnitsky, Y. L., V. V. Mozhaev, A. B. Belova, M. V. Sergeeva, and K. Martinek.** 1991. Denaturation Capacity - a New Quantitative Criterion for Selection of Organic-Solvents as Reaction Media in Biocatalysis. European Journal of Biochemistry **198:**31-41.

76. **Khmelnitsky, Y. L., S. H. Welch, D. S. Clark, and J. S. Dordick.** 1994. Salts Dramatically Enhance Activity of Enzymes Suspended in Organic-Solvents. Journal of the American Chemical Society **116:**2647-2648.

77. **Kim, J., D. S. Clark, and J. S. Dordick.** 2000. Intrinsic effects of solvent polarity on enzymic activation energies. Biotechnology and Bioengineering **67:**112-116.

78. **Klibanov, A. M., A. N. Semenov, G. P. Samokhin, and K. Martinek.** 1978. Enzymatic-Reactions in Water-Organic Mixtures - Criterion for Choosing an Optimal Organic-Solvent. Bioorganicheskaya Khimiya **4:**82-87.

79. **Korkegian, A., M. E. Black, D. Baker, and B. L. Stoddard.** 2005. Computational thermostabilization of an enzyme. Science **308:**857-860.

80. **Korkhin, Y., A. J. Kalb, M. Peretz, O. Bogin, Y. Burstein, and E. Frolow.** 1999. Oligomeric integrity - The structural key to thermal stability in bacterial alcohol dehydrogenases. Protein Science **8:**1241-1249.

81. **Kumar, R., S. Singh, and O. V. Singh.** 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J Ind Microbiol Biotechnol **35:**377-91.

82. **Kumar, S., and R. Nussinov.** 2001. How do thermophilic proteins deal with heat? Cellular and Molecular Life Sciences **58:**1216-1233.

83. **Laane, C.** 1987. Medium-engineering for Bio-organic Synthesis. Biocatalysis **1:**17-22.

84. **Laane, C., S. Boeren, K. Vos, and C. Veeger.** 1986. Rules for Optimization of Biocatalysis in Organic Solvents. Biotechnology and Bioengineering **30:**81-87.

85. **Lehmann, M., C. Loch, A. Middendorf, D. Studer, S. F. Lassen, L. Pasamontes, A. van Loon, and M. Wyss.** 2002. The consensus concept for thermostability engineering of proteins: further proof of concept. Protein Engineering **15:**403-411.

86. **Lehmann, M., L. Pasamontes, S. F. Lassen, and M. Wyss.** 2000. The consensus concept for thermostability engineering of proteins. Biochimica Et Biophysica Acta, Protein Structure and Molecular Enzymology **1543:**408-415.

87. **Lehmann, M., and M. Wyss.** 2001. Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. Current Opinion in Biotechnology **12:**371-375.

88. **Li, Y., D. A. Drummond, A. M. Sawayama, C. D. Snow, J. D. Bloom, and F. H. Arnold.** 2007. A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments. Nat Biotechnol **25:**1051-6.

89. **Lowry, O. H., and J. V. Passonneau.** 1972. A flexible system of enzymatic analysis. Academic Press, New York,.

90. **Lowry, O. H., J. V. Passonneau, and M. K. Rock.** 1961. The stability of pyridine nucleotides. J Biol Chem **236:**2756-9.

91. **Lowry, O. H., N. R. Roberts, and J. I. Kapphahn.** 1957. The fluorometric measurement of pyridine nucleotides. J Biol Chem **224:**1047-64.

92. **Lynd, L. R., P. J. Weimer, W. H. van Zyl, and I. S. Pretorius.** 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev **66:**506-77, table of contents.

93. **Makino, Y., S. Negoro, I. Urabe, and H. Okada.** 1989. Stability-Increasing Mutants of Glucose-Dehydrogenase from Bacillus-Megaterium Iwg3. Journal of Biological Chemistry **264:**6381-6385.

94. **Malkov, S., M. V. Zivkovic, M. V. Beljanski, and S. D. Zaric.** 2005. Correlations of Amino Acids with Secondary Structure Types:

Connection with Amino Acid Structure. Los Alamos National Laboratory, Preprint Archive, Quantitative Biology 1-13.

95. **Mansfeld, J., and R. Ulbrich-Hofmann.** 2007. The stability of engineered thermostable neutral proteases from Bacillus stearothermophilus in organic solvents and detergents. Biotechnol Bioeng **97:**672-9.

96. **Mansfeld, J., G. Vriend, B. W. Dijkstra, O. R. Veltman, B. VandenBurg, G. Venema, R. UlbrichHofmann, and V. G. H. Eijsink.** 1997. Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. Journal of Biological Chemistry **272:**11152-11156.

97. **Martinek, K., A. M. Klibanov, G. P. Samokhin, A. N. Semenov, and I. V. Berezin.** 1977. Preparative Enzymatic-Synthesis in Biphasic Aqueous-Organic System. Bioorganicheskaya Khimiya **3:**696-702.

98. **Martinek, K., A. N. Semenov, and I. V. Berezin.** 1980. Enzyme-Catalyzed Peptide-Synthesis in Biphase Water-Organic System. Doklady Akademii Nauk Sssr **254:**121-123.

99. **Matsumura, M., G. Signor, and B. W. Matthews.** 1989. Substantial increase of protein stability by multiple disulphide bonds. Nature **342:**291-3.

100. **Matthews, B. W., H. Nicholson, and W. J. Becktel.** 1987. Enhanced Protein Thermostability from Site-Directed Mutations That Decrease the Entropy of Unfolding. Proceedings of the National Academy of Sciences of the United States of America **84:**6663-6667.

101. **Mehling, A., U. F. Wehmeier, and W. Piepersberg.** 1995. Application of Random Amplified Polymorphic DNA (Rapd) Assays in Identifying Conserved Regions of Actinomycete Genomes. Fems Microbiology Letters **128:**119-125.

102. **Mihovilovic, M. D.** 2006. Enzyme mediated Baeyer-Villiger oxidations. Current Organic Chemistry **10:**1265-1287.

103. **Miyano, H., T. Toyooka, and K. Imai.** 1985. Further-Studies on the Reaction of Amines and Proteins with 4-fluoro-7-nitrobenzo-2-oxa-1,3-diazole. Analytica Chimica Acta **170:**81-87.

104. **Moore, J. C., D. J. Pollard, B. Kosjek, and P. N. Devine.** 2007. Advances in the enzymatic reduction of ketones. Accounts of Chemical Research **40:**1412-1419.

105. **Morley, K. L., and R. J. Kazlauskas.** 2005. Improving enzyme properties: when are closer mutations better? TRENDS in Biotechnology **23:**231-237.

106. **Mozhaev, V. V., Y. L. Khmelnitsky, M. V. Sergeeva, A. B. Belova, N. L. Klyachko, A. V. Levashov, and K. Martinek.** 1989. Catalytic Activity and Denaturation of Enzymes in Water Organic Cosolvent Mixtures - Alpha-Chymotrypsin and Laccase in Mixed Water Alcohol, Water Glycol and Water Formamide Solvents. European Journal of Biochemistry **184:**597-602.

107. **Nagao, T., Y. Makino, K. Yamamoto, I. Urabe, and H. Okada.** 1989. Stability-Increasing Mutants of Glucose-Dehydrogenase. Febs Letters **253:**113-116.

108. **Nagata, S., K. Tanizawa, N. Esaki, Y. Sakamoto, T. Ohshima, H. Tanaka, and K. Soda.** 1988. Gene Cloning and Sequence Determination of Leucine Dehydrogenase from Bacillus-Stearothermophilus and Structural Comparison with Other Nad(P)+-Dependent Dehydrogenases. Biochemistry **27:**9056-9062.

109. **Ohshima, T., S. Nagata, and K. Soda.** 1985. Purification and Characterization of Thermostable Leucine Dehydrogenase from Bacillus-Stearothermophilus. Archives of Microbiology **141:**407-411.

110. **Oka, M., Y. S. Yang, S. Nagata, N. Esaki, H. Tanaka, and K. Soda.** 1989. Overproduction of Thermostable Leucine Dehydrogenase of Bacillus-Stearothermophilus and Its One-Step Purification from Recombinant Cells of Escherichia-Coli. Biotechnology and Applied Biochemistry **11:**307-311.

111. **Omidinia, E., A. Samadi, H. Taherkhani, S. Khatami, N. Moazami, R. R. Pouraie, and Y. Asano.** 2002. Cloning and expression of Bacillus sphaericus phenylalanine dehydrogenase gene in Bacillus subtilis cells: purification and enzyme properties. World Journal of Microbiology & Biotechnology **18:**593-597.

112. **Otten, L. G., and W. J. Quax.** 2005. Directed evolution: selecting today's biocatalysts. Biomolecular Engineering **22:**1-9.

113. **Owusu, R. K., and D. A. Cowan.** 1989. Correlation between microbial protein thermostability and resistance to denaturation in aqueous:organic solvent two-phase systems. Enzyme Microbiology and Technology **11:**568-574.

114. **Pace, C. N., G. R. Grimsley, J. A. Thomson, and B. J. Barnett.** 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J Biol Chem **263:**11820-5.

115. **Panke, S., M. Held, and M. Wubbolts.** 2004. Trends and innovations in industrial biocatalysis for the production of fine chemicals. Current Opinion in Biotechnology **15:**272-279.

116. **Pauly, H. E., and G. Pfleiderer.** 1975. D-Glucose Dehydrogenase from Bacillus-Megaterium M-1286 - Purification, Properties and Structure. Hoppe-Seylers Zeitschrift Fur Physiologische Chemie **356:**1613-1623.

117. **Percival Zhang, Y. H., M. E. Himmel, and J. R. Mielenz.** 2006. Outlook for cellulase improvement: screening and selection strategies. Biotechnol Adv **24:**452-81.

118. **Podar, M., and A. L. Reysenbach.** 2006. New opportunities revealed by biotechnological explorations of extremophiles. Current Opinion in Biotechnology **17:**250-255.

119. **Polizzi, K. M., A. S. Bommarius, J. M. Broering, and J. F. Chaparro-Riggers.** 2007. Stability of biocatalysts. Current Opinion in Chemical Biology **11:**220-225.

120. **Polizzi, K. M., J. F. Chaparro-Riggers, E. Vázquez-Figueroa, and A. S. Bommarius.** 2006. Structure-guided consensus approach to create a more thermostable penicillin G acylase. Biotechnology Journal **1:**531-536.

121. **Pollard, D., M. Truppo, J. Pollard, C. Y. Chen, and J. Moore.** 2006. Effective synthesis of (S)-3,5-bistrifluoromethylphenyl ethanol by asymmetric enzymatic reduction. Tetrahedron-Asymmetry **17:**554-559.

122. **Pollard, D. J., and J. M. Woodley.** 2007. Biocatalysis for pharmaceutical intermediates: the future is now. Trends in Biotechnology **25:**66-73.

123. **Porollo, A., and J. Meller.** 2007. Prediction-based fingerprints of protein-protein interactions. Proteins **66:**630-45.

124. **Porollo, A., and J. Meller.** 2007. Solvent accessibility based Protein-Protein Interface iDEntification and Recognition

125. **Putt, K. S., and P. J. Hergenrother.** 2004. An enzymatic assay for poly(ADP-ribose) polymerase-1 (PARP-1) via the chemical quantitation of NAD(+): application to the high-throughput screening of small molecules as potential inhibitors. Anal Biochem **326:**78-86.

126. **Rabinovich, M. L., M. S. Melnick, and A. V. Bolobova.** 2002. The structure and mechanism of action of cellulolytic enzymes. Biochemistry (Mosc) **67:**850-71.

127. **Ramaley, R. F., and N. Vasantha.** 1983. Glycerol Protection and Purification of Bacillus-Subtilis Glucose-Dehydrogenase. Journal of Biological Chemistry **258:**2558-2565.

128. **Reetz, M. T.** 2004. Controlling the enantioselectivity of enzymes by directed evolution: Practical and theoretical ramifications. Proceedings of the National Academy of Sciences of the United States of America **101:**5716-5722.

129. **Reetz, M. T., M. Bocola, J. D. Carballeira, D. X. Zha, and A. Vogel.** 2005. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. Angewandte Chemie, International Edition **44:**4192-4196.

130. **Reetz, M. T., and J. D. Carballeira.** 2007. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. Nat Protoc **2:**891-903.

131. **Reetz, M. T., J. D. Carballeira, J. Peyralans, H. Hobenreich, A. Maichele, and A. Vogel.** 2006. Expanding the substrate scope of enzymes: Combining mutations obtained by CASTing. Chemistry-a European Journal **12:**6031-6038.

132. **Reetz, M. T., J. D Carballeira, and A. Vogel.** 2006. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angewandte Chemie, International Edition **45:**7745-7751.

133. **Saitou, N., and M. Nei.** 1987. The Neighbor-Joining Method - a New Method for Reconstructing Phylogenetic Trees. Molecular Biology and Evolution **4:**406-425.

134. **Sanwal, B. D., and M. W. Zink.** 1961. L-Leucine Dehydrogenase of Bacillus Cereus. Archives of Biochemistry and Biophysics **94:**430-&.

135. **Scapin, G., M. Cirilli, S. G. Reddy, Y. Gao, J. C. Vederas, and J. S. Blanchard.** 1998. Substrate and inhibitor binding sites in Corynebacterium glutamicum diaminopimelate dehydrogenase. Biochemistry **37:**3278-3285.

136. **Schmid, A., J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, and B. Witholt.** 2001. Industrial biocatalysis today and tomorrow. Nature **409:**258-268.

137. **Schulein, M.** 2000. Protein engineering of cellulases. Biochim Biophys Acta **1543:**239-252.

138. **Sekimoto, T., T. Fukui, and K. Tanizawa.** 1994. Involvement of Conserved Lysine-68 of Bacillus-Stearothermophilus Leucine Dehydrogenase in Substrate-Binding. Journal of Biological Chemistry **269:**7262-7266.

139. **Sekimoto, T., T. Fukui, and K. Tanizawa.** 1994. Role of the Conserved Glycyl Residues Located at the Active-Site of Leucine Dehydrogenase from Bacillus-Stearothermophilus. Journal of Biochemistry **116:**176-182.

140. **Sekimoto, T., T. Matsuyama, T. Fukui, and K. Tanizawa.** 1993. Evidence for Lysine-80 as General Base Catalyst of Leucine Dehydrogenase. Journal of Biological Chemistry **268:**27039-27045.

141. **Semenov, A. N., and K. Martinek.** 1980. Enzymatic-Synthesis in Biphasic Aqueous-Organic Systems .3. The Shift of Ionic Equilibria. Bioorganicheskaya Khimiya **6:**1559-1571.

142. **Semenov, A. N., K. Martinek, and I. V. Berezin.** 1980. Enzymatic-Synthesis in Biphasic Water-Organic Systems .2. The Chemical-Equilibrium Shift. Bioorganicheskaya Khimiya **6:**600-608.

143. **Serdakowski, A. L., and J. S. Dordick.** 2008. Enzyme activation for organic solvents made easy. Trends in Biotechnology **26:**48-54.

144. **Serov, A. E., E. R. Odintzeva, I. V. Uporov, and V. I. Tishkov.** 2005. Use of ramachandran plot for increasing thermal stability of bacterial formate dehydrogenase. Biochemistry-Moscow **70:**804-808.

145. **Serrano, L.** 2000. The relationship between sequence and structure in elementary folding units, p. 49-85, Advances in Protein Chemistry, Vol 53.

146. **Shin, J. S., and B. G. Kim.** 1999. Asymmetric synthesis of chiral amines with omega-transaminase. Biotechnology and Bioengineering **65:**206-211.

147. **Shin, J. S., and B. G. Kim.** 2001. Comparison of the omega-transaminases from different microorganisms and application to

production of chiral amines. Bioscience Biotechnology and Biochemistry **65:**1782-1788.

148. **Shin, J. S., and B. G. Kim.** 2002. Exploring the active site of amine : pyruvate aminotransferase on the basis of the substrate structure-reactivity relationship: How the enzyme controls substrate specificity and stereo selectivity. Journal of Organic Chemistry **67:**2848-2853.

149. **Shin, J. S., and B. G. Kim.** 1998. Kinetic modeling of omega-transamination for enzymatic kinetic resolution of alpha-methylbenzylamine. Biotechnology and Bioengineering **60:**534-540.

150. **Shin, J. S., and B. G. Kim.** 1999. Modeling of the kinetic resolution of alpha-methylbenzylamine with omega-transaminase in a two-liquid-phase system. Enzyme and Microbial Technology **25:**426-432.

151. **Simossis, V. A., and J. Heringa.** 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. Nucleic Acids Research **33:**W289-W294.

152. **Simossis, V. A., J. Kleinjung, and J. Heringa.** 2005. Homology-extended sequence alignment. Nucleic Acids Research **33:**816-824.

153. **Simpson, E. R., J. K. Meldrum, R. Bofill, M. D. Crespo, E. Holmes, and M. S. Searle.** 2005. Engineering enhanced protein stability through beta-turn optimization: Insights for the design of stable peptide beta-hairpin systems. Angewandte Chemie, International Edition **44:**4939-4944.

154. **Smallwood, I. M.** 1996. Handbook of organic solvent properties. Arnold;Halsted Press, London;New York.

155. **Steipe, B.** 2004. Consensus-based engineering of protein stability: From intrabodies to thermostable enzymes, p. 176-186, Protein Engineering, vol. 388.

156. **Steipe, B., B. Schiller, A. Pluckthun, and S. Steinbacher.** 1994. Sequence Statistics Reliably Predict Stabilizing Mutations in a Protein Domain. Journal of Molecular Biology **240:**188-192.

157. **Strickler, S. S., A. V. Gribenko, T. R. Keiffer, J. Tomlinson, T. Reihle, V. V. Loladze, and G. I. Makhatadze.** 2006. Protein stability and surface electrostatics: A charged relationship. Biochemistry **45:**2761-2766.

158. **Takagi, H., T. Takahashi, H. Momose, M. Inouye, Y. Maeda, H. Matsuzawa, and T. Ohta.** 1990. Enhancement of the thermostability of subtilisin E by introduction of a disulfide bond engineered on the basis of structural comparison with a thermophilic serine protease. J Biol Chem **265:**6874-8.

159. **Tararov, V. I., and A. Borner.** 2005. Approaching highly enantioselective reductive amination. Synlett**:**203-211.

160. **Tararov, V. I., R. Kadyrov, T. H. Riermeier, C. Fischer, and A. Borner.** 2004. Direct Reductive Amination *versus* Hydrogenation of

Intermediates - A Comparison. Advanced Synthesis & Catalysis **346:**561-565.

161. **Thornton, J. M.** 1981. Disulfide Bridges in Globular Proteins. Journal of Molecular Biology **151:**261-287.

162. **Tishkov, V. I., and V. O. Popov.** 2006. Protein engineering of formate dehydrogenase. Biomolecular Engineering **23:**89-110.

163. **Tsotsou, G. E., A. E. Cass, and G. Gilardi.** 2002. High throughput assay for cytochrome P450 BM3 for screening libraries of substrates and combinatorial mutants. Biosens Bioelectron **17:**119-31.

164. **Tuazon, P. T., and S. L. Johnson.** 1977. Free radical and ionic reaction of bisulfite with reduced nicotinamide adenine dinucleotide and its analogues. Biochemistry **16:**1183-8.

165. **Valivety, R. H., P. J. Halling, and A. R. Macrae.** 1992. Reaction-Rate with Suspended Lipase Catalyst Shows Similar Dependence on Water Activity in Different Organic-Solvents. Biochimica Et Biophysica Acta **1118:**218-222.

166. **Valivety, R. H., P. J. Halling, A. D. Peilow, and A. R. Macrae.** 1994. Relationship between Water Activity and Catalytic Activity of Lipases in Organic Media - Effects of Supports, Loading and Enzyme Preparation. European Journal of Biochemistry **222:**461-466.

167. **Vanhooke, J. L., J. B. Thoden, N. M. Brunhuber, J. S. Blanchard, and H. M. Holden.** 1999. Phenylalanine dehydrogenase from Rhodococcus sp. M4: high-resolution X-ray analyses of inhibitory ternary complexes reveal key features in the oxidative deamination mechanism. Biochemistry **38:**2326-39.

168. **Vazquez-Figueroa, E., J. Chaparro-Riggers, and A. S. Bommarius.** 2007. Development of a thermostable glucose dehydrogenase by a structure-guided consensus concept. Chembiochem **8:**2295-301.

169. **Vedha-Peters, K., M. Gunawardana, J. D. Rozzell, and S. J. Novick.** 2006. Creation of a broad-range and highly stereoselective D-amino acid dehydrogenase for the one-step synthesis of D-amino acids. J Am Chem Soc **128:**10923-9.

170. **Wakarchuk, W. W., W. L. Sung, R. L. Campbell, A. Cunningham, D. C. Watson, and M. Yaguchi.** 1994. Thermostabilization of the Bacillus circulans xylanase by the introduction of disulfide bonds. Protein Eng **7:**1379-86.

171. **Watanabe, K., T. Ohkuri, S. Yokobori, and A. Yamagishi.** 2006. Designing thermostable proteins: Ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. Journal of Molecular Biology **355:**664-674.

172. **Weil, J., P. Westgate, K. Kohlmann, and M. R. Ladisch.** 1994. Cellulose pretreatments of lignocellulosic substrates. Enzyme Microb Technol **16:**1002-4.

173. **Wetzel, R.** 1987. Harnessing Disulfide Bonds Using Protein Engineering. Trends in Biochemical Sciences **12:**478-482.

174. **Wichmann, R., and D. Vasic-Racki.** 2005. Cofactor regeneration at the lab scale, p. 225-260, Technology Transfer in Biotechnology: From Lab to Industry to Production, vol. 92.

175. **Wolfenden, R., and M. J. Snider.** 2001. The depth of chemical time and the power of enzymes as catalysts. Acc Chem Res **34:**938-45.

176. **Wong, C. H., D. G. Drueckhammer, and H. M. Sweers.** 1985. Enzymatic Vs Fermentative Synthesis - Thermostable Glucose-Dehydrogenase Catalyzed Regeneration of Nad(P)H for Use in Enzymatic-Synthesis. Journal of the American Chemical Society **107:**4028-4031.

177. **Wong, C. H., and G. M. Whitesides.** 1981. ENZYME-CATALYZED ORGANIC-SYNTHESIS - NAD(P)H COFACTOR REGENERATION BY USING GLUCOSE-6-PHOSPHATE AND THE GLUCOSE-6-PHOSPHATE-DEHYDROGENASE FROM LEUCONOSTOC-MESENTEROIDES. Journal of the American Chemical Society **103:**4890-4899.

178. **Wong, T. S., F. H. Arnold, and U. Schwaneberg.** 2004. Laboratory evolution of cytochrome p450 BM-3 monooxygenase for organic cosolvents. Biotechnol Bioeng **85:**351-8.

179. **Wunderlich, M., A. Martin, and F. X. Schmid.** 2005. Stabilization of the cold shock protein CspB from Bacillus subtilis by evolutionary optimization of coulombic interactions. Journal of Molecular Biology **347:**1063-1076.

180. **Yamamoto, K., G. Kurisu, M. Kusunoki, S. Tabata, I. Urabe, and S. Osaki.** 2001. Crystal structure of glucose dehydrogenase from Bacillus megaterium IWG3 at 1.7 angstrom resolution. Journal of Biochemistry **129:**303-312.

181. **Zavodszky, P., J. Kardos, A. Svingor, and G. A. Petsko.** 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. Proceedings of the National Academy of Sciences of the United States of America **95:**7406-7411.

182. **Zhang, T., E. Bertelsen, and T. Albert.** 1994. Entropic Effects of Disulfide Bonds on Protein Stability. Nature Structural Biology **1:**434-438.

183. **Zhao, H. M., and W. A. van der Donk.** 2003. Regeneration of cofactors for use in biocatallysis. Current Opinion in Biotechnology **14:**583-589.

VITA

EDUARDO VAZQUEZ-FIGUEROA

Eduardo Vázquez-Figueroa is a native of Bayamón, Puerto Rico. He graduated in 1998 from Antilles High School and graduated with a B.S. in Chemical Engineering from the University of Puerto Rico Mayagüez in 2003. Eduardo is highly experienced in many research fields. He has participated in five internship programs: in electrochemistry at the University of Cornell, in polymer applications at the Massachusetts Institute of Technology and at Kimberly-Clark, and twice in industrial biocatalysis with Merck & Co. Inc. At GA Tech, Eduardo has been the recipient of several graduate fellowships, including the prestigious National Science Foundation Graduate Research Fellowship. He is a coauthor on five publications, one patent and one patent application. His research, conducted in the biocatalysis laboratory of Prof. Andreas S. Bommarius, focuses on the improvement of dehydrogenase enzymes by various protein engineering techniques. In 2007, Eduardo was honored with the Best Teaching Assistant award. When Eduardo is not working he enjoys reading, movies, and dancing. Eduardo is en route to Merck & Co. Inc. at Barceloneta, Puerto Rico where he will be joining the Global Technical Operations team.