

# Computational design of *Candida boidinii* xylose reductase for altered cofactor specificity

George A. Khoury, Hossein Fazelinia, Jonathan W. Chin,  
Robert J. Pantazes, Patrick C. Cirino, and Costas D. Maranas\*

Department of Chemical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

Received 2 April 2009; Revised 22 June 2009; Accepted 11 August 2009

DOI: 10.1002/pro.227

Published online 19 August 2009 proteinscience.org

**Abstract:** In this study we introduce a computationally-driven enzyme redesign workflow for altering cofactor specificity from NADPH to NADH. By compiling and comparing data from previous studies involving cofactor switching mutations, we show that their effect cannot be explained as straightforward changes in volume, hydrophobicity, charge, or BLOSUM62 scores of the residues populating the cofactor binding site. Instead, we find that the use of a detailed cofactor binding energy approximation is needed to adequately capture the relative affinity towards different cofactors. The implicit solvation models Generalized Born with molecular volume integration and Generalized Born with simple switching were integrated in the iterative protein redesign and optimization (IPRO) framework to drive the redesign of *Candida boidinii* xylose reductase (CbXR) to function using the non-native cofactor NADH. We identified 10 variants, out of the 8,000 possible combinations of mutations, that improve the computationally assessed binding affinity for NADH by introducing mutations in the CbXR binding pocket. Experimental testing revealed that seven out of ten possessed significant xylose reductase activity utilizing NADH, with the best experimental design (CbXR-GGD) being 27-fold more active on NADH. The NADPH-dependent activity for eight out of ten predicted designs was either completely abolished or significantly diminished by at least 90%, yielding a greater than 10<sup>4</sup>-fold change in specificity to NADH (CbXR-REG). The remaining two variants (CbXR-RTT and CBXR-EQR) had dual cofactor specificity for both nicotinamide cofactors.

**Keywords:** computational protein design; IPRO; cofactor switching; protein engineering; xylitol; cofactor specificity; nicotinamide

## Introduction and Background

The ability of enzymes to catalyze chemical reactions with great specificity, efficiency, and selectivity provides the basis of metabolism in all living organisms. By carefully redesigning metabolism through enzyme

modification, many desired biocatalytic transformations can be efficiently carried out in a variety of microbial production hosts. Proteins have been previously computationally designed to bind new ligands,<sup>1</sup> proteins,<sup>2</sup> and nucleic acids,<sup>3</sup> to improve protein stability,<sup>4,5</sup> as well as to introduce novel enzymatic activity,<sup>6,7</sup> demonstrating that the fundamental rudiments of molecular recognition can adequately be captured via computational design. The systematic fine-tuning of molecular recognition between proteins and ligands finds many biotechnological applications ranging from improved catalytic activity,<sup>8</sup> improved protein thermostability,<sup>9–11</sup> genetic circuits,<sup>12</sup> biosensors,<sup>13,14</sup> chiral separations,<sup>15</sup> the construction of novel enzymes with alternative functionality,<sup>16,17</sup> the creation of gene

---

Conflict of Interest: The authors declare no conflicts of interest.

Grant sponsor: National Science Foundation Award; Grant number: CBET-063992; Grant sponsor: REU; Grant number: EEC-0353569.

\*Correspondence to: Costas D. Maranas, Department of Chemical Engineering, 112 Fenske Laboratory, The Pennsylvania State University, University Park, PA 16802. E-mail: costas@psu.edu

**Table I.** Summary<sup>a</sup> of NAD(P)(H) Cofactor Engineering Studies Extending from Marohnic *et al.*<sup>22</sup>

Source	Enzyme	Specificity change	Mutation(s) <sup>b</sup>	Reference(s)
<i>Candida tenuis</i>	Xylose reductase	NADPH → NADH	K274R, K274G, K274M, S275A, N276D, R280H, K274R/N276D	26,27
<i>Corynebacterium</i>	2,5-diketo-D-gluconic acid	NADPH → NADH	K232G, R235G, R238H and F22Y/RS233T/R235E/A272G	29,30
<i>Escherichia coli</i>	Glutathione reductase	NADPH → NADH	A179G/A183G/V197E/ R198M/ K199F/H200D/R204P	31
<i>Escherichia coli</i>	Ketol acid reductoisomerase	NADPH → NADH	R68D, K69L, K75V, R76D	32
<i>Neurospora crassa</i>	Nitrate reductase	NADPH → NADH	S920D/R932S	33
<i>Pichia stipitis</i>	Xylose reductase	NADPH → NADH	K270M, K270S/ S271G/N272P/ R276F	25,28
<i>Pseudomonas fluorescens</i>	p-hydroxybenzoate hydroxylase	NADPH → NADH	R33S/Q34R/P35R/D36A/ Y37E	34
<i>Rattus norvegicus</i>	Cytochrome p450 reductase	NADPH → NADH	S596D	35
<i>Saccharomyces cerevisiae</i>	17β-hydroxysteroid dehydrogenase	NADPH → NADH	Y49D	36
<i>Sinorhizobium morelense</i>	1,5-anhydro-D-fructose reductase	NADPH → NADH	A13G/S33D	37
<i>Anabaena. sp.</i> (strain PCC 7119)	Ferredoxin: NADP+ reductase	NADP+ → NAD+	S223D	38
<i>Escherichia coli</i>	Isocitrate dehydrogenase	NADP+ → NAD+	C201I/C332Y/K344D/Y345I/ V351A/Y391K/R395S	39
<i>Thermus thermophilus</i>	Isocitrate dehydrogenase	NADP+ → NAD+	K283D/Y284I/N287G/V288I/ I290A	40
<i>Vibrio harveyi</i>	Aldehyde dehydrogenase	NADP+ → NAD+	T175D, T175E, T175S, T175N, T175Q	41
<i>Bacillus stearothermophilus</i>	L-lactate dehydrogenase	NADH → NADPH	I51K/D52S	42
<i>Rattus norvegicus</i>	Cytochrome b5 reductase	NADH → NADPH	D239T	22
<i>Spinacia oleracea</i>	Nitrate reductase	NADH → NADPH	E864S/F876R	43
<i>Thermus thermophilus</i>	β-isopropylmalate dehydrogenase	NADH → NADPH	D236R/D289K/I290A/ A296V/ G337Y	44
<i>Bacillus stearothermophilus</i>	D-lactate dehydrogenase	NAD+ → NADP+	D175A	45
<i>Bacillus stearothermophilus</i>	Glyceraldehyde-3-phosphate dehydrogenase	NAD+ → NADP+	D32A/L187A/P188S	46
<i>Gluconobacter oxydans</i>	xylitol dehydrogenase	NAD+ → NADP+	D38S/M39R	47
<i>Homo sapien</i>	Human mitochondrial NAD(P)-dependent malic enzyme	NAD+ → NADP+	Q362K	48
<i>Pichia stipitis</i>	Xylitol dehydrogenase	NAD+ → NADP+	D207A/I208R/F209S/N211R	24
<i>Pseudomonas stutzeri</i>	Phosphite dehydrogenase	NAD+ → NADP+	E175A/A176R	23
<i>Saccharomyces cerevisiae</i>	Formate dehydrogenase	NAD+ → NADP+	D196A/Y197R	49
<i>Thermus thermophilus</i>	Isopropylmalate dehydrogenase	NAD+ → NADP+	S226R/D278K/I279Y/A285V/ P324T/P325Y/G328E/G329R/ S330L	50
<i>Tramitichromis intermedius</i>	Leucine dehydrogenase	NAD+ → NADP+	D203A/I204R/D210R	51

<sup>a</sup> In all studies, structural analysis was used to determine residues to be mutated. Mutations were introduced by site-directed mutagenesis except for Liang *et al.*,<sup>28</sup> who used a combinatorial saturation mutagenesis approach.

<sup>b</sup> The best mutants reported in each study are summarized in this table. Multiple mutations occurring in a single mutant are separated by “/”. Commas are used to separate individual mutants.

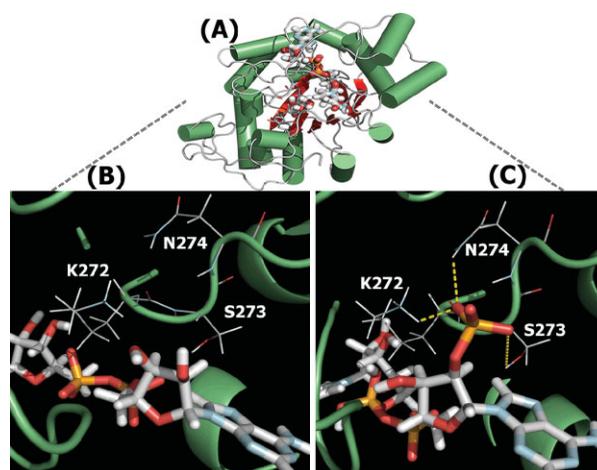
switches<sup>18</sup> and signal transduction pathways.<sup>19,20</sup> Many of the aforementioned applications require the enzymes to operate under unnatural conditions (e.g., at elevated temperatures or in nonaqueous environments), and/or possess altered cofactor or substrate specificity.<sup>21</sup> Even with these successes, predictably changing a protein's cofactor specificity has not been reported via a systematic computational workflow.

In the past few years, there have been many reported successes of enzyme redesign for altered cofactor specificity utilizing structural analysis with site-directed mutagenesis as their method for redesign. Table I summarizes the best identified mutations involved in changing cofactor specificity (extending an earlier compilation).<sup>22</sup> Key successful redesigns include the work of Woodyer *et al.*<sup>23</sup> that succeeded in

relaxing the cofactor specificity of *Pseudomonas stutzeri* phosphite dehydrogenase from 100-fold in favor of nicotinamide adenine dinucleotide (NAD<sup>+</sup>) to three-fold in favor of nicotinamide adenine dinucleotide phosphate (NADP<sup>+</sup>) using homology modeling and site-directed mutagenesis to identify and construct a double mutant. This double mutant showed potential as an efficient *in vitro* NAD(P)(H) regeneration system for reductive biocatalysis.<sup>23</sup> Watanabe *et al.*<sup>24</sup> used site-directed mutagenesis to change cofactor specificity of a *Pichia stipitis* NAD<sup>+</sup>-dependent xylitol dehydrogenase (PsXDH) from NAD<sup>+</sup> to NADP<sup>+</sup> as part of an efficient biomass-ethanol conversion system. Their designs yielded greater activity for NADP<sup>+</sup> than NAD<sup>+</sup> after redesign. Kostrzynska *et al.*<sup>25</sup> found that in the aldo-keto reductase (AKR) family of enzymes, the IPKS (Ile-Pro-Lys-Ser) motif is strictly conserved. They utilized site-directed mutagenesis at a conserved Lys-270 in *P. stipitis* xylose reductase (PsXR) to conclude that it binds to the 2'-phosphate of the NADPH (reduced form of NADP<sup>+</sup>). Site-directed mutagenesis-based studies also successfully pinpointed sets of mutations leading to complete reversal of *Candida tenuis* xylose reductase (CtXR) cofactor specificity from NADPH to NADH (reduced form of NAD<sup>+</sup>).<sup>26,27</sup> Similarly, Liang *et al.*<sup>28</sup> used a semirational approach called combinatorial active site saturation (CASTing) to switch cofactor preference from NADPH to NADH in PsXR.

Purely experimental design efforts relying on combinatorial library construction and screening have been successful for a number of cofactor alteration studies (see Table I), however, the lessons learned do not easily generalize to other systems. To address the lack of a systematic procedure, we introduce a generally applicable computational workflow based on the iterative protein redesign and optimization algorithm (IPRO).<sup>52</sup> The approach is tested for the xylose reductase enzyme from the yeast *Candida boidinii* (CbXR). Xylose reductase catalyzes the reduction of the open chain form of D-xylose to xylitol.

Xylose reductase belongs to the AKR superfamily.<sup>53,54</sup> The AKR superfamily shares a common ( $\alpha/\beta$ )<sub>8</sub>-barrel fold without a Rossmann-fold motif and their members show varied preferences for NADPH over NADH.<sup>55</sup> The active site, conserved in both structure and sequence in nearly all AKRs, is situated in a deep cavity inside the ( $\alpha/\beta$ )<sub>8</sub> barrel, and is defined by a tetrad of catalytic residues. In CtXR, these residues are Asp-46, Tyr-51, Lys-80, and His-113,<sup>56</sup> and are homologous to Asp-45, Tyr-50, Lys-79, and His-112 in CbXR. Previous studies<sup>56–59</sup> of AKRs have identified the functional role these residues have on the catalytic mechanism, but because they are further than 12 Å from the residues involved in determining cofactor specificity, they have minimal effect on cofactor binding. Structures of the apo- and holo- forms of CtXR have been determined to 2.2 Å resolution.<sup>55</sup> This



**Figure 1.** (A) The structure of the homology modeled CbXR with NADPH bound and D-xylose situated in a deep cavity inside the ( $\alpha/\beta$ )<sub>8</sub> barrel. (B) The cofactor binding pocket of wild-type CbXR containing NADH with no hydrogen bonding interactions near the 2'-hydroxyl group. (C) The cofactor binding pocket of CbXR containing hydrogen bonding interactions within 2.5 Å of the 2'-phosphate of NADPH. These hydrogen bonding interactions are important for the specificity of CbXR for NADPH over NADH. This figure was made using PyMOL (Delano Scientific).

enzyme selectively binds NADPH over NADH by roughly 20-fold.<sup>27</sup> In contrast, CbXR (62% homologous to CtXR) is strictly an NADPH-dependent enzyme. The structure of the homology modeled CbXR is shown in Figure 1(A), with NADPH bound and D-xylose situated inside the ( $\alpha/\beta$ )<sub>8</sub> barrel. In Figure 1(B), the cofactor binding pocket is shown with no hydrogen bonding interactions observed between wild-type CbXR and NADH. In Figure 1(C), hydrogen bond interactions are shown between the 2'-phosphate in NADPH and the surrounding residues Lys-272, Ser-273, and Asn-274. Alignment of AKRs reveals a conserved Lys residue near position 274 (amino acid position 274 in CtXR; position 272 in CbXR), which plays a critical role in cofactor binding.<sup>26</sup> One notable exception is the presence of an Arg residue rather than Lys at position 276 of the XR from *C. parapsilosis*, which prefers NADH as its cofactor.<sup>60</sup> Leitgeb *et al.* showed that replacement of Lys-274 with Arg in CtXR results in reversal of cofactor specificity for NADH over NADPH.<sup>26</sup>

Xylitol has been listed among the top value-added platform chemical products of biomass refining.<sup>61</sup> The production of xylitol from xylose by engineered *Escherichia coli* growing on glucose and expressing a xylose reductase from either *C. boidinii*, *C. tenuis*, *P. stipitis*, or *Saccharomyces cerevisiae* was recently studied.<sup>62</sup> Of the enzymes tested, functional expression of CbXR in *E. coli* resulted in the highest titers of xylitol production. It is unclear whether this is related to its strict requirement for NADPH or whether it is more a

**Table II.** NAD(P)(H) Binding Pockets Structurally Aligned with Combinatorial Extension

Protein	PDB code	Sequence positions aligned	RMSD binding pocket (Å)	Sequence identity of binding pocket (%) (20–50 residues)	Overall RMSD (Å)
CtXR	1MI3	270–290	0.0	100.0	0.0
GR	1GER	170–210	1.9	6.2	4.9
KARI	1YRL	60–80	2.9	12.5	4.4
PHBH	1PBB	20–50	2.5	12.5	4.5
2,5-DKG	1A80	220–250	0.8	50	1.3
1,5-AFR	2GLX	2–49	1.7	18.8	3.7
IDH	2D1C	280–300	3.25	12.5	4.7
GAPDH	1GD1	180–200	3.26	6.2	5.28
PsXDH	<sup>a</sup>	200–220	1.42	15.8	4.9
Ferredoxin: NADP <sup>+</sup> reductase	1QUF	220–240	3.31	6.2	5.15
L-LDH	1LDB	46–70	1.88	0	5.31

<sup>a</sup> This structure was generated utilizing the SWISS-MODEL first approach homology modeling method as there was no initial PDB crystal structure available.<sup>70,71</sup>

function of its expression characteristics. We explored whether xylitol production could be improved by enabling the use of NADH for xylose reduction. In addition to a number of other strategies recently explored,<sup>63</sup> we sought a variant of CbXR with either dual cofactor specificity or specificity toward NADH (which could then be coexpressed with wild-type CbXR). In addition, due to the higher stability of NADH relative to NADPH,<sup>29</sup> and the higher cost of NADPH regeneration compared to NADH generation,<sup>64</sup> a NADH-utilizing CbXR variant may prove industrially useful. We initially constructed the K272R mutation in CbXR and found this mutant to be active on NADH, while NADPH activity was weakened by fivefold. However, NADH-utilizing activity was less than 5% of the wild-type enzyme's activity with NADPH. We therefore sought to use computational design to more effectively engineer mutants with activity toward NADH.

The goal of this work was to explore the computational design of CbXR to bind (and subsequently oxidize) NADH as its cofactor. We first extracted and analyzed data from various cofactor usage alteration studies to pinpoint key interactions, factors, and trends that are discernable when performing cofactor switches between these particular substrates. We next validated the use of a computationally-derived interaction energy as a reasonable objective function and binding free energy surrogate by correlating it to published experimental binding results. This surrogate of cofactor affinity was found to correlate ( $R^2 = 72\%$ ) with experimental activities for a system previously designed using IPRO.<sup>52</sup> Our working hypothesis was that computationally generated sets of mutations that improve binding of NADH to CbXR will lead to mutants that exhibit enzymatic activity on NADH. Next, we modified the IPRO framework as presented by Saraf *et al.*<sup>52</sup> to improve modeling accuracy by adding implicit solvation models to drive the identification of sets of mutations that have increased affinity for NADH as evidenced by improved interaction energies,

as well as increased stability for the CbXR mutants relative to the wild-type. Lastly, we constructed and experimentally tested the best variants predicted by IPRO to assess the value of computations to drive redesign.

## Results

### Analysis of results from previous cofactor engineering studies

We first explored whether the experimentally observed binding affinities for NAD(P)(H) and/or enzymatic activities requiring these cofactors can be explained by using simple metrics such as residue volume, charge and hydrophobicity. Net charge,<sup>65</sup> hydrophobicity,<sup>66</sup> and side-chain volume<sup>67</sup> data for all amino acids were collected. A structural alignment was performed for the nicotinamide binding pockets targeted by mutational studies of the following proteins: glutathione reductase (GR),<sup>31</sup> ketol acid reductoisomerase (KARI),<sup>32</sup> *p*-hydroxybenzoate hydroxylase (PHBH),<sup>34</sup> 2,5-diketo-D-gluconic acid (2,5-DKG),<sup>29,30</sup> 1,5-anhydro-D-fructose reductase (1,5-AFR),<sup>37</sup> isocitrate dehydrogenase (IDH),<sup>39</sup> glyceraldehyde-3-phosphate dehydrogenase (GAPDH),<sup>46</sup> *P. stipitis* xylitol dehydrogenase (PsXDH),<sup>24</sup> ferredoxin: NADP<sup>+</sup> reductase,<sup>38</sup> and L-lactate dehydrogenase (L-LDH).<sup>42</sup> The nicotinamide cofactor binding pockets of these proteins were aligned to the NADPH binding pocket of CtXR using Combinatorial Extension.<sup>68</sup> These proteins were chosen as they are well characterized and most had high resolution crystal structures available. The structural alignments used RCSB Protein Data Bank (PDB)<sup>69</sup> crystal structures to provide the atomic coordinates for all structures except for PsXDH, which was constructed via the SWISS-MODEL first approach method.<sup>70,71</sup> The results of the different nicotinamide binding pockets structurally aligned with Combinatorial Extension are shown in Table II. Significant structural similarity in the nicotinamide cofactor binding pockets was found across the enzymes used based on their root of mean square deviation (RMSD)

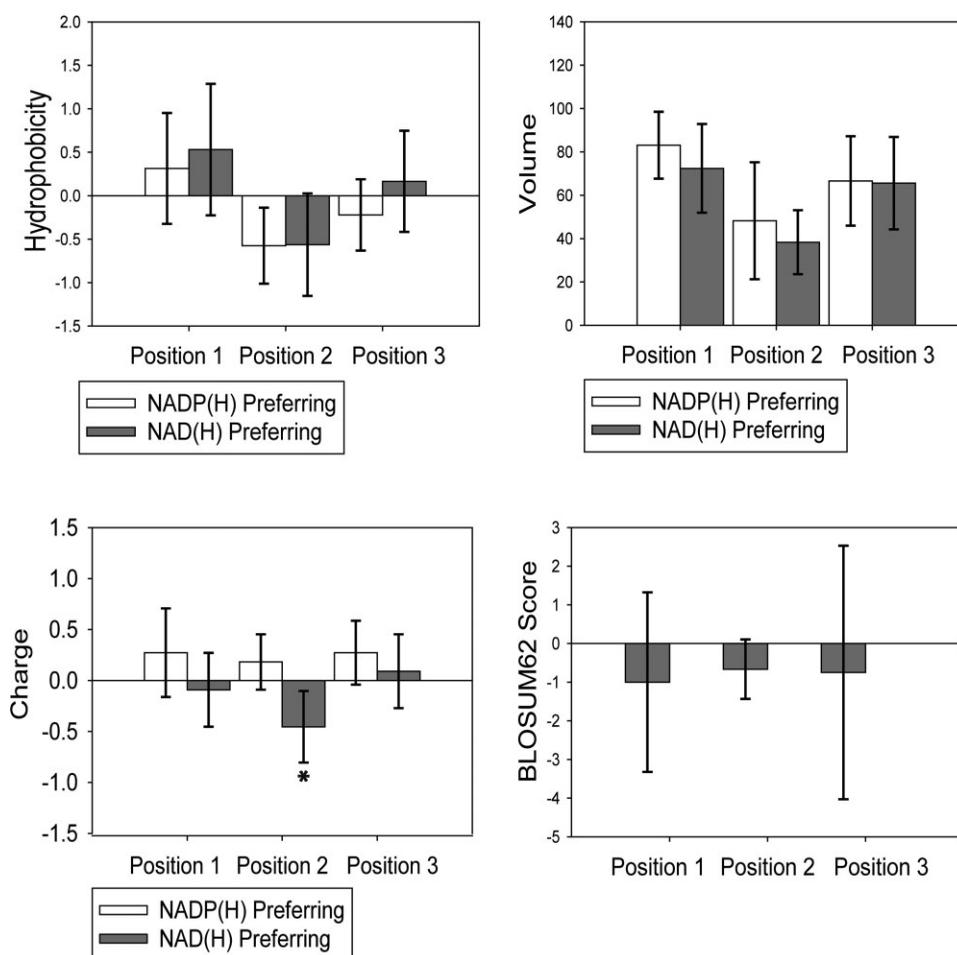
**Table III.** Residues Used in Calculating Average Properties of NAD(P)H-Binding Residues

Protein	NADP(H) → NAD(H) preferring		
	Position 1	Position 2	Position 3
CtXR	Lys→Arg	Ser→Ala	Asn→Asp
GR	Val	Gly	Ala
KARI	Leu	Arg→Asp	Lys→Leu
PHBH	Glu	Arg→Ser	Gln→Arg
2,5-DKG	Lys→Gly	Ser	Val
1,5-AFR	Met	Ser→Asp	Thr
IDH	Lys→Asp	Tyr→Iso	Ala
GAPDH	Ala→Leu	Ser→Pro	His
PsXDH	Val	Ala→Asp	Arg→Iso
Ferredoxin:NADP+ reductase	Iso	Ser→Asp	Arg
l-LDH	Lys→Iso	Ser→Asp	Ala

values. In contrast, no significant sequence alignment occurred in the binding pockets of the sampled proteins and that of CtXR, except in 2,5-DKG, where there is 50% sequence similarity.

Based on previous mutational studies performed on CtXR, positions Lys-274, Ser-275, and Asn-276 emerged as key locations to mutate to increase cofactor specificity for NADH over NADPH.<sup>27,53,72</sup> We defined Positions 1–3 as the residues that are aligned to K274, S275, and N276, respectively. Positions 1–3 are nearby the phosphate group in NADPH, but are over 12Å from the hydride transfer site in the catalytic mechanism, highlighting that these positions affect cofactor specificity and affinity but are not directly involved in the reaction. Next, a statistical analysis on charge, hydrophobicity, and volume was performed for each design position structurally aligned to CtXR in both their NADP(H) and NAD(H)-preferring forms for the residues listed in Table III. This allowed us to discern whether any of those metrics played an identifiable role in cofactor specificity. For each position, we calculated the average value of each parameter, as depicted in Figure 2 with error bars representing 95% confidence intervals.

While differences exist between the average values for charge, hydrophobicity, and volume, the average



**Figure 2.** Comparison of average hydrophobicity, volume, charge, and BLOSUM62 score for all design positions. Error bars are shown for a 95% confidence interval. No statistically significant signal was found except for charge in position 2, where NADH-preferring residues were found to be more negative than NADPH-preferring residues, which is consistent with previous reports in the literature.

values are well within their confidence intervals for the mean, indicating no statistically significant signal. In position 2, enzymes preferring NAD(H) over NADP(H) were on average more negative compared with NADP(H)-preferring enzymes, which is consistent with what would be expected. The more positively charged residues electrostatically interact with the negatively charged phosphate of the adenosine ribose in NADP(H). The residues that are more negative in the NAD(H)-preferring enzymes may be compensating for the lack of the negative 2'-phosphate present in NADP(H) and are stabilizing the 2'-OH in the enzymes' NAD(H)-bound form.<sup>32,34</sup> In addition, we performed a similar analysis using the BLOSUM62<sup>73</sup> scores of the mutations in each position leading to altered cofactor specificity. The BLOSUM62 scores reported are based on the change in amino acid when going from NADPH-preferring to NADH-preferring residues. There was no statistically significant difference in the average scores per position. Notably, the mutations considered resulted, on average, in negative BLOSUM62 scores, indicating generally nonconservative mutations.<sup>74</sup> These results do not mean that charge, hydrophobicity, volume, and BLOSUM62 scores do not have an effect on the affinity for different cofactors. Instead, they imply that the effect of each factor separately is not monotonic or even discernible in isolation of all other metrics. Therefore, straightforward approaches using size, charge or hydrophobicity observations to suggest successful enzymatic redesigns cannot be successfully applied. Given the insufficiency of simple metrics to drive redesign, we next explored whether calculated binding affinities could be used to support enzyme redesign.

### **Comparison of calculated interaction energies of enzyme-NAD(P)(H) complexes versus affinities**

Here we explore whether cofactor interaction energy is an adequate surrogate of cofactor specificity to drive computational cofactor alteration. To test this, we contrasted calculated interaction energy values (through CHARMM<sup>75,76</sup>) with published kinetic parameter data from a study aimed at changing specificity from NADPH to NADH in CtXR.<sup>27</sup> We compare the results of interaction energy changes calculated with and without solvation effects to determine whether the substantially increased computational cost needed for solvation is necessary.

The crystal structure of CtXR with NADH bound (PDB:1MI3) provided the starting coordinates for this analysis.<sup>69</sup> For this complex, we imposed a harmonic restraint to all nonhydrogen atoms with a force constant of 0.1 and mass weighting enabled. The CHARMM force field was applied and the complexes were energy minimized using the Adopted Basis-set Newton-Raphson (ABNR)<sup>75</sup> method with the Generalized Born with a simple switching implicit solvent

model (GBSW).<sup>77,78</sup> The energy function in CHARMM accounts for forces from van der Waals interactions, bond stretching, bond angles, dihedral (torsion) angles, improper dihedral angles, electrostatics, and solvation. All minimizations converged successfully within the iteration limit. The interaction energy for the minimized wild-type complex was calculated using the Generalized Born with molecular volume integration (GBMV)<sup>79,80</sup> implicit solvent model as:

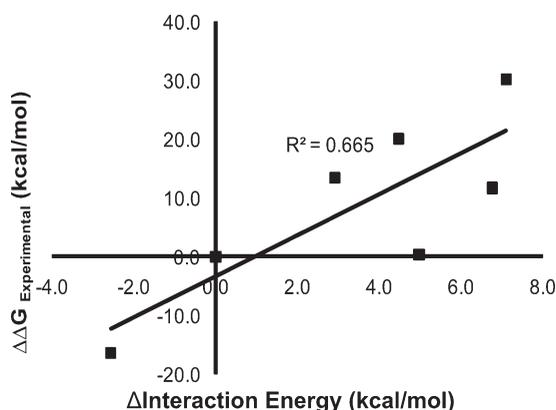
$$\text{Interaction energy} = \text{energy of complex} - \text{energy of Apo enzyme} - \text{energy of cofactor} \quad (1)$$

The interaction energy represents the intermolecular component of total energy. The minimized wild-type structure was then mutated in positions relevant to altering cofactor specificity for NAD(P)H as reported in the literature.<sup>27</sup> Each mutated structure was minimized and had its interaction energy calculated using the same methods applied to the wild-type structure.  $\Delta$ Interaction energy was then calculated as:

$$\Delta\text{Interaction energy} = \text{interaction energy}_{(\text{mutant})} - \text{interaction energy}_{(\text{wild-type})} \quad (2)$$

Here we used GBSW in all energy minimizations to ensure proper packing of hydrophobic cores while GBMV was used to approximate the solvation component for the interaction energy calculations. In the GBMV method,<sup>79,80</sup> the effective Born radius is computed by numerical integration of the molecular volume. The Coulomb field approximation includes a higher order correction term to improve agreement with the radii calculated from solving the Poisson-Boltzmann equation. The GBMV method was used for the interaction energy calculations because it is highly accurate but still more tractable in an iterative form than other options, such as solving the Poisson-Boltzmann equation. The minimizations utilized the GBSW model, as the GBMV model may utilize a sharp molecular surface representation for some systems, which would lead to large fluctuations in energy and cause stability problems in the simulations. GBSW is very similar to GBMV, but it is 2–3 times faster since it replaces the computationally expensive molecular surface calculation with a simple smoothing function at the dielectric boundary. GBSW's use of a smoothed dielectric boundary allows the change in polarization forces to vary more smoothly compared to GBMV.

For mutations changing cofactor specificity from NADPH to NADH, Figure 3 shows the calculated changes in interaction energy from (wild-type to mutant) including solvation against experimental ground state binding energy data outlined in Petschacher *et al.*<sup>27</sup> The correlation coefficient value is equal to 67%. This implies the calculated interaction energy



**Figure 3.** Changes in experimental ground state binding energies from Petschacher *et al.*<sup>27</sup> versus our calculated changes in interaction energies. Shown are the changes in interaction energy with solvation showing reasonable correlation with the experimental data ( $R^2 = 67\%$ ), whereas changes in interaction energy without solvation correlated significantly less with the experimental data ( $R^2 = 24\%$ ) (data not shown).

explains 67% of the variance in the experimental binding data. Although not a fully quantitative description, this is generally sufficient for rank ordering of different enzyme redesigns. When eliminating the implicit solvation model GBMV from the energy calculations, the correlation was reduced to 24% (data not shown), implying the need to include solvation effects in enzyme redesign. With these observations, we next modified and used the IPRO<sup>52</sup> framework to account

for solvation based on the GBSW and GBMV models to explore redesigns for CbXR.

### Computational predictions using IPRO

Using the modified IPRO, we were able to generate variants of CbXR with improved interaction energies for NADH by targeting the design positions Lys-272, Ser-273, and Asn-274 in the NADPH binding pocket.

The wild-type interaction energies of CbXR-NADH and CbXR-NADPH were calculated to be  $-232$  kcal/mol and  $-339$  kcal/mol, respectively, and the interaction energy improvements towards NADH as a result of mutations predicted by IPRO for the top 10 designs are reported in Table IV. The mutants generated improvements in interaction energies for NADH by up to 78% relative to the wild-type and were selected among the  $20^3$  ( $= 8000$ ) possible combinations of mutations. The interaction energies of the redesigned variants with the native cofactor NADPH were also calculated to assess the effect of the NADH binding improving mutations on the retention or abolishment of affinity for NADPH. Notably, we found that mutations in position 272 to methionine to be most effective at suppressing binding affinity based on an increase in interaction energy for the native cofactor. This is in agreement with the experimental results derived by Petschacher *et al.*<sup>27</sup> who found that the K274M mutation in the homologous CtXR increases NADPH dissociation and reduces the catalytic efficiency of CtXR utilizing NADPH. The increased hydrophobicity of the methionine side chain relative to

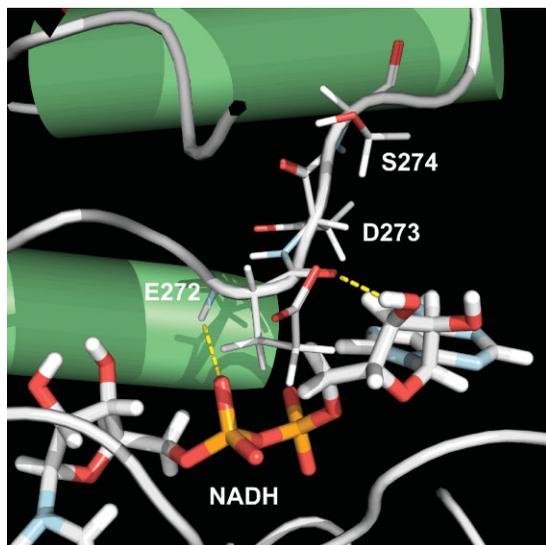
**Table IV.** Computational and Experimental Results

Mutations	$\Delta$ Interaction energy <sub>NADH</sub> (kcal/mol)	$\Delta$ Interaction energy <sub>NADPH</sub> (kcal/mol)	Activity with NADH (mU/mg protein)	Activity with NADPH (mU/mg protein)	Specificity (NADH/NADPH)	$\Delta$ Charge
Wild-Type	0	0	$0.7 \pm 2.2$	$78.6 \pm 4.7$	0.01	0
EDS	-181	42	$12.8 \pm 2.1$	$5.5 \pm 2.7$	2.4	-3
EDR	-133	29	$1.9 \pm 0.8$	<0.1	>19	-2
MGD	-129	181	$17.1 \pm 1.1$	$1.9 \pm 0.6$	8.9	-2
GGD	-126	195	$19.0 \pm 0.8$	$4.3 \pm 0.8$	4.4	-2
EQR	-103	-10	$4.1 \pm 0.6$	$72.3 \pm 5.2$	0.06	-1
RTT	-102	-30	$14.4 \pm 2.3$	$109.3 \pm 16.5$	0.13	0
MES	-99	180	$1.81 \pm 0.4$	<0.1	>18	-2
MAE	-92	266	$5.4 \pm 1.3$	<0.1	>54	-2
REG	-79	135	$11.2 \pm 3.1$	<0.1	>112	-1
RSE	-70	15	$10.8 \pm 1.2$	$29.5 \pm 0.4$	0.37	-1
R	-60	-9	$7.4 \pm 2.5$	$14.7 \pm 5.4$	0.5	0
Negative controls						
RNI	-45	74	<0.1	ND	ND	0
KKG	101	232	<0.1	ND	ND	1
RHC	-73	127	<0.1	ND	ND	0

<sup>a</sup> The top designs predicted by IPRO with their changes in interaction energies are reported. The mutation labels (e.g., EDS) correspond to positions 272, 273, and 274 respectively in CbXR. The wild-type interaction energy with NADH was calculated as  $-232$  kcal/mol and with NADPH as  $-339$  kcal/mol.

<sup>b</sup> The NADH and NADPH-linked activities of the CbXR variants are reported in this table for comparison with the computational predictions. Values of <0.1 indicate activity could not be detected above the background value in the absence of xylose. ND indicates the activity was not determined.

<sup>c</sup> The net local change in charge was calculated as a result of mutation and provided since charge was found to be important in determining cofactor specificity.



**Figure 4.** CbXR-EDS binding pocket containing NADH. The mutated residues Glu-272, Asp-273, and Ser-274 are labeled. Hydrogen bonding interactions are observed within 2.5 Å between the negative Glu-272 and the 3'-OH from NADH. This figure was made using PyMOL (Delano Scientific).

lysine may imply that the orientation of the methionine side chain with respect to bulk water is not favored.<sup>27,53</sup>

In Figure 1(B), no hydrogen bonding interactions were present between NADH and the design positions chosen in CbXR. In contrast, as shown in Figure 4, the best computationally-derived design, CbXR-EDS (involving three point mutations K272E, S273D, and N274S) improved the interaction energy by  $-181$  kcal/mol while forming a number of new hydrogen bonds between CbXR and the NADH. These newly formed hydrogen bonds likely explain the acquired affinity for NADH. Note that a hydrogen bond with the glutamic acid at position 272 stabilizes the 3'-OH of NADH (near the 2'-phosphate position of NADPH).

In nine out of the ten variants generated, including CbXR-EDS, the net charge change of the residues in the three positions considered is negative relative to the wild-type, with the change in CbXR-EDS being greatest (-3). This is in agreement with the results of the statistical analysis presented above. Presumably, this is because the more positively charged residues in the NADP(H)-bound enzymes electrostatically interact with the negatively charged phosphate of the adenosine ribose. The residues with a higher net negative charge change in the NAD(H)-preferring enzymes, specifically the Asp and Glu residues, are thought to provide a significant portion of substrate specificity for NAD(H) by hydrogen bonding to one or both of the 2'- and 3'-OH and to compensate for the lack of a partially negative 2'-phosphate present in NADP(H).<sup>23,32,34,81</sup> Also, in three of the top five designs, position 272 was mutated to glutamic acid,

indicating that this may be a critical mutation in changing the cofactor specificity of this enzyme.

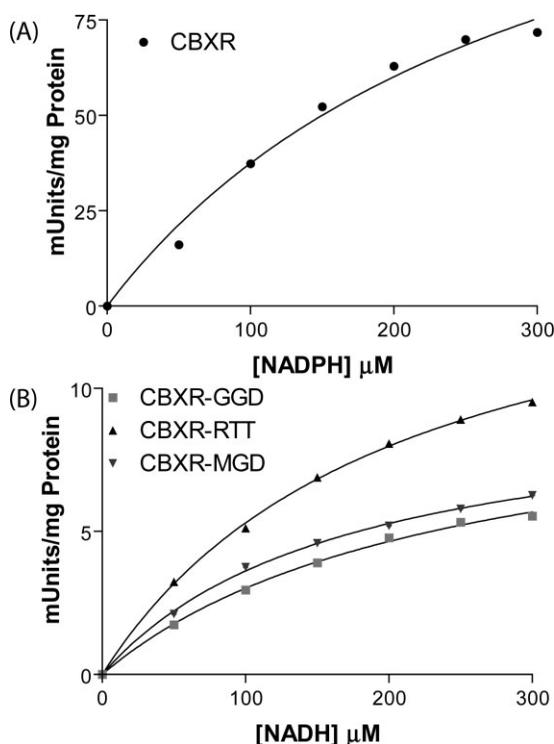
Interestingly, CbXR-EQR and CbXR-RTT increased binding affinity for NADH, as required by IPRO, but also increased binding affinity for the original cofactor NADPH. Of the mutants generated, CbXR-RTT was the only design in which the net charge change as a result of mutation in the three design positions did not change. Comparing CbXR-RTT to the wild-type, there is no significant change in hydrophobicity or side-chain volume in any of the residues compared to the wild-type. Conservative increases in side-chain volume as a result of the mutations may slightly increase van der Waals and hydrogen bonding interactions to fine-tune the enzyme to bind NADH as well, without disrupting the original hydrogen bonding network and positive charge preference of the 2'-phosphate of NADPH

With these computationally-predicted designs, we next experimentally assessed the effect of the predicted mutations on cofactor preference to assess the efficacy of our computational predictions.

### Experimental results

We experimentally constructed the top 10 predicted designs to test the computational procedure and also shed light onto the functional significance of mutations in the binding pocket of CbXR. One additional mutant (CbXR-R) was also constructed by mutating Lys-272 to Arg. The wild-type lysine in this position provides a positive charge for NADPH binding and the mutation of this residue to Arg was previously shown to change the cofactor specificity of CbXR from NADPH to NADH.<sup>26,27</sup> As negative controls, we also constructed three mutants not predicted by IPRO (CbXR-RNI, -KKG, -RHC).

Specific activities ( $\mu\text{mol}/\text{min}/\text{mg}$ ) of clarified cell lysates containing the engineered CbXR mutants in the presence of  $300 \mu\text{M}$  NADH and  $300 \text{ mM}$  D-xylose were measured and are presented in Table IV. Wild-type CbXR, as expected, clearly showed activity for NADPH ( $78.6 \text{ mU}/\text{mg}$  protein) and less than  $1 \text{ mU}/\text{mg}$  protein of activity for NADH. Interestingly, while all top 10 predicted designs clearly displayed some levels of NADH-linked enzymatic activity, all three negative controls exhibited a complete loss of reductase activity. Interaction energy calculations were performed on the negative controls for completeness. Notably, CbXR-KKG was calculated to have worse affinity for both cofactors, which is consistent with the observation from results presented above for mutants having a net positive charge change. CbXR-RNI and CbXR-RHC were calculated to have increased affinity for NADH and decreased affinity for NADPH, with zero net charge change. The lack of NADPH activity for these mutants with no local charge change bolsters the importance of charge in determining specificity and affinity for cofactor.



**Figure 5.** Michaelis-Menten plot for (A) wild type CbXR with NADPH and (B) three tested variants of engineered CbXR with NADH.

Experimental results for redesigning the cofactor binding pocket of CbXR for NADH specificity confirmed a number of important computationally predicted redesign trends. Cofactor specificity of this enzyme is markedly influenced by different amino acid substitutions in three design positions. Replacement of Lys-272 by Arg, which was previously shown to completely reverse nicotinamide cofactor specificity in CtXR,<sup>26</sup> also yielded NADH activity in CbXR while weakening the NADPH-linked catalytic activity (by approximately fivefold; see Table IV). While this mutant did not make the top 10 predicted designs, off-line interaction energy calculations showed a significant  $-60$  kcal/mol (26%) improvement in interaction energy toward NADH relative to the wild-type CbXR. The effect of this mutation on NADH binding is clearly dependent on amino acids in positions 273 and 274. In the presence of Arg at position 272, mutation of Ser273 and Asn274 to larger and more hydrophobic amino acids in CbXR-RNI and CbXR-RHC resulted in a complete loss of reductase activity, whereas smaller and more hydrophilic amino acids at these positions

in CbXR-REG, CbXR-RTT, and CbXR-RSE exhibited improved enzymatic activity.

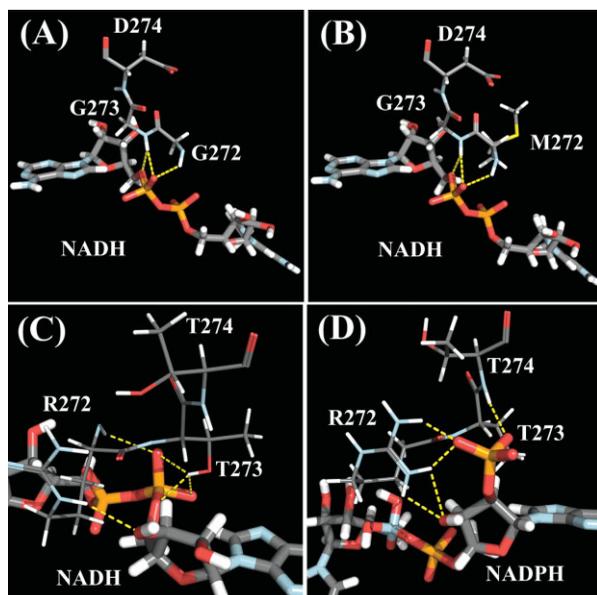
In agreement with the computational results, methionine in position 272 is found to improve binding and activity for NADH while abolishing activity for NADPH. It appears that more negatively charged residues in the design positions help to explain the observed cofactor affinity alterations, as the net charge change for the three residues in CbXR-MGD, CbXR-MAE, and CbXR-MES is negative relative to wild-type. This may serve to compensate for the lack of a partially negative 2'-phosphate in NADP(H).<sup>23,32,34,81</sup>

Of the mutants experimentally tested, only CbXR-RTT showed activity toward NADH and also increased activity for NADPH. This is consistent with the computational results in that the binding affinity for both cofactors was increased for this mutant (Table IV). CbXR-EQR was predicted computationally to have a small increase in affinity for NADPH while also binding NADH. Experimental results revealed a slight decrease ( $\sim 8\%$ ) in activity for NADPH while introducing novel activity for NADH. Cofactor specificity of the designed mutants was measured as the ratio of activity on NADH versus NADPH. Seven of the ten predicted mutations exhibited specificity values greater than one, indicating greater specificity for NADH. Four mutants (EDR, MES, MAE, REG) exhibited completely diminished ( $< 0.1$  mU/mg) activity on NADPH, most likely as a result of local charge repulsion between the 2'-phosphate and the more negative residues in the design region. Mutant CbXR-REG exhibited a greater than  $10^4$ -fold change in substrate specificity from NADPH to NADH.

The variants that showed the highest activity toward NADH, (i.e., CbXR-GGD, CbXR-MGD, and CbXR-RTT) were further analyzed by determining their Michaelis kinetic parameters for NADH and NADPH in the presence of saturating concentrations of D-xylose (300 mM). Data were fitted to the Michaelis-Menten equation for a single substrate using nonlinear least squares regression as shown in Figure 5.  $K_m$  and  $V_{max}$  values are listed in Table V. The  $K_m$  values for the mutant enzymes and wild-type CbXR are comparable, however, the  $V_{max}$  values for these mutants are approximately one order of magnitude lower than the one for the wild-type enzyme. This suggests that NADH binding strength for these mutants is comparable to that of NADPH to the wild-type, and that IPRO successfully improved substrate binding. Figure 6 highlights the computationally-predicted enzyme-cofactor interactions for the best three mutant enzymes. NADH binding is suggested to be stabilized by a network of hydrogen

**Table V.** Michaelis-Menten Constants for Wild-type and Mutant CbXR

	Engineered CbXR with NADH			Wild-type CbXR with NADPH
	CbXR-GGD	CbXR-RTT	CbXR-MGD	CbXR
$K_m$ ( $\mu M$ )	$238 \pm 24$	$205 \pm 12$	$169 \pm 12$	$307 \pm 87$
$V_{max}$ (mUnits/mg)	$10 \pm 0.6$	$16 \pm 0.5$	$9 \pm 0.3$	$152 \pm 25$



**Figure 6.** Structures of redesigned NAD(P)H binding pockets. (A) CbXR-GGD and (B) CbXR-MGD establish new hydrogen bond interactions between the mutated residues in CbXR and the bridging phosphates in NADH. The net charge change of these mutations is negative which may serve to compensate for the lack of negative 2'-phosphate in NADH. The mutations to glycine may serve to add conformational flexibility in the backbone to allow proper positioning of the NADH. CbXR-RTT, the mutation predicted by IPRO that was experimentally found to have dual cofactor specificity, bound to NADH (C) and NADPH (D). New hydrogen bond interactions are shown stabilizing the 3'-phosphate in NADH and NADPH from Arg-272, which may be the cause of the dual cofactor specificity. In NADPH, new hydrogen bonds are found to stabilize the 2'-phosphate group from Arg-272 and Thr-274. A neutral net change in charge is thought to contribute to dual cofactor specificity as well. All hydrogen bonds shown are within 2.5 Å. This figure was made using PyMOL (Delano Scientific).

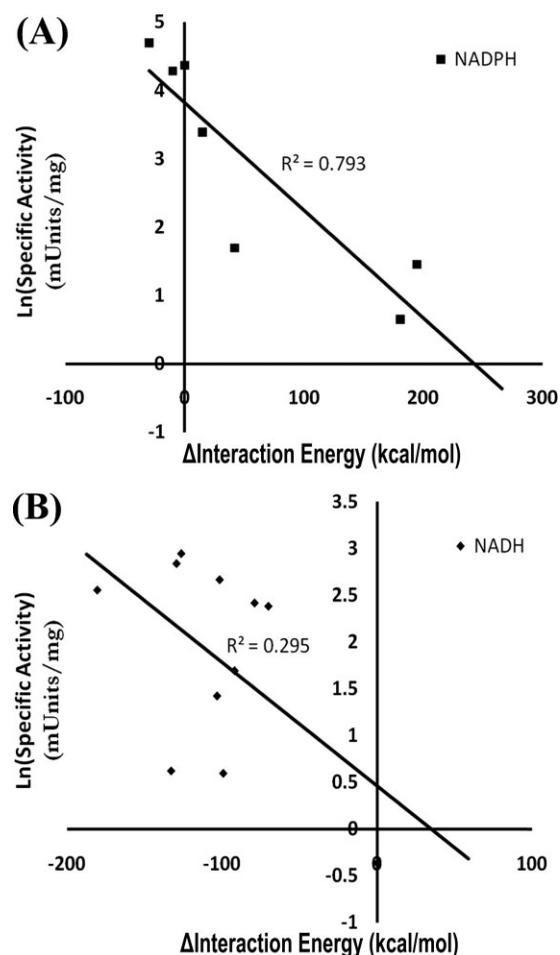
bonds, absent in the wild-type enzyme, as well as van der Waals interactions between the side chains of residues in the design positions and the 2'- and 3'-OH groups in NADH. In CbXR-GGD [Fig. 6(A)] and CbXR-MGD [Fig. 6(B)], new hydrogen bonding interactions were established between the new residues and bridging phosphate groups in NADH. It is interesting that mutations to glycine were selected, perhaps to introduce conformational flexibility that allows better placement of the new cofactor in the binding pocket. In CbXR-RTT, new hydrogen bonding interactions appear to stabilize the 3'-hydroxyl group both for NADH [Fig. 6(C)] and NADPH [Fig. 6(D)]. New hydrogen bonds from Arg-272 and Thr-274 are found to stabilize the 2'-phosphate group in NADPH. These mutations yield a net neutral charge change, which may be why both cofactors can be bound without substantial electrostatic resistance.

Figure 7 plots the natural log of specific activity against interaction energy for all mutants. For NADPH

[Fig. 7(A)] there was a 79% correlation, and only 30% for NADH [Fig. 7(B)]. The difference in the ability of the interaction energy to predict differences in activity toward the two cofactors may be related to the fact that the position of NADPH is based on crystallographic data,<sup>55</sup> while NADH was computationally docked using ZDOCK (Version 2.3),<sup>82</sup> causing some of the catalytic atoms to be positioned suboptimally for the reaction to occur. An alternate explanation for this difference in correlations is that mutations that improve NADH binding may also disrupt xylose reduction to some extent, in which case activity will not necessarily correlate with interaction energy.

## Discussion

Redox enzyme variants with dual or switched cofactor preference are useful choices in metabolic engineering studies to better understand the role of cofactor utilization in strain performance. To date, strategies to engineer nicotinamide cofactor specificity have mainly relied on structural analysis and site-directed



**Figure 7.** Plots of the natural log of specific activity toward NADPH (A) or NADH (B) versus interaction energy for CbXR mutants described in this study. The correlation coefficient for mutants yielding activity for NADPH is 79%, whereas the correlation is only 30% for NADH.

mutagenesis (see Table I). Despite a number of successes, a systematic computational workflow to drive design of cofactor specificity has been absent.

In this article, using a modified IPRO workflow we identified sets of mutations that changed the nicotinamide cofactor specificity of CbXR from its physiological preference for NADPH, to the alternate cofactor NADH. We used calculated interaction energies to determine the increased or decreased affinities of CbXR variants for both nicotinamide cofactors, which were verified by our experimental results. Modifying the computational framework to account for implicit solvation<sup>77,78,79</sup> effects, we conclude that the increased computational expense needed to account in detail for solvation using GBSW and GBMV was warranted, as manifested by the successful experimental redesigns. Seven out of ten mutants proposed computationally to have increased affinity toward NADH were verified experimentally to bind and show significant activity toward NADH. Two variants identified by IPRO (i.e., CbXR-EQR and CbXR-RTT) led to dual cofactor specificity with preference for NADPH. Our results suggest interaction energies can successfully serve to introduce activity towards a new cofactor. Nevertheless, reaching the activity levels of the wild-type enzyme using the native cofactor for the redesigned enzymes using the new cofactor remains a challenge.<sup>23</sup> For example, for CtXR, Petschacher *et al.*<sup>27</sup> through site-directed mutagenesis was able to achieve increased catalytic efficiency for the alternate cofactor, but only at 27% of the native cofactor's efficiency (mutant K274R). Additional engineering efforts are therefore necessary to further increase activity toward NADH by expanding the list of positions for mutation. Specifically, it may be necessary to proactively design the catalytic atoms in the binding pocket.

Given there were only three design positions, we believe the reported top ten designs are a good representation of the top performing ones. The rotamer/residue selection step in IPRO converges to the globally optimal solution for the randomly perturbed  $\phi$  and  $\psi$  angles, however, a rigorous mathematical proof is not possible given the reliance on a simulated annealing step after every backbone relaxation/redocking step. Our computational results showed that the CbXR variants binding NADH are characterized by a net negative charge change in the binding pocket. We suggest that this net negative charge change coupled with the predicted new hydrogen bonding interactions between the mutants and NADH are important factors in ushering the change in CbXR's cofactor specificity. This is consistent with what has been observed in the literature: more negative residues in the binding pocket of NAD(H)-preferring enzymes compensate for the lack of partially negative 2'-phosphate of the NADP(H).<sup>32,34,81</sup> In summary, the computational procedure presented here can serve as a powerful tool for introducing enzyme activity toward a non-native

cofactor. It can be applied to other enzyme-cofactor systems, and the methodology can be extended to engineer specificity toward oxidized or reduced nicotinamide cofactors, as well as to non-nicotinamide cofactors of interest such as AMP and GMP.

The modified IPRO algorithm is available at <http://maranas.che.psu.edu>.

## Materials and Methods

### Modified IPRO computational procedure

The IPRO framework, which was previously developed by our group,<sup>52,83</sup> performs enzyme redesign by optimally identifying mutations in the protein sequence using energy-based scoring functions. The modified IPRO algorithm is available for download at <http://maranas.che.psu.edu>. In this effort, we added the implicit solvation models GBSW and GBMV to the minimization and interaction energy calculation steps, respectively. The steps of the algorithm are as follows. First, design positions are selected, and the torsion angles in a small region around a design position of the backbone are perturbed by up to  $\pm 5$  degrees. The vast majority of evolutionary engineering studies over the past 10 years involve simple uphill walks on the plot of fitness versus sequence.<sup>84</sup> As a result, the positions chosen for redesign of CbXR were Lys-272, Ser-273, and Asn-274 after structurally aligning CbXR with CtXR using Combinatorial Extension<sup>68</sup> between residues 200–290 and based on previous cofactor engineering studies on CtXR.<sup>26,27</sup> Next, all amino acid rotamers consistent with these torsion angles are selected at each position from the Dunbrack rotamer library.<sup>85,86</sup> For the design positions, the rotamers considered include all amino acids, whereas for non-design positions, the possible rotamers are only those from the native amino acid. Next, rotamer-rotamer and rotamer-backbone energies are calculated for all of the selected rotamers in the previous step using the energy function presented in Kuhlman *et al.*<sup>5</sup> A mixed-integer linear programming formulation is then used to select the optimal combination of rotamers in the design window such that the energy is minimized for the torsion angles considered. The backbone of the protein is next relaxed through energy minimization with the GBSW implicit solvation model to allow the backbone to adjust to the new side chains. The ligand position is then readjusted in the next step with respect to the modified backbone and side chains using the Fast-Fourier Transform ZDOCK docking software<sup>82</sup> (Version 2.3) with constraints added to block residues 8Å from the binding pocket from being considered in the docking step. The interaction energy of the protein-ligand complex is next evaluated with the GBMV implicit solvation model and the move is accepted or rejected based on whether the interaction energy has been improved relative to the best design thus far with the Metropolis criteria<sup>87</sup> to escape local

minima. Please refer to Saraf *et al.*<sup>52</sup> for further details of the algorithm. Here IPRO was used to identify the optimal set of rotamers or residues on CbXR in the NADPH binding pocket necessary to increase the affinity for NADH over NADPH.

Although a high-resolution crystal structure of CbXR has not been determined, the amino acid sequence of CtXR<sup>88</sup> is sufficiently similar to that of CbXR<sup>89</sup> to act as a plausible model for CbXR (62% sequence similarity). The model structure of CbXR was constructed by homology modeling through Modeller using defined geometrical restraints between the conserved atoms of binding pocket residues and the cofactor obtained from the homologous CtXR crystal structure with NADPH bound (PDB: 1K8C).<sup>55</sup>

IPRO was performed with the modifications for solvation on a Linux PC cluster using eight 3.06GHz Xeon CPUs with 4GB RAM for 2 CPU days to improve the interaction energies of CbXR for NADH. In each iteration, interaction energy calculations took ~6 seconds of CPU time per evaluated mutant.

### Experimental procedure

The redesigned proteins were constructed using standard site-directed mutagenesis techniques<sup>90</sup> and all sequences were verified by DNA sequencing. Proteins were then expressed in *E. coli* BL21 as follows: Seed cultures (10 mL in LB medium containing 50 µg/mL kanamycin) were grown at 37°C to an OD<sub>600</sub> of ~2.0 and were used to inoculate cultures by dilution to a final OD<sub>600</sub> of 0.1 in 125 mL of LB (50 µg/mL kan). When the cultures (at 37°C) reached an OD<sub>600</sub> of 0.6–0.7, protein expression was initiated by adding 1.0 mM IPTG and transferring the cell cultures to 30°C. After 9 hours of induction, cells were pelleted by centrifugation at 3200 g for 20 min, washed twice with 25 mL of 50 mM potassium phosphate buffer (pH 7.5). Cell pellets were stored at –20°C until use. The cell pellets were resuspended to a final OD<sub>600</sub> of 100 in ice-cold lysis buffer (50 mM potassium phosphate buffer (pH 7.5), 4 mM MgCl<sub>2</sub>, 3.3 µg/mL DNase I). Cells were lysed by three passes through a French Pressure cell press, and centrifuged at 4°C, 3750 g for 25 min to separate cellular debris. The resulting supernatant contained the soluble xylose reductase.

Xylose reductase activity was measured in 96-well microtiter plates using a Spectra Max Plus384 plate reader. A typical enzymatic reaction contained 300 mM xylose, 300 µM β-NADPH or 300 µM β-NADH, 50 mM potassium phosphate buffer (pH 7.5), 40 µL cell lysate supernatant and 5 mM KCN (to reduce background dehydrogenase activity) in 200 µL total final volume. Reduction in the β-NADH or β-NADPH concentration was monitored by the decrease in absorbance at 340 nm [extinction coefficient ~6.2 (mM cm)<sup>-1</sup>]. Reactions were initiated by adding reduced cofactor and measurements were taken every 3 seconds for 90 seconds. One unit is defined as the

enzyme activity that consumes 1 µmol of NADH or NADPH in one minute (background activity in the absence of xylose is subtracted). Total protein concentration was measured using the Quick Start™ Bradford protein assay protocol (Bio-Rad laboratories) based on binding of Coomassie Blue dye to proteins. Bovine serum albumin was used as a standard.

### References

1. Looger LL, Dwyer MA, Smith JJ, Hellinga HW (2003) Computational design of receptor and sensor proteins with novel functions. *Nature* 423:185–190.
2. Kortemme T, Baker D (2004) Computational design of protein-protein interactions. *Curr Opin Chem Biol* 8: 91–97.
3. Ashworth J, Havranek JJ, Duarte CM, Sussman D, Monnat RJ, Stoddard BL, Baker D (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* 441:656–9.
4. Dahiyat BI, Mayo SL (1997) De novo protein design: fully automated sequence selection. *Science* 278:82–87.
5. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science* 302:1364–1368.
6. Jiang L, Althoff EA, Clemente FR, Doyle L, Rothlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF, Hilvert D, Houk KN, Stoddard BL, Baker D (2008) De novo computational design of retro-aldol enzymes. *Science* 319:1387–1391.
7. Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, Dechance J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, Baker D (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453:190–195.
8. Varadarajan N, Gam J, Olsen MJ, Georgiou G, Iverson BL (2005) Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. *Proc Natl Acad Sci USA* 102:6855–6880.
9. Korkegian A, Black ME, Baker D, Stoddard BL (2005) Computational thermostabilization of an enzyme. *Science* 308:857–860.
10. Johannes TW, Woodyer RD, Zhao H (2005) Directed evolution of a thermostable phosphite dehydrogenase for NAD(P)H regeneration. *Appl Environ Microbiol* 71: 5728–5734.
11. McLachlan MJ, Johannes TW, Zhao H (2008) Further improvement of phosphite dehydrogenase thermostability by saturation mutagenesis. *Biotechnol Bioeng* 99: 268–274.
12. Hasty J, Mcmillen D, Collins JJ (2002) Engineered gene circuits. *Nature* 420:224–230.
13. Vercillo NC, Herald KJ, Fox JM, Der BS, Dattelbaum JD (2007) Analysis of ligand binding to a ribose biosensor using site-directed mutagenesis and fluorescence spectroscopy. *Protein Sci* 16:362–368.
14. Hellinga HW, Marvin JS (1998) Protein engineering and the development of generic biosensors. *Trends Biotechnol* 16:183–189.
15. Maier NM, Franco P, Lindner W (2001) Separation of enantiomers: needs, challenges, perspectives. *J Chromatogr A* 906:3–33.
16. Bolon DN, Mayo SL (2001) Enzyme-like proteins by computational design. *Proc Natl Acad Sci USA* 98: 14274–14279.
17. Zanghellini A, Jiang L, Wollacott AM, Cheng G, Meiler J, Althoff EA, Rothlisberger D, Baker D (2006) New

- algorithms and an in silico benchmark for computational enzyme design. *Protein Sci* 15:2785–2794.
18. Chockalingam K, Chen ZL, Katzenellenbogen JA, Zhao HM (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc Natl Acad Sci USA* 102:5691–5696.
  19. Koh JT (2002) Engineering selectivity and discrimination into ligand-receptor interfaces. *Chem Biol* 9:17–23.
  20. Tang SY, Fazelinia H, Cirino PC (2008) AraC regulatory protein mutants with altered effector specificity. *J Am Chem Soc* 130:5267–5271.
  21. Moore GL, Maranas CD (2004) Computational challenges in combinatorial library design for protein engineering. *AIChE J* 50:262–272.
  22. Marohnic CC, Bewley MC, Barber MJ (2003) Engineering and characterization of a NADPH-utilizing cytochrome b<sub>5</sub> reductase. *Biochemistry* 42:11170–11182.
  23. Woodyer R, Van Der Donk WA, Zhao H (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. *Biochemistry* 42:11604–11614.
  24. Watanabe S, Kodaki T, Makino K (2005) Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. *J Biol Chem* 280:10340–10349.
  25. Kostrzynska M, Sopher CR, Lee H (1998) Mutational analysis of the role of the conserved lysine-270 in the *Pichia stipitis* xylose reductase. *FEMS Microbiol Lett* 159:107–112.
  26. Leitgeb S, Petschacher B, Wilson DK, Nidetzky B (2005) Fine tuning of coenzyme specificity in family 2 aldo-keto reductases revealed by crystal structures of the Lys-274→Arg mutant of *Candida tenuis* xylose reductase (AKR2B5) bound to NAD<sup>+</sup> and NADP<sup>+</sup>. *FEBS Lett* 579:763–767.
  27. Petschacher B, Leitgeb S, Kavanagh KL, Wilson DK, Nidetzky B (2005) The coenzyme specificity of *Candida tenuis* xylose reductase (AKR2B5) explored by site-directed mutagenesis and x-ray crystallography. *Biochem J* 385:75–83.
  28. Liang L, Zhang J, Lin Z (2007) Altering coenzyme specificity of *Pichia stipitis* xylose reductase by the semi-rational approach CASTing. *Microb Cell Fact* 6:36.
  29. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S (2002) Alteration of the specificity of the cofactor-binding pocket of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A. *Protein Eng* 15:131–140.
  30. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S (2002) Optimizing an artificial metabolic pathway: engineering the cofactor specificity of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase for use in vitamin C biosynthesis. *Biochemistry* 41:6226–6236.
  31. Scrutton NS, Berry A, Perham RN (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343:38–43.
  32. Rane MJ, Calvo KC (1997) Reversal of the nucleotide specificity of ketol acid reductoisomerase by site-directed mutagenesis identifies the NADPH binding site. *Arch Biochem Biophys* 338:83–89.
  33. Shiraishi N, Croy C, Kaur J, Campbell WH (1998) Engineering of pyridine nucleotide specificity of nitrate reductase: mutagenesis of recombinant cytochrome b reductase fragment of *Neurospora crassa* NADPH:Nitrate reductase. *Arch Biochem Biophys* 358:104–115.
  34. Eppink MH, Overkamp KM, Schreuder HA, Van Berkel WJ (1999) Switch of coenzyme specificity of *p*-hydroxybenzoate hydroxylase. *J Mol Biol* 292:87–96.
  35. Elmore CL, Porter TD (2002) Modification of the nucleotide cofactor-binding site of cytochrome P-450 reductase to enhance turnover with NADH in Vivo. *J Biol Chem* 277:48960–48964.
  36. Kristan K, Stojan J, Adamski J, Rizner TL (2007) Rational design of novel mutants of fungal 17 beta-hydroxy steroid dehydrogenase. *J Biotechnol* 129:123–130.
  37. Dambe TR, Kuhn AM, Brossette T, Giffhorn F, Scheidig AJ (2006) Crystal structure of NADP(H)-dependent 1,5-anhydro-D-fructose reductase from *Sinorhizobium morelense* at 2.2 Å resolution: construction of a NADH-accepting mutant and its application in rare sugar synthesis. *Biochemistry* 45:10030–10042.
  38. Medina M, Luquita A, Tejero J, Hermoso J, Mayoral T, Sanz-Aparicio J, Grever K, Gomez-Moreno C (2001) Probing the determinants of coenzyme specificity in ferredoxin-NADP<sup>+</sup> reductase by site-directed mutagenesis. *J Biol Chem* 276:11902–11912.
  39. Chen R, Greer A, Dean AM (1995) A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proc Natl Acad Sci USA* 92:11666–11670.
  40. Yaoi T, Miyazaki K, Oshima T, Komukai Y, Go M (1996) Conversion of the coenzyme specificity of isocitrate dehydrogenase by module replacement. *J Biochem (Tokyo)* 119:1014–1018.
  41. Zhang L, Ahvazi B, Szittner R, Vrieling A, Meighen E (1999) Change of nucleotide specificity and enhancement of catalytic efficiency in single point mutants of *Vibrio harveyi* aldehyde dehydrogenase. *Biochemistry* 38:11440–11447.
  42. Holmberg N, Ryde U, Bulow L (1999) Redesign of the coenzyme specificity in L-lactate dehydrogenase from *Bacillus stearothermophilus* using site-directed mutagenesis and media engineering. *Protein Eng* 12:851–856.
  43. Barber MJ (2000) Altered pyridine nucleotide specificity of spinach nitrate reductase. *FASEB J* 14:A1416–A1416.
  44. Miller SP, Lunzer M, Dean AM (2006) Direct demonstration of an adaptive constraint. *Science* 314:458–461.
  45. Bernard N, Johnsen K, Holbrook JJ, Delcour J (1995) D175 discriminates between NADH and NADPH in the coenzyme binding site of *Lactobacillus delbrueckii* subsp. *Bulgaricus* D-lactate dehydrogenase. *Biochem Biophys Res Commun* 208:895–900.
  46. Clermont S, Corbier C, Mely Y, Gerard D, Wonacott A, Branlant G (1993) Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. *Biochemistry* 32:10178–10184.
  47. Ehrensberger AH, Elling RA, Wilson DK (2006) Structure-guided engineering of xylitol dehydrogenase cosubstrate specificity. *Structure* 14:567–575.
  48. Hsieh JY, Liu GY, Chang GG, Hung HC (2006) Determinants of the dual cofactor specificity and substrate cooperativity of the human mitochondrial NAD(P)(+)-dependent malic enzyme-functional roles of glutamine 362. *J Biol Chem* 281:23237–23245.
  49. Serov AE, Popova AS, Fedorchuk VV, Tishkov VI (2002) Engineering of coenzyme specificity of formate dehydrogenase from *Saccharomyces cerevisiae*. *Biochem J* 367:841–847.
  50. Chen R, Greer A, Dean AM (1996) Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase. *Proc Natl Acad Sci USA* 93:12171–12176.
  51. Galkin A, Kulakova L, Ohshima T, Esaki N, Soda K (1997) Construction of a new leucine dehydrogenase with preferred specificity for NADP<sup>+</sup> by site-directed mutagenesis of the strictly NAD<sup>+</sup>-specific enzyme. *Protein Eng* 10:687–690.

52. Saraf MC, Moore GL, Goodey NM, Cao VY, Benkovic SJ, Maranas CD (2006) IPRO: an iterative computational protein library redesign and optimization procedure. *Biophys J* 90:4167–4180.
53. Kavanagh KL, Klimacek M, Nidetzky B, Wilson DK (2003) Structure of xylose reductase bound to NAD<sup>+</sup> and the basis for single and dual co-substrate specificity in family 2 aldo-keto reductases. *Biochem J* 373:319–326.
54. Jez JM, Penning TM (2001) The aldo-keto reductase (AKR) superfamily: an update. *Chem Biol Interact* 130:499–525.
55. Kavanagh KL, Klimacek M, Nidetzky B, Wilson DK (2002) The structure of apo and holo forms of xylose reductase, a dimeric aldo-keto reductase from *Candida tenuis*. *Biochemistry* 41:8785–8795.
56. Kratzer R, Wilson DK, Nidetzky B (2006) Catalytic mechanism and substrate selectivity of aldo-keto reductases: insights from structure-function studies of *Candida tenuis* xylose reductase. *IUBMB Life* 58:499–507.
57. Bohren KM, Grimshaw CE, Lai CJ, Harrison DH, Ringe D, Petsko GA, Gabbay KH (1994) Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme kinetics and crystal structure of the Y48H mutant enzyme. *Biochemistry* 33:2021–2032.
58. Barski OA, Gabbay KH, Grimshaw CE, Bohren KM (1995) Mechanism of human aldehyde reductase: characterization of the active site pocket. *Biochemistry* 34:11264–11275.
59. Schlegel BP, Jez JM, Penning TM (1998) Mutagenesis of 3 alpha-hydroxysteroid dehydrogenase reveals a “push-pull” mechanism for proton transfer in aldo-keto reductases. *Biochemistry* 37:3538–3548.
60. Lee JK, Koo BS, Kim SY (2003) Cloning and characterization of the xyl1 gene, encoding an NADH-preferring xylose reductase from *Candida parapsilosis*, and its functional expression in *Candida tropicalis*. *Appl Environ Microbiol* 69:6179–6188.
61. Werpyp T, Petersen G, Aden A, Bozell J, Holladay J, White J, Manheim A (2004) Top value added chemicals from biomass, volume I: results of screening for potential candidates from sugars and synthesis gas.
62. Cirino PC, Chin JW, Ingram LO (2006) Engineering *Escherichia coli* for xylitol production from glucose-xylose mixtures. *Biotechnol Bioeng* 95:1167–1176.
63. Chin JW, Khankal R, Monroe CA, Maranas CD, Cirino PC (2009) Analysis of NADPH supply during xylitol production by engineered *Escherichia coli*. *Biotechnol Bioeng* 102:209–220.
64. Chenault HK, Whitesides GM (1987) Regeneration of nicotinamide cofactors for use in organic synthesis. *Appl Biochem Biotechnol* 14:147–197.
65. Klein P, Kanehisa M, Delisi C (1984) Prediction of protein function from sequence properties. Discriminant analysis of a data base. *Biochim Biophys Acta* 787:221–226.
66. Cid H, Bunster M, Canales M, Gazitua F (1992) Hydrophobicity and structural classes in proteins. *Protein Eng* 5:373–375.
67. Krigbaum WR, Komoriya A (1979) Local interactions as a structure determinant for protein molecules: II. *Biochim Biophys Acta* 576:204–248.
68. Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 11:739–747.
69. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucl Acids Res* 28:235–242.
70. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201.
71. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucl Acids Res* 31:3381–3385.
72. Petschacher B, Nidetzky B (2005) Engineering *Candida tenuis* xylose reductase for improved utilization of NADH: antagonistic effects of multiple side chain replacements and performance of site-directed mutants under simulated in vivo conditions. *Appl Environ Microbiol* 71:6390–6393.
73. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89:10915–10919.
74. Jonson PH, Petersen SB (2001) A critical view on conservative mutations. *Protein Eng* 14:397–402.
75. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217.
76. Mackerell J, Brooks AD, Brooks B, Nilsson CL, Roux L, Won B, Karplus Y. CHARMM: the energy function and its parameterization with an overview of the program. In: Schleyer PVR, Allinger NL, Ed. (1998) The encyclopedia of computational chemistry. Chichester: Wiley, pp 271–277.
77. Im W, Lee MS, Brooks CL (2003) Generalized born model with a simple smoothing function. *J Comput Chem* 24:1691–702.
78. Born M (1920) Volumes and hydration warmth of ions. *Z Fur Phys* 1:45–48.
79. Lee MS, Salsbury FR, Brooks CL (2002) Novel generalized born methods. *J Chem Phys* 116:10606–10614.
80. Born M (1920) Der aufbau der materie: dreis aufsätze über moderne atomistik und elektronentheorie. Berlin: Springer, 3, 81, 2113.
81. Carugo O, Argos P (1997) NADP-dependent enzymes. I: conserved stereochemistry of cofactor binding. *Proteins* 28:10–28.
82. Chen R, Li L, Weng Z (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins* 52:80–87.
83. Fazelinia H, Cirino PC, Maranas CD (2007) Extending iterative protein redesign and optimization (IPRO) in protein library design for ligand specificity. *Biochem Biophys Res Commun* 357:2120–2130.
84. Tracewell CA, Arnold FH (2009) Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol* 13:3–9.
85. Dunbrack RL, Karplus M (1993) Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *J Mol Biol* 230:543–574.
86. Dunbrack RL, Karplus M (1994) Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nat Struct Biol* 1:334–340.
87. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092.
88. Hacker B, Habenicht A, Kiess M, Mattes R (1999) Xylose utilization: cloning and characterization of the xylose reductase from *Candida tenuis*. *Biol Chem* 380:1395–403.
89. Kang MH, Ni H, Jeffries TW (2003) Molecular characterization of a gene for aldose reductase (CbXYL1) from *Candida boidinii* and its expression in *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* 105:265–276.
90. An YF, Ji JF, Wu WF, Lv A, Huang RB, Wei YT (2005) A rapid and efficient method for multiple-site mutagenesis with a modified overlap extension PCR. *Appl Microbiol Biotechnol* 68:774–778.