JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

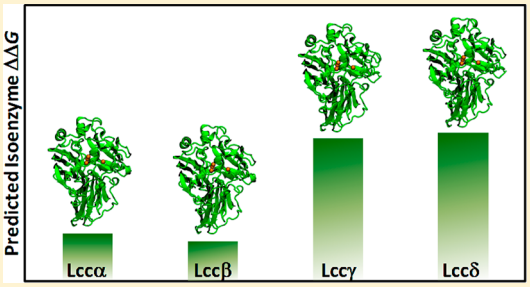
Stability Mechanisms of Laccase Isoforms using a Modified FoldX Protocol Applicable to Widely Different Proteins

Niels J. Christensen and Kasper P. Kepp*

Technical University of Denmark, DTU Chemistry, Kemitorvet 206, DK-2800 Kongens Lyngby, Denmark

Supporting Information

ABSTRACT: A recent computational protocol that accurately predicts and rationalizes protein multisite mutant stabilities has been extended to handle widely different isoforms of laccases. We apply the protocol to four isoenzymes of *Trametes versicolor* laccase (TvL) with variable lengths (498−503 residues) and thermostability (T_{opt} ∼ 45−80 °C) and with 67−77% sequence identity. The extended protocol uses (i) statistical averaging, (ii) a molecular-dynamics-validated "compromise" homology model to minimize bias that causes proteins close in sequence to a structural template to be too stable due to having the benefits of the better sampled template (typically from a crystal structure), (iii) correction for hysteresis that favors the input template to overdestabilize, and (iv) a preparative protocol to provide robust input sequences of equal length. The computed $\Delta\Delta G$ values are in good agreement with the major trends in experimental stabilities; that is, the approach may be applicable for fast estimates of the relative stabilities of proteins with as little as 70% identity, something that is currently extremely challenging. The computed stability changes associated with variations are Gaussian-distributed, in good agreement with experimental distributions of stability effects from mutation. The residues causing the differential stability of the four isoforms are consistent with a range of compiled laccase wild type data, suggesting that we may have identified general drivers of laccase stability. Several sites near Cu, notably 79, 241, and 245, or near substrate, mainly 265, are identified that contribute to stability-function trade-offs, of relevance to the search for new proficient and stable variants of these important industrial enzymes.



INTRODUCTION

Understanding the drivers of protein stability is one of the most important challenges of biochemistry, closely related to the protein-folding problem,[1−5] and is also of enormous importance in biotechnology, where increased protein stability vastly improves the usefulness of many industrial enzymes.[6−9]

Stability mechanisms are not universal and must be explored for each protein class separately,[3,4] and they also depend on temperature (T), pH, solvent properties, and ionic strength (IS).[1,4] Properties known to often correlate with enhanced stability include internal electrostatics,[10] hydrophobic packing of the protein,[1,5,11] optimized packing of loops and turns,[12,13] and post-translational modifications of the protein surface.[14,15] The challenge of obtaining both highly stable and proficient enzymes for industrial purposes is further complicated by common stability-function trade-offs during mutation.[16] Understanding when and how this trade-off can be circumvented would be of substantial value to the protein engineering field.

Rational methods may help us to understand protein stability drivers and to discover new, highly stable and proficient proteins to be produced in the laboratory.[17,18] Rational prediction of stability based on free energy decomposition is extremely challenging because interactions are coupled, although early pessimistic conclusions[19] have been later partly remedied, both via decomposition theory[20,21] and from the results produced by such methods.[22] Most importantly, modern protein stability predictors can now use empirically fitted free energy components to produce total free energies with errors of ∼1 kcal/mol for single-site mutants.[23,24]

For such approaches to be of practical interest, it is our opinion that they should predict with few false positives stability rankings of widely different proteins far beyond single-point mutations: Prediction of single-site mutant stabilities is of limited use considering the currently available, fast experimental mutagenesis procedures. Instead, less homologous proteins constitute almost all of the relevant protein space but are inaccessible by experimental techniques due to the prohibitive number of generations needed to produce them or due to dead ends separating them from the starting wild type, preventing consecutive rounds of mutations from producing the desired, optimal proteins.

Unfortunately, current state-of-the-art protein stability predictors are of limited use outside their parametrization range, even for single-site mutations.[23,25] However, we have recently shown[26] that one of the most accurate single-mutation stability predictors, FoldX,[24] can be modified to handle also multisite mutants with up to 11 mutated sites for complex proteins (tested on laccases) outside the parametrization range of the predictor, when a number of issues are solved, notably

Received: March 18, 2013
Published: June 11, 2013

   dx.doi.org/10.1021/ct4002152 | J. Chem. Theory Comput. 2013, 9, 3210−3223

Table 1. Data for Four Studied Laccase Isozymes (Lccα, Lccβ, Lccγ, and Lccδ)[a]

isozyme name	protein length	GenBank No.	gene name	charge pH 5/7	related structure	k_{cat} (s^{-1})	K_M (μM)	pH_{opt}	T_{opt} (°C)
Lccα	499	AY693776	α[39]	−6/−4	1GYC (99% ID)	55	22	1.9	75
Lccβ	499	Y18012	β	−19/−14	1KYA (98% ID)	203	88	2.3	80
Lccγ	503	D84235	CVLG1[40]	−31/−30	n/a	647	359	3.1	60
Lccδ	498	X84683	lcc1[41]	−35/−30	n/a	798	2262	3.1	45

[a]Characterization by ABTS Assays. Data from ref 27. K_M and k_{cat} are reported at pH = 5.

template bias, hysteresis, statistical noise, and the quality of the structural template used as input for the computations.

Here, we investigate if such protocols can in simple, reproducible ways be further modified to handle less homologous proteins of potential value to protein engineering, by focusing on four laccase isoforms from the fungus *Trametes versicolor*.[27] Laccases are multifunctional, often highly stable and complex multidomain multicopper oxidases[28−30] with exquisite oxidation capabilities,[31−33] used for a variety of purposes such as degradation of lignin,[34,35] bioremediation,[29,36] and oxidative bleaching.[32,34,37,38] Their complexity, importance, and high stabilities render them particularly suitable as test cases for stability predictors. Furthermore, some fungi have several (typical 2−5) isoforms that vary substantially in terms of function and stability,[38] providing a favorable starting point for understanding stability-function trade-offs for a single organism with one life style, physiology, and habitat.

A phylogenetic analysis[39] of the reported sequences of *Trametes versicolor* laccases (TvL) produced four isoenzyme groups termed α, β, γ, and δ with genes within each group sharing at least 97% sequence identity. Later,[27] a gene (*lcc*α,[39] *lcc*β, *lcc*γ,[40] and *lcc*δ[41]) from each group was heterologously expressed in *Pichia pastori* to yield the isoenzymes Lccα, Lccβ, Lccγ, and Lccδ. Some TvLs are known to have very high redox potentials, of the order of ∼800 mV.[30] Biochemical characterization classified the isoenzymes into two groups: Lccα and Lccβ had lower activity toward 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) but higher stability (both at low pH and elevated T) compared to Lccγ and Lccδ.[27] Within the two groups, Lccβ was more stable/less active than Lccα, and Lccγ was more stable/less active than Lccδ, although these differences were less pronounced than the differences between the groups. Glycosylation coverage was similar in the isoenzymes and could thus not explain the differences in isoenzyme properties. These data thus constitute a prominent example of stability-function trade-off of particular interest in our ongoing efforts to predict and analyze laccase (and, more generally, complex protein) stabilities.

Compared to previous computational studies of relative protein stability, the present study poses two new challenges of substantial sequence variation (sequence identity variation of 67−77%) and, importantly, different sequence lengths (from 498 to 503 residues). We show that a modified FoldX protocol, when corrected for sequence length and critical issues relating to structural template bias and sampling, provides robust rankings of the stabilities of the four isoforms and gives significant insight into drivers of their stability differences, of potential use in experimental optimization of these important proteins.

COMPUTATIONAL DETAILS

Laccase Isoenzymes. The data set for the four TvL isoenzymes[27] (Lccα, Lccβ, Lccγ, and Lccδ) including available sequences and stability/activity data are compiled in Table 1. Charges were calculated with the PROPKA software.[42] Sequence identities between the four laccase isoenzymes and selected homologous crystal structures are shown in Table 2.

Table 2. Sequence Identity (%) Matrix for Lccα, Lccβ, Lccγ, and Lccδ and Homologous Crystal Structures (PDB-ID: 3FPX, 1GYC, 1KYA) Used as FoldX Templates

	Lccβ	Lccγ	Lccδ	3FPX	1GYC	1KYA
Lccα	79	70	69	79	99	80
Lccβ		72	71	88	80	98
Lccγ			77	72	71	71
Lccδ				70	70	71
3FPX					80	88
1GYC						80

Strategy for Computing Relative Free Energies of Folding. The relative stability of a computed "mutant" protein relative to a reference protein, often a wild type (WT), is defined as

$$\begin{aligned}\Delta\Delta G &= \Delta G_M - \Delta G_{WT}\\ &= -RT\ln(K_M/K_{WT})\\ &= -RT\ln([F]_M[U]_{WT}/[U]_M[F]_{WT})\end{aligned}$$

ΔG_M and ΔG_{WT} are the free energies of folding for the "mutant" and reference "wild type" proteins, K_M and K_{WT} are the effective two-state folding equilibrium constants, and $[F]$ or $[U]$ denote concentrations of folded (F) or unfolded (U) protein at equilibrium. Experimental standard errors of $\Delta\Delta G$ are of the order of ∼0.5 kcal/mol.[43]

Given a 3D structure template and a series of amino acid sequences, FoldX constructs 3D models for the sequences and predicts $\Delta\Delta G$ relative to an internally adjusted version of the input structure, requiring the same length for all sequences. As some of the four TvL isoenzymes vary in sequence length, a protocol to handle this difference is necessary before comparing more diverse proteins. However, Lccα and Lccβ do have the same length (499 residues) and can thus also be used directly with the standard FoldX program.

We therefore considered two cases separately: First, we used FoldX to predict $\Delta\Delta G$ for Lccα and Lccβ using a 499 residue homologous crystal structure as structural input. In the second case, we provided a protocol to enable same-length sequence comparison of all four isoforms. The later protocol requires the generation of a homology (threading) model for structural input (*vide infra*). Several approaches and conditions were evaluated to enable a realistic comparison of the isoforms and test sensitivity to changes in procedure and parameters, and the stabilities were in all cases compared to the experimental optimal-activity temperatures (T_{opt}).

Computing $\Delta\Delta G$ for Lccα and Lccβ. The sequences of Lccα and Lccβ were obtained from GenBank[44] using the

accession numbers in Table 1. Searching the protein data bank[45,46] for a suitable 499-residue crystal structure revealed three candidates with reasonable homology (Table 2): the 1.90 Å *Trametes versicolor* α-laccase (PDB-ID: 1GYC),[47] the 2.40 Å *Trametes versicolor* β-laccase (PDB-ID: 1KYA),[48] and the 1.80 Å *Trametes hirsuta* laccase (PDB-ID: 3FPX).[49] These structures were all investigated as input templates due to the previously described structural sensitivity of FoldX, notably the energy bias toward the input structure.[26]

As a structural template for FoldX predictions, the 3FPX structure has advantages compared to 1GYC and 1KYA: First, 3FPX has a better resolution (1.80 Å) than 1GYC (1.90 Å) and 1KYA (2.40 Å). Second and more importantly, the maximum sequence identity between the TvL isoenzymes and 3FPX is 88% (Lccβ) compared to 98% between 1KYA and Lccβ and 99% between 1GYC and Lccα (see Table 2). Thus compared to 1GYC and 1KYA, 3FPX assumes a more balanced position in sequence space with respect to the isoenzymes, reducing the initial bias toward any one protein. Accordingly, FoldX predictions from 3FPX provided much more realistic $\Delta\Delta G$ values for Lccα and Lccβ whereas other choices remained biased (Supporting Information, Table S1). Thus, to handle the template bias, 3FPX was used as template for the calculations.

Computing $\Delta\Delta G$ for Lccα, Lccβ, Lccγ, and Lccδ. To compute $\Delta\Delta G$ consistently for all four isoenzymes (Lccα, Lccβ, Lccγ, and Lccδ), it was required that the sequences were modified to have the same length as a common, unbiased structural template used for computation. Thus, we first performed a structural alignment and homology modeling. The crystal structure used as template (3FPX) was loaded into VMD,[50] and four homology models for Lccα, Lccβ, Lccγ, and Lccδ were obtained from the ModBase database.[51] The crystal structure and homology models were structurally aligned using the STAMP[52] alignment tool of the MultiSeq[53] plugin. Except for one insertion in a β-sheet, the variation in sequence length was due to variable loop lengths; that is, sequence editing could be restricted to mainly loosely structured parts of the proteins.

Next, guided by the sequence alignment (Supporting Information, Figure S1) resulting from structural alignment, we produced equal sequence lengths using two protocols: Protocol I deleted all residues of a sequence that were insertions relative to at least one other protein leading to 495 residues in the new sequences. Protocol II differed from Protocol I by omitting deletions potentially affecting secondary structure. Such problematic residues were in Protocol II substituted with Gly across all sequences in order to maintain backbone length, leading to 496 residues in the new sequences. With reference to Figure S1, Supporting Information, Protocol I and II deleted the loop residues 431−433 in 3FPX, Lccα, and Lccβ, 184−185, 391−395, and 437 in Lccγ, and 390 and 432 in Lccδ. In addition, Protocol I deleted the β-sheet residue at position 379 in 3FPX, Lccα, Lccβ, and Lccδ, whereas Protocol II inserted a glycine at this position in all five sequences.

Whereas the solvent-exposed insertions were expected to have limited structural impact, the β-sheet insertion might affect the secondary or tertiary structure. This was tested by keeping the β-sheet backbone intact in Protocol II and comparing the two results (see the Supporting Information, Figure S2). The 495- and 496-residue sequences (and structures, *vide infra*) obtained from Protocol I and Protocol II, respectively, are in the following designated by the original sequence/structure identifier followed by a prime: Lccα', Lccβ', Lccγ', and Lccδ' refer to the modified same-length TvL isoenzyme sequences and 3FPX′ refers to the modified sequence of the crystal structure template.

To generate an input structure for $\Delta\Delta G$ computation of Lccα', Lccβ', Lccγ', and Lccδ', the 3FPX′ sequences obtained from Protocols I and II were threaded into the 3FPX crystal structure using Phyre2.[54] The resulting 3FPX′ models were then imported into Maestro[55] and treated with the Protein Preparation Wizard to assign bond orders and establish disulfide bridges. Cu atoms and coordinating oxygens were inserted into the structures using the 3FPX crystal structure coordinates. A charge of +1 was assigned to all Cu atoms and oxygen ligands to Cu at the T2 and T3 sites were modeled as water molecules. To allow comparison with crystal structures, the numbering of amino acid residues in the modified sequences and structures follows 3FPX.pdb.

FoldX Computations. The relative free energies of folding, $\Delta\Delta G$, were predicted with FoldX version 3b51[24] using the ⟨BuildModel⟩ command to mutate an input structure (template) into target structures (isoenzyme models) specified as amino acid sequences. As in our previous study,[26] we repeated each $\Delta\Delta G$ computation five times to improve sampling by inserting five copies of each sequence in the FoldX sequence input text file (mutant_file.txt).

First, default FoldX parameters were used for temperature (298 K), ionic strength (0.05 M), and pH (7), and subsequently, the sensitivity of results to these parameters were tested by computing also the free energies for conditions resembling more the experimental assay (T = 350 K, 0.2 M, pH = 5). Previously,[26] we observed that FoldX, when applied to complex proteins outside the parametrization range, over-stabilized the input structure, and we found that the average $\Delta\Delta G$ obtained from successive forward/backward mutations substantially reduced this template bias to provide more realistic and consistent free energies in good agreement with experimental stabilities for a range of multisite mutants. Consequently, the FoldX calculations herein were performed with both forward and backward mutation, and $\Delta\Delta G$ has been reported in tables for both directions and as the numerical average for a forward−backward cycle (we call this free energy estimate the "hysteresis-corrected" $\Delta\Delta G$).

To obtain the contribution $\Delta\Delta G_i$ of each amino acid to the total $\Delta\Delta G$ and to the individual energy terms, we performed FoldX stability calculations for the mutant structures of interest and for the corresponding WT structure using the ⟨Stability⟩ and ⟨SequenceDetail⟩ keywords. $\Delta\Delta G$ and the component energy terms were calculated by subtracting the energy terms of the WT from those of the mutant. To improve sampling, all per-residue FoldX calculations were replicated five times both forward and in reverse (once for each of the FoldX generated structures in the nonresidue resolved runs). Average values across these runs are reported in all cases. The difference between the largest and the smallest per-residue $\Delta\Delta G$-contribution across the four Lcc′ sequences ($\Delta\Delta G_{i\text{max-min}}$) was evaluated at each of the 496 sequence positions.

Molecular Dynamics Simulations. The *Trametes hirsuta* crystal structure[49] (PDB-ID: 3FPX) and the homology model 3FPX′ from Protocol II were prepared for MD simulations using the protein preparation wizard of Maestro.[55,56] Each structure was prepared in two different protonation states assigned with PROPKA[42] corresponding to pH = 5 as is typical for *Trametes hirsuta* laccase oxidation assays[57] and pH = 7 as employed in the crystallization of 3FPX.[49] The proteins were solvated in a 15-Å cubic TIP3P box with charge-neutralizing

counterions (Na^+) in addition to an ionic background of 0.3 M NaCl, followed by Desmond[58] steepest descent minimization to a gradient of 1 kcal mol^{-1} $Å^{-1}$. System details are summarized in Table 3. Additional minimization with the default presimulation protocol in Desmond consisted of (1) minimization with solute restraints, (2) minimization without restraints, (3) Berendsen[59] NVT simulation at 10 K with small time steps and restraints on heavy solute atoms, (4) Berendsen NPT simulation at T = 10 K with restraints on heavy solute atoms, (5) Berendsen NPT simulation with restraints on heavy solute atoms, and (6) unrestrained Berendsen NPT simulation.

Table 3. System Details for MD Simulations of *Trametes hirsuta* laccase (PDB-ID: 3FPX) and the Homology Model 3FPX′

	pH	no. TIP3P	no. atoms	protein charge	no. Na^+, no. Cl^-
3FPX	5	29404	95966	−11	180, 169 (0.32 M NaCl)
	7	29370	95864	−16	185, 169 (0.32 M NaCl)
3FPX′	5	29417	95967	−11	180, 169 (0.32 M NaCl)
	7	29383	95865	−16	185, 169 (0.32 M NaCl)

Production-run NPT simulations with periodic boundary conditions were carried out for each system for 20 ns, using the Nose−Hoover[60,61] chain thermostat and the Martyna−Tobias−Klein[62] barostat with isotropic coupling and a relaxation time of 2.0 ps. Equations of motion were integrated using the RESPA[63] integrator with bonded, near, and far time steps of 2.0 fs, 2.0 fs, and 6.0 fs, respectively. A cutoff of 9 Å was used for nonbonded interactions while long-range electrostatics was treated with the smooth-particle mesh Ewald method[64] with a tolerance of 10^{-9}. The MD trajectory was saved to disk at 10 ps intervals. Backbone RMSD curves were calculated in VMD[50] relative to the 3FPX crystal structure and 3FPX′ homology model. After structural alignment, root-mean-square fluctuations (RMSF) were calculated for the residues in 3FPX and 3FPX′ using the converged last 10 ns of the trajectories.

The OPLS-2005[65,66] force field was employed to describe the protein. Water and sodium and chloride ions were described with the TIP3P[67] potential and the free-energy consistent alkali-and halide-ion OPLS potentials,[68] respectively. Using a realistic ionic strength is important as physiological ions (notably Cl^-) have a stabilizing effect on the protein backbone, viz. the Hofmeister salting-out effect that changes T_m by typically tens of degrees[69,70] and change the conformational preference between protein states[71,72] (i.e., MD-simulations without realistic ionic strengths will substantially underestimate stability and folding propensity). Thus, realistic, self-consistent ion potentials, that is, those developed for OPLS, are necessary to describe the balance between the tendencies of peptide side chains and ions to hydrate and mutually interact, respectively.[73]

Sulfate, carbohydrates at N-glycosylation sites, and crystal waters more than 5 Å from hetero atoms were deleted followed by addition of hydrogen atoms. The 3FPX structure contains three oxygen species coordinated to the T2 copper, between the two T3 coppers, and between the T2 copper and one T3 copper. While the former two species are also found in the *Trametes versicolor* laccase crystal structures 1GYC[47] and 1KYA,[48] the third species may represent an intermediate oxidation state of the laccase. Thus, the two former were modeled as water bound to Cu whereas the third species was removed. Consistent with our previous study,[26] angles and bond distances involving all copper atoms were fixed as in the crystal structure, and a charge of +1 was assigned to copper.

RESULTS AND DISCUSSION

Stabilities of Isoforms Lccα and Lccβ using Non-modified Sequences. The *Trametes hirsuta* laccase crystal structure (PDB-ID: 3FPX) was selected as FoldX structural input, as discussed in the Computational Details section. The structures of the highly homologous *Trametes versicolor* α-laccase[47] and β-laccase[48] displayed substantial template bias toward their homologous counter parts, confirming the previously identified[26] template bias. First, we investigated the properties of Lccα and Lccβ using our protocol described previously,[26] that is, without sequence length modifications and homology modeling.

ΔΔG was computed for Lccα and Lccβ from the 3FPX crystal structure template as described in the Computational Details section. As a test of the sensitivity of the method to the type of local sampling, we used both FoldX-repaired and nonrepaired structures. The first generation of mutations, named 3FPX.pdb → Lccα and 3FPX.pdb → Lccβ in Table 4,

Table 4. Computed ΔΔG (kcal/mol) for Lccα and Lccβ Using 3FPX as Structural Template

	3FPX.pdb → Lccα	Lccα→ Lccβ	Lccβ → Lccα	ΔΔG[a]
no repair	17.3	6.5	2.3	2.1
repair	21.6	6.5	3.2	1.6
	3FPX.pdb → Lccβ	Lccβ → Lccα	Lccα → Lccβ	ΔΔG
no repair	18.6	6.1	5.1	0.5
repair	23.8	6.7	4.7	1.0

[a]The value of Lccα → Lccβ minus the value of Lccβ → Lccα, divided by 2.

yielded large positive (destabilizing) values of ΔΔG of 17.3 and 18.6 kcal/mol for Lccα and Lccβ, respectively. These unrealistic destabilizations confirm a strong bias toward the template structure, and the repair protocol produces even more bias of 4−5 kcal/mol, consistent with previous observations.[26] Despite this bias, the similar ΔΔGs for Lccα and Lccβ within ∼2 kcal/mol indicate that this simple FoldX protocol recovers similar stabilities for the two isoenzymes, in agreement with the T_{opt} values of 75 °C and 80 °C. However, given the issues of such a calculation, this agreement could be fortuitous.

To investigate this further, we computed a new set of free energies, now using the first FoldX-generated structures of Lccα and Lccβ for mutation, both Lccα → Lccβ and Lccβ → Lccα. This produced ΔΔG values of 6.5 and 6.1 kcal/mol, respectively. Notably, computing these free energies directly from corresponding crystal structures 1GYC and 1KYA produced larger template biases: Mutating 1GYC to Lccα and Lccβ yielded ΔΔGs of 3.1 and 40.3 kcal/mol, respectively, while the same mutants generated from 1KYA resulted in ΔΔGs of 8.3 and 5.7 kcal/mol, respectively (see the Supporting Information, Table S1). Instead, using the 3FPX crystal structure template as a compromise reference state, bias was reduced and free energies were in a more realistic range. The substantially positive ΔΔG values reflect remaining bias toward the respective templates, but the biases are of similar magnitudes. The difference between using the repair protocol

or not is accordingly also reduced as overall template bias is reduced. When repeating this calculation one more time by computing Lccβ → Lccα and Lccα → Lccβ using the new templates, free energies were reduced from 6−8 kcal/mol to the more realistic 2−5 kcal/mol. In other words, bias was further reduced by a reverse mutation run, in agreement with the concept of "hysteresis" in the thermodynamic cycle.

Our previous[26] protocol for estimating ΔΔG was to use the average of the free energies from forward and backward mutations to eliminate hysteresis, in analogy with free energy perturbation techniques. Using this procedure gives free energies that favor Lccα by ∼2 kcal/mol in the calculations starting from 3FPX.pdb → Lccα, and ∼1 kcal/mol in favor of Lccβ when this protein was the first mutated. A final, template-bias, sampling, and hysteresis-corrected estimate is thus that Lccα is ∼1 kcal/mol more stable than Lccβ. As explained above, a standard FoldX protocol (single run mutation from a structurally repaired model) would never have produced such a reasonable result: Mutation from the 1GYC and 1KYA templates predicted ΔΔG differences between Lccα and Lccβ of 42.3 and 5.8 kcal/mol, respectively (Supporting Information, Table S6), with a number of errors prevailing. The estimate from the improved protocol includes statistical averaging over five runs. Considering the standard deviations of such repetitions (e.g., ∼1.4 kcal/mol for 3FPX → Lccα and ∼0.6 kcal/mol for for 3FPX → Lccβ) the five-run averages are more robust free-energy estimates than single-run results.

Computing ΔΔG for Lccα′, Lccβ′, Lccγ′, and Lccδ′. After having discussed some issues with the calculation of stabilities using FoldX, we now turn to the computation of ΔΔG for all four isoforms using our modified protocol. As mentioned, FoldX requires similar lengths of mutant sequences and the structural template. We tested two deletion protocols (Protocol I and II, described in the Computational Details section) to circumvent this problem. Up to 10 total deletions were required to produce equal sequence lengths. Furthermore, all but one of the sites occurred in solvent-exposed loops, thus minimizing secondary structure disruption due to the sequence deletions.

Applying both protocols followed by FoldX calculations (five-run averages, without hysteresis) showed that Protocol II overall produced lower energies than Protocol I and correctly ranked ΔΔG for Lccγ and Lccδ (Supporting Information, Figure S3). As shown below, this protocol also produced encouraging and useful results for the whole data set, mainly because the modification has small effect relative to the large total sequence variation between the four isoforms. The results in the remainder of this article are thus based on analysis of computed data resulting from structures produced from the homology models after sequence-editing according to Protocol II.

Using these prepared structures, stability calculations are reported in Table 5 as five-averaged results from forward (3FPX′ → Lcci′), backward (Lcci′ → 3FPX′), and forward−backward−forward (3FPX′ → Lcci′) at FoldX default conditions (0.05 M, pH 7, 298 K, top four rows) and at conditions resembling the optimal protein conditions (0.2M, pH 5, 350 K, bottom four rows). Note that these values are relative to the sequence-length-corrected and optimized 3FPX′ homology model, which is a hypothetical protein; that is, the absolute numbers are meaningless, only the relative numbers should be considered.

Table 5. Computed ΔΔG (kcal/mol) for *Trametes versicolor* Laccase Isoenzymes Lccα′, Lccβ′, Lccγ′, and Lccδ′ Showing Forward (3FPX′ → Lcci′), Backward (Lcci′→ 3FPX′), and Forward−Backward−Forward (3FPX′→ Lcci′) Mutations at 0.05 M, pH 7, 298 K and 0.2M, pH 5, 350 K[a]

i	3FPX′ → Lcci′	Lcci′ → 3FPX′	3FPX′ → Lcci′	ΔΔG[b]
		0.05 M, pH 7, 298 K		
α	21.1	−11.7	19.5	16.4
β	21.6	−15.0	19.6	18.3
γ	61.0	−48.9	60.5	55.0
δ	66.1	−47.9	57.9	57.0
		0.2 M, pH 5, 350 K		
α	24.0	−13.8	22.1	18.9
β	19.0	−12.6	17.1	15.8
γ	63.4	−50.4	62.2	56.9
δ	66.9	−51.2	64.0	59.1

[a]All reported values are averages of five subsequent runs. [b]The value of 3FPX′→ Lcci′, minus the value of Lcci′→ 3FPX′ divided by 2. Results from calculations with the repair protocol are reported in Table S2 in the Supporting Information.

Under both conditions, the relative stabilities (ΔΔG) for the isoenzymes are ordered as Lccα ∼ Lccβ ≫ Lccγ ∼ Lccδ, in agreement with the experimental observation of two classes based on properties and phylogeny (Lccα, Lccβ) and (Lccγ, Lccδ).[27] If the 3FPX′ homology model had been subject to a standard FoldX-repair protocol (i.e., application of repair to the structural template, but no repetition- or hysteresis-averaging of the predictions), the results would have been substantially less convincing (see Supporting Information, Table S2), consistent with our previous observations[26] that a modified protocol using either crystal structures with bias- and hysteresis correction or MD-simulated structural input provides markedly higher correlation with experimental data than the standard FoldX protocol.

The results confirm that the computational setup provides results in good qualitative agreement with experimental stabilities. The two experimental groups of stability across 35 °C of T_{opt} are well reproduced. Importantly, this performance may be markedly improved (to possibly 1 °C error) by quantitative structure−property relationships (QSPR) using the FoldX energy terms.[26] Thus, although the protocol is qualitative or at best semiquantitative and currently only applicable to laccases, it constitutes a promising approach to fast computational screening of approximate stability trends for proteins of limited homology, much smaller than that of single-site or simple multisite mutants usually applicable to stability prediction. Notably, our protocol circumvents the problem that free energy errors (and errors in stability ranking) for standard predictors increase monotonously with the number of simultaneous mutations studied, allowing assessment of larger parts of sequence space.

Molecular Dynamics Simulations of 3FPX and 3FPX′. Previously,[26] MD snapshots were shown to provide extremely variable and essentially useless properties, including free energies from FoldX, whereas crystal structure accuracy was recovered when FoldX free energies were averaged over ∼50 snapshots well-separated (to avoid strong time-correlation) over the equilibrated part of a trajectory that remained in the crystal structure native conformation (RMSD ∼1−2 Å).[26] This conclusion most likely relates to other MD-derived properties as well. In the present work, the homology model 3FPX′ was

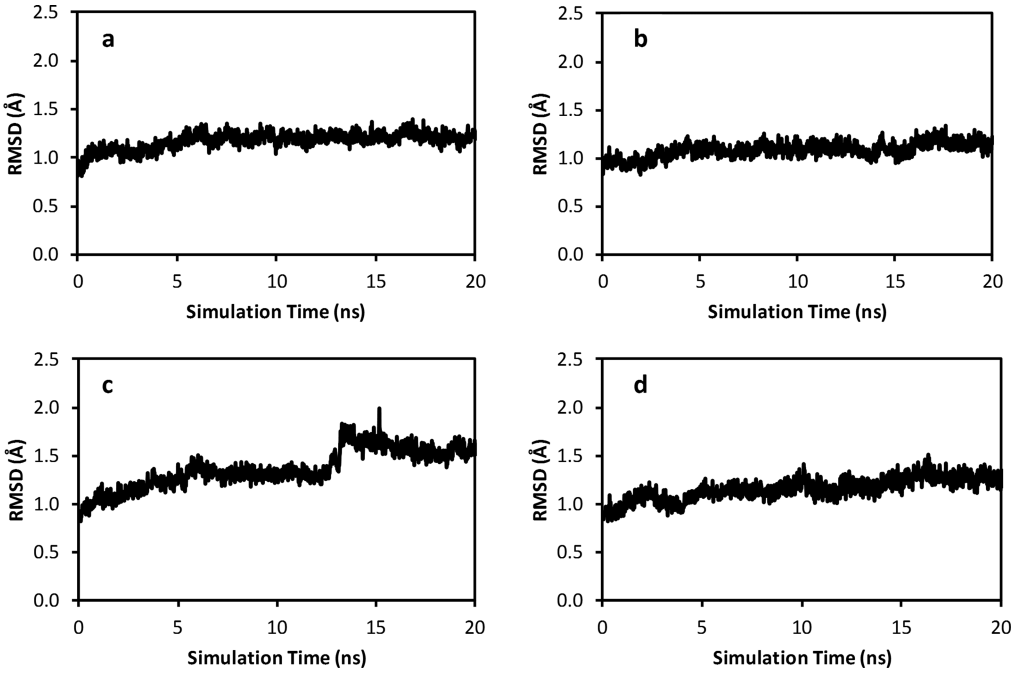


Figure 1. Backbone RMSD curves for 20 ns NPT simulations. (a) 3FPX at pH 7, (b) 3FPX′ at pH 7, (c) 3FPX at pH 5, and (d) 3FPX′ at pH 5.

threaded and does not suffer from the same inaccuracy; that is, it preserves the essential time-averaged crystal structure quality (*vide supra*). Thus, MD is not required for structural averaging but may validate that the unbiased high-quality sequence-modified homology model 3FPX′ is structurally and dynamically similar to the 3FPX crystal structure. Furthermore, MD, when used with realistic ionic strength, is a sensitive tool for probing structural perturbations of fungal laccases[74] and may also rationalize dynamic effects not available on the crystal diffraction time scale.

Backbone RMSD plots for the four simulations (3FPX and 3FPX′ at pH 5 and pH 7) are shown in Figure 1. At pH 7, the 3FPX and 3FPX′ RMSD plots (Figure 1, a and b, respectively) are well-behaved and indicate structural equilibration after ∼5 ns with equilibrium RMSDs (averaged over the last 10 ns) of ∼1.2 and ∼1.1 Å, respectively. At pH = 5 the RMSD plot for the 3FPX′s simulation (Figure 1d) is also well-behaved with an equilibrium RMSD of ∼1.2 Å whereas the RMSD for the 3FPX simulation (Figure 1c) displays a notable transition from ∼1.3 to ∼1.7 Å at ∼13 ns. Such transitions are common in simulations, in particular in zero-ionic strength environment, high *T*, or low pH, and reflect departure from the initially sampled conformational space (statistics using several such states are meaningless). The MD simulations thus show that the modification of the 3FPX crystal structure using our Protocol II to produce the 3FPX′ homology model preserves the conformational ensemble close to the crystal structure on the 20-ns time scale; that is, the 3FPX′ model is dynamically robust.

Physico-chemical Causes of Differential Stability: ΔΔ*G* Components for Lcc*α*′, Lcc*β*′, Lcc*γ*′, and Lcc*δ*′. After having showed that the described protocol provides robust, semiquantitative results for the relative stabilities of the laccase isoforms, we now turn to the more biologically relevant question of what causes these major differences in stability observed experimentally.

The hysteresis-averaged FoldX energy terms comprising the ΔΔ*G* values in the last column of Table 5 are listed in Table 6.

Table 6. Computed ΔΔ*G* (kcal/mol) and Energy Components (FoldX Acronyms in Parentheses) Lcc*α*′, Lcc*β*′, Lcc*γ*′, and Lcc*δ*′, and $E_{max\text{-}min}$—the Maximum Difference for an Energy Term Across the Four Isoenzymes

energy term	Lcc*α*′	Lcc*β*′	Lcc*γ*′	Lcc*δ*′	$E_{max\text{-}min}$
backbone hydrogen bonds (BackHbond)	3.6	2.4	9.3	7.0	6.9
side-chain hydrogen bonds (SideHbond)	12.9	7.7	11.8	4.1	8.8
van der Waals energy (energy-vdw)	−3.3	4.1	7.6	2.8	10.9
electrostatic energy (Electro)	−0.9	1.2	3.0	2.5	3.9
polar solvation enery (energy_SolvP)	1.6	−4.8	−5.7	0.1	7.3
hydrophobic solvation energy (energy_SolvH)	−7.5	5.0	6.8	4.7	14.3
van der Waals repulsion (vdW clash)	6.8	3.5	17.1	20.9	17.4
side-chain entropy (entrop_sc)	−0.6	−4.8	−7.4	−3.6	6.8
backbone entropy (entropy_mc)	1.8	2.7	8.1	13.5	11.7
intraresidue torsional repulsion (energy_torsion)	2.3	2.4	5.2	6.6	4.3
helix dipole moment (energy_dipole)	−0.4	−0.5	−0.2	−0.6	0.4
energy of ionization (energy ionisation)	0.1	−0.5	−0.8	−0.9	1.0

The five terms differing the most between any two isoenzyme models are the total energy of inter-residue van der Waals repulsion (17.4 kcal/mol higher for Lcc*δ*′ than for Lcc*β*′), the hydrophobic solvation term (14.3 kcal/mol separating Lcc*γ*′ and Lcc*α*′), the backbone entropy (11.7 kcal/mol separating Lcc*δ*′ and Lcc*β*′), the total van der Waals energy (10.9 kcal/mol difference between Lcc*γ*′ and Lcc*α*′), and the side-chain hydrogen bond energies (8.8 kcal/mol difference between

dx.doi.org/10.1021/ct4002152 | J. Chem. Theory Comput. 2013, 9, 3210−3223

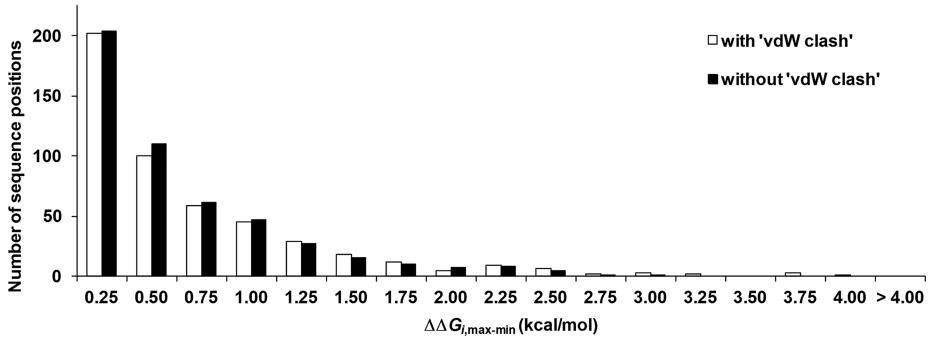


Figure 2. Distributions of maximum free energy variations $\Delta\Delta G_{i,\text{max-min}}$ across all four isozymes. The differences were evaluated with (white) and without (black) the FoldX "vdW clash" term.

Lccα' and Lccδ'). For the remaining terms, the maximum difference between two isoenzyme models is between 0.4 and 7.3 kcal/mol.

Table 6 shows that some FoldX energy components contribute substantially more than others to the differential stability of the (Lccα', Lccβ') and (Lccγ', Lccδ') isoenzyme pairs: van der Waals interactions are responsible for more than 10 kcal/mol difference, backbone entropy for 5−10 kcal/mol, roughly, and backbone hydrogen bonds for ∼5 kcal/mol. The electrostatic energy is also significantly larger for Lccγ' (3.0 kcal/mol) and Lccδ' (2.5 kcal/mol) compared to Lccα' (−0.9 kcal/mol) and Lccβ' (1.2 kcal/mol). These values anticorrelate with the PROPKA charges calculated for the ModBase models (Table 1) at pH 5 ($r = -0.99$, $R^2 = 0.98$) and pH 7 ($r = -0.96$, $R^2 = 0.94$), and we observe anticorrelation ($r = -0.68$, $R^2 = 0.46$) between the FoldX electrostatic energy and T_{opt}, suggesting that electrostatics may correlate well with stability of the laccases.

The van der Waals repulsion term ("vdW clash") may represent physically relevant unfavorable interactions as well as errors inherent to the geometry of the homology models due to the large sequence variation going from the 3FPX′ structural template and the modified TvL isoenzymes. As discussed in Supporting Information (Effect of van der Waals Repulsion section and Figure S3), neglecting this term did not destroy the correlation with experimental data. Thus, the following discussion is based on the ΔΔG excluding the "vdW clash" term from consideration.

Residue Contributions to ΔΔG. As explained in the Computational Details section, at each residue position i, we evaluated the difference $\Delta\Delta G_{i,\text{max-min}}$ between the maximum and minimum residue contribution to ΔΔG across Lccα', Lccβ', Lccγ', and Lccδ'. The histograms in Figure 2 show that $\Delta\Delta G_{i,\text{max-min}} > 1$ kcal/mol for ∼100, or about 20% of the sites in the laccases. Conversely, for ∼80% of the residues, $\Delta\Delta G_{i,\text{max-min}}$ is <1 kcal/mol. These results are in good agreement with the empirically known skewed Gaussian distribution of stability effects of single-site mutations across proteins, with a spread of ∼1 kcal/mol.[16] Thus, in the studied fungal laccases, a large part of the total energy difference is caused by only a few sites whereas the majority of sites are nearly "neutral" in terms of stability.

Residues near Copper Sites. To understand the substantial differences in stability of the four laccase isoforms, we started by inspecting sequence changes occurring near the functionally important copper sites, where electrons are transferred from substrate (at the T1 site) to oxygen (at the trinuclear T2/T3 site). One would expect most residues close to these coppers to be evolutionarily conserved due to the structural and energetic constraints imposed by the sites. To this end, all residues within 7 Å of any Cu were listed (Supporting Information, Table S3).

As expected, residues close to the copper sites are highly conserved across the isozymes: Sequence variation occurs only at four sites 68, 79, 241, and 425, of the 42 sites within 7 Å of Cu. The nonconserved sites are shown in Figure 3. Sites 79 and

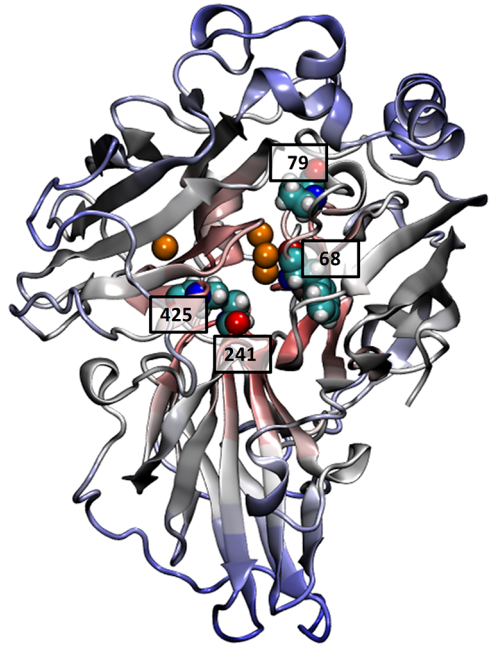


Figure 3. Location of the four nonconserved sites within 7 Å of any copper atom (orange spheres) 68, 79, 241, and 425 across the four isoenzymes, based on the constructed 3FPX′ homology model.

241 are particularly interesting as they are located in turns close to the T3 site: In the 3FPX′ homology model, residue 79 interacts with the three histidines coordinating one Cu and residue 241 packs against the three histidines coordinating the other Cu of T3.

Furthermore, the residue at position 425 is a structural neighbor to residue 241, and both residues are in contact in the 3FPX′ model. Because of this situation, sequence variation at positions 79, 241, and 425 may potentially influence enzyme function by modulating electrostatics or dynamics near T3. In contrast, although position 68 is located close to the T2 solvent channel, the side chain is directed away from the copper atoms.

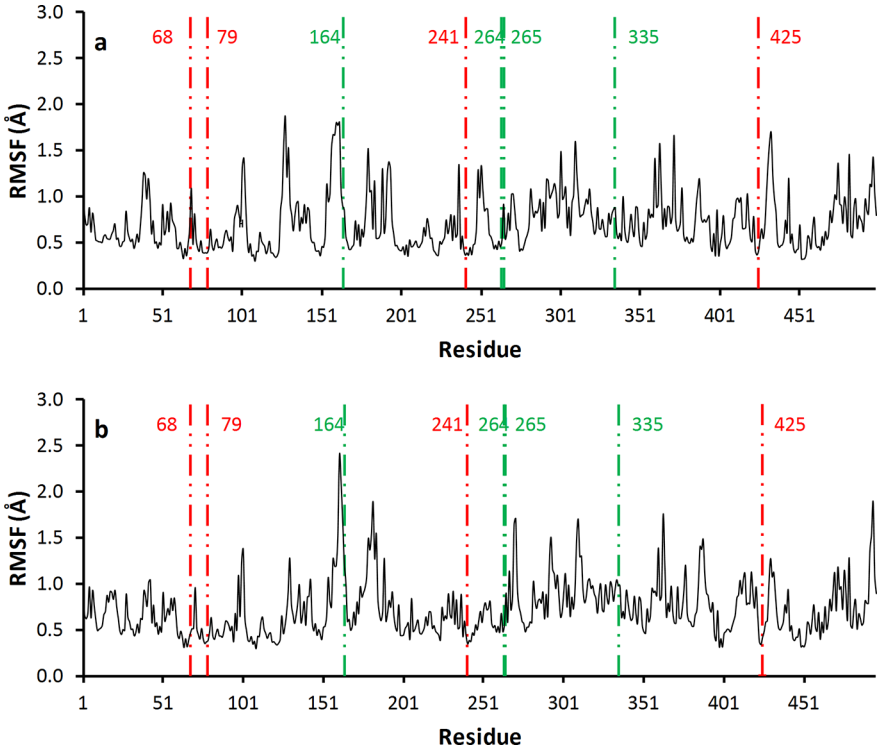


Figure 4. Computed root-mean-square fluctuations (RMSF) for the last 10 ns of the MD simulations of (a) 3FPX and (b) 3FPX′ both protonated with PROPKA corresponding to pH 7. Nonconserved residues near T3-coppers (68, 79, 241, and 425) and near the substrate binding site (164, 264, 265, and 335) are indicated with red and green vertical lines, respectively.

The RMSF time series from the MD simulations at pH 7 (Figure 4, numerical data in Table S7, Supporting Information) are consistent with these positions of the residues: In the case of both 3FPX′ and 3FPX, reduced fluctuations occur for the buried residues 79, 241, and 425 (RMSF ~ 0.4 Å) compared to the more exposed position 68 (RMSF ~0.5 - 0.7 Å).

We then analyzed the free energy variations of all residues within 7 Å of Cu. For the conserved residues where residues are identical, $\Delta\Delta G_{i,\text{max-min}}$ is, as expected, small (<0.2 kcal/mol) or moderately small (<1 kcal/mol), as for positions 424 (0.9 kcal/mol, a neighbor to 425), 206 (0.5 kcal/mol), and 451 (0.4 kcal/mol). However, the nonconserved sites within 7 Å of Cu contribute to variations in stability, compared to the average mutation, with $\Delta\Delta G_{i,\text{max-min}}$ values ranging from 2.4 kcal/mol (site 79), over 1.5 kcal/mol (site 425) to slightly less than 1 kcal/mol (sites 241 and 68). This implies that some of the functionally important variations also have large stability effects, to be investigated further below.

Analysis of positions 79, 241, and 245 close to T3 shows that Ala79 destabilizes Lccγ' and Lccδ' by 2.1 and 2.4 kcal/mol respectively, relative to Pro79 in Lccα' and Lccβ', principally due to penalties by backbone entropy (~0.9 kcal/mol) and polar solvation free energy (~0.8 kcal/mol). The fact that the destabilization occurs in both the unstable laccases relative to the stable isoforms, and that the protocol computes these effects to be similar, shows that the approach is robust as a way to analyze main contributors to differential total protein stability. It also concurs with the observation that proline generally stabilizes turns.[13] The finding that proline at site 79 is a substantial contributor to stability, and plausibly function via its position, may indicate a role in stability-function trade-off observed experimentally for the laccase isoforms, of potential relevance to protein engineering.

In contrast, at position 241, Gly *stabilizes* Lccγ' and Lccδ' by ~0.8 kcal/mol relative to Ala in Lccα' and Lccβ', consistent with general preferences for turn-composition, and the computed numbers are robust for both classes of isozymes. Position 425 only differs in Lccγ' where Thr425 destabilizes the isoenzyme by ~1.5 kcal/mol relative to Val425 in the other three isoenzymes. The destabilization mainly results from hydrophobic solvation (1.0 kcal/mol), that is, the differential advantage of burying Val vs Thr. For the last of the nonconserved sites within 7 Å of Cu, site 68, Ile in Lccγ' is destabilized by ~0.7 kcal/mol relative to Phe in the other three isoenzymes for similar reasons. Overall, the predicted net effect of positions 79, 241, and 245 in Lccγ' and Lccδ' is destabilization relative to Lccα' and Lccβ', which correlates with the experimentally observed stability trends. This suggests that these sites are responsible for the observed trade-off between stability and k_{cat}, as k_{cat}, which is 3−13 times higher in the less stable γ− and δ−isoforms, will probably be affected by these variations close to copper sites.

Sequence Variations near the Substrate Binding Site. No crystal structures have been published for the four TvL isoenzymes studied here, which is the reason we obtained homology models as described in the Computational Details section. However, a crystal structure[48] with a bound substrate (2,5-xylidine) is available for the homologous *Trametes versicolor* laccase 1KYA.pdb. Inspection of this structure shows that the substrate cavity is formed by five looplike regions (positions 153−166, 203−207, 264−275, 332−337, 384−395) and a helical region (positions 453−461).

To identify the potential sites involved in substrate binding and, thus, the determination of K_M (whereas the residues close to copper would be anticipated to affect the redox potential and k_{cat}), we identified substrate-binding residues as those within 7

Table 7. Sequence Variations Contributing ≥2 kcal/mol to Differential Stability of Any Two Laccase Isoforms[a]

residue no.	Lccα'	Lccβ'	Lccγ'	Lccδ'	$\Delta\Delta G_{i,\text{max-min}}$	location
27	V (0.0)	V (−0.1)	A (2.3)	A (2.2)	2.5	partially exposed C-terminus of β-strand
79	P (0.0)	P (0.0)	A (2.1)	A (2.4)	2.4	near T3 Cu
136	R (0.2)	L (−1.8)	L (−1.8)	L (−1.8)	2.1	exposed C-terminus of α-fragment
177	S (1.0)	S (0.7)	F (−0.9)	F (−1.0)	2.0	exposed loop
192	H (0.3)	P (−2.2)	Q (0.3)	Q (0.1)	2.5	exposed β-turn
209	Y (−1.9)	H (0.0)	F (−2.8)	Y (−2.0)	2.8	buried β-strand
223	V (0.0)	T (2.3)	V (−0.1)	A (2.4)	2.5	C-terminus of β-strand/start of buried turn
265	F (−0.3)	F (0.1)	T (1.9)	A (1.3)	2.1	exposed, substrate site, discussed above
269	G (0.0)	G (0.0)	D (2.0)	S (2.4)	2.4	exposed, in a substrate site loop
346	P (0.1)	P (0.0)	P (0.0)	S (2.2)	2.1	exposed loop, bottom T3 channel inlet
356	L (0.0)	I (2.2)	L (1.4)	L (1.3)	2.2	exposed helical segment
373	P (−0.4)	S (1.6)	T (1.3)	A (1.4)	2.1	exposed β-sheet
422	F (−0.3)	F (0.0)	R (2.2)	R (2.1)	2.5	T2 solvent channel entrance
478	N (0.2)	N (0.1)	N (0.2)	T (2.3)	2.1	loop to C-terminal helix
494	L (0.2)	R (2.0)	L (−0.1)	L (−0.1)	2.1	last residues of C-terminal helix

[a]Residues are numbered according to the crystal structure of *Trametes hirsuta* laccase (PDB-ID: 3FPX). $\Delta\Delta G_i$ is indicated in parentheses following one-letter amino acid codes.

Å of the 2,5-xylidine substrate in the 3FPX′ homology model superimposed on chain A of 1KYA. The list (Supporting Information, Table S4), excludes those positions within 7 Å of Cu that were already discussed above. For the conserved residue positions in the putative substrate binding region, $\Delta\Delta G_{i,\text{max-min}}$ is, as expected, small (0.2−0.3 kcal/mol) at positions 162, 163, 165, 332, but moderate at position 334 ($\Delta\Delta G_{i,\text{max-min}}$ = 0.8 kcal/mol). Even when the same residue occurs across isozymes, sometimes it occupies different conformations or interact differently with the surroundings (e.g., interact with sites that do vary across isozymes) to produce larger variation in free energy.

For the four nonconserved positions, substantial $\Delta\Delta G_{i,\text{max-min}}$ is found at position 265 ($\Delta\Delta G_{i,\text{max-min}}$ = 2.1 kcal/mol), and moderate stability effects occur at sites 335 (0.8 kcal/mol) and 264 (0.6 kcal/mol), whereas position 164 has little importance for stability ($\Delta\Delta G_{i,\text{max-min}}$ = 0.1 kcal/mol). The large $\Delta\Delta G_{i,\text{max-min}}$ at position 265 is most pronounced from Phe265 in Lccα to Thr265 in Lccγ ($\Delta\Delta G_i$ = 2.1 kcal/mol) but also the change from Phe265 in Lccα or Lccβ to Ala265 in Lccδ is significantly destabilizing ($\Delta\Delta G_i$ = 1.6 kcal/mol from Lccα to Lccδ). The destabilization results, as in the examples discussed above, from hydrophobic and van der Waals free energies relating to the advantage of burying the more hydrophobic Phe. Residue 265 is solvent-exposed and has substantial B-factors (based on 1GYC), demonstrating the dynamics at this site; that is, the computed role of this site should be treated with caution. The increased dynamics due to solvent exposure for binding site residues is also evident from the substantial MD RMSF values at positions 164 (∼0.9 Å), 264 (∼0.5−0.6 Å), 265 (0.9 Å), and 335 (0.9−1.0 Å) (see the Supporting Information, Table S7).

Structural inspection indicates that, compared to Phe265, the smaller side chains of Thr265 and, in particular, Ala265 may allow the binding of larger substrates with less steric clashes, possibly explaining partly the stronger binding (lower K_M) of these isozymes. This would indicate a stability-function trade-off relating to K_M, not k_{cat}, also of potential relevance to laccase optimization and to understand the molecular drivers of the experimentally observed large anticorrelation between K_M and T_{opt} (r = −0.90, R^2 = 0.81). Thus, the calculations suggest that stability may be traded for a larger and more electrostatically favorable ABTS-binding site. The moderate effect on stability for other sites suggests that these may be engineered to optimize K_M without substantial structural destabilization.

Functionally Non-Important Sites Contributing Substantially to ΔΔG. As discussed above, positions most likely affecting enzyme function (k_{cat} or K_M) were highly conserved, but some important variations suggested stability-function trade-offs potentially explaining the experimentally observed kinetic parameters, although this requires further experimental confirmation by site-directed mutagenesis. However, to be discussed below, the major contributions to $\Delta\Delta G$ across the four laccase isozymes come from sites more than 7 Å from copper and substrate-binding sites, partly because the latter form a minority of all sites and partly because they tend to be conserved.

Based in Figure 2, we examined all sequence positions where $\Delta\Delta G_{i,\text{max-min}} \geq 2.0$ kcal/mol, as seen in Table 7 (a corresponding list without subtraction of van der Waals repulsions is included in Table S5 of the Supporting Information). This corresponds to the sequence positions yielding the highest 3% of $\Delta\Delta G_{i,\text{max-min}}$. The functionally important sequence positions 79 and 265 discussed above are listed also, as they are among the sites providing the top 3% differential stability.

The largest contribution to $\Delta\Delta G$ (−2.8 kcal/mol) is found at position 209 with a variation from His in Lccβ' to Phe in Lccγ'. Position 209 is partially buried and occurs at the beginning of a β-strand. Inspection of the component energy terms shows that hydrophobic solvation and van der Waals interactions again explain most of this effect, consistent with the large increase in hydrophobicity with Phe in the buried site. Positions 27, 136, 177, and 223 reflect the same mechanism of stabilization: Site 27 occurs at the intersection between the second β-strand of 1KYA and the following loop. Ala at this position (Lccγ', Lccδ') is ∼2.5 kcal/mol more destabilizing than Val (Lccα', Lccβ'), primarily due to hydrophobic solvation. Position 136 occurs in a partly solvent-exposed helical fragment, and Arg in Lccα' is disfavored by ∼2 kcal/mol relative to Leu found in the other three isoenzymes. The output structures of the calculations show favorable packing with hydrophobic neighbors for Leu136 in Lccβ', Lccγ', and Lccδ', with hydrophobic solvation causing most of the differential

Table 8. Key Residues of Importance from Full Sequence Alignment of TvL Isoenzymes (Lccα, Lccβ, Lccγ, and Lccδ) and Selected Fungal Laccases[a]

laccase	stability	27	79	136	177	192	209	223	265	269	346	356	373	422	478	494
3FPX[82]		V	P	R	A	K	H	V	F	G	P	L	V	F	N	L
Lcca[83]	T_{opt}: 75 °C (ABTS)	V	P	R	S	H	Y	V	F	G	P	L	P	F	N	L
Lccb	T_{opt}: 80 °C (ABTS)	V	P	L	S	P	H	T	F	G	P	I	S	F	N	R
Lccg[84]	T_{opt}: 60 °C (ABTS)	A	A	L	F	Q	F	V	T	D	P	L	T	R	N	L
Lccd[85]	T_{opt}: 45 °C (ABTS)	A	A	L	F	Q	Y	A	A	S	S	L	A	R	T	L
rLac1[86]	$t_{1/2}$: 60 min at 60 °C (DMP)	A	P	L	F	Q	W	A	N	G	P	L	T	Q	N	L
rLac2[86]	$t_{1/2}$: 25 min at 60 °C (DMP)	A	P	D	Y	S	Y	A	L	G	P	L	T	Q	N	L
MqL[87]	T_{opt}: 75 °C (SGZ)	V	P	L	S	K	H	A	S	N	P	L	S	Y	N	L
PM1L[88]	T_{opt}: 80 °C (GUA)	V	P	L	S	K	H	A	S	N	P	L	S	Y	N	L
POX1[89]		A	P	L	F	S	F	A	L	G	A	L	S	R	~	L
POX2[90]	T_{opt}: 50 °C (ABTS)	A	P	L	Y	S	F	A	L	G	P	L	E	R	~	L
POX3[91]	"unstable"	A	P	L	S	Q	Y	T	A	G	S	L	T	R	E	Y
POXA1b[92]	$t_{1/2}$: 3 h at 60 °C	A	P	L	S	K	Y	A	S	G	P	L	D	R	N	S
Lac1b[93]	T_{opt}: 60 °C (SGZ)	A	P	L	N	L	W	V	S	G	P	L	E	V	N	L
3T6V[94]	T_{opt}: 75 °C (ABTS)	A	P	L	T	V	Y	T	S	G	P	C	S	W	N	A
1V10[95]	T_{opt}: 40 °C	A	P	L	N	S	Y	V	N	G	P	L	S	R	N	N
1HFU[96]	T_{opt}: 60 −70 °C	V	A	L	Y	Q	W	V	K	G	S	M	E	K	N	L
4A2F[97]	T_{opt}: 65 − 70 °C	V	P	L	S	K	Y	A	S	N	P	L	S	Y	N	L
LAP2[98]	T_{opt}: 50 − 60 °C (ABTS)	V	P	L	S	A	Y	T	F	G	P	I	S	F	N	L
TsL[99]	$t_{1/2}$: 2.3 h at 60 °C	V	P	L	A	H	H	V	G	G	P	L	T	F	N	L

[a]3FPX (PDB-ID: 3FPX): *Coriolus hirsutus* laccase. Lcca (AAW29420.1), Lccb (CAA77015.1), Lccg (BAA23284.1), and Lccd (CAA59161.1): *Trametes versicolor* laccase isoforms. rLac1 (AFI57924.1) and rLac2 (AFI57925.1): *Ceriporiopsis rivulosa* laccase isoforms. MqL (AAF06967.1): *Marasmius quercophilus* laccase, PM1 (CAA78144.1): PM1 laccase. POX1 (CAA78144.1), POX2 (PRF: 1587216), POX3 (CAR48260.1), and POXA1b (CAA06291.1): *Pleurotus ostreatus* laccase isoforms. Lac1B (ACL93462.1): *Cerrena gunicolor* laccase. 3T6V (3T6V): *Steccherinum ochraceum* laccase. 1V10 (CAE81289): *Rigidoporus microporus* laccase. 1HFU (1HFU): *Coprinus cinereus* laccase. 4A2F (AAF70119.2): *Coriolopsis gallica* laccase. LAP2 (AAM18407.1): *Trametes pubescens* laccase. TsL (ACO51010.1): *Trametes sanguinea* laccase.

stability (1.8 kcal/mol). Position 177 is in the loop comprised by residues 172−185. Phe in Lccγ′ and Lccδ′ is ∼2 kcal/mol more stable than Ser in Lccα′ and Lccβ′. Position 223 is buried in a β-turn in the wall of the T2 water channel. Val is favored at this position in Lccα′ and Lccγ′ compared to Thr (2.3 kcal/mol) and Ala (2.4 kcal/mol) in Lccβ′ and Lccδ′, respectively. Again, hydrophobic solvation destabilizes Thr223 and Ala223. In addition, Thr223 and Ala223 are destabilized by ∼1 kcal/mol, respectively. In all these cases, hydrophobic (de)solvation of buried or even partly exposed sites is a major determinant of differential laccase stability, relevant to future engineering of fungal laccases.

A different and interesting case is seen for Position 192, which is located in a solvent-exposed β-turn. Only Lccβ′ has Pro at this position resulting in 2.3−2.5 kcal/mol stabilization compared to the other three isoenzymes. This proline is thus also suggested to be a relevant trait in a stability-optimized laccase combining the best of all four isoforms.

Position 269 is found in the same loop as residue 265 discussed above. Although located distantly from the substrate, mutations at position 269 may influence substrate binding through changed dynamics and/or structure of the substrate binding loop. FoldX predicts stabilization of Gly at this position in Lccα′ and Lccβ′ by 2.0−2.4 kcal/mol relative to D269 in Lccγ′ and S269 in Lccδ′. The destabilization of the latter isoforms primarily results from positive contributions from backbone entropy to Lccγ′ (1.8 kcal/mol) and Lccδ′ (1.6 kcal/mol).

Of other interesting variations, position 346 occurs in a solvent-exposed loop at the bottom of the T3 solvent channel. The corresponding channel in *Melanocarpus albomyces* laccase has been associated with enzyme function.[75] Only Lccδ′ has Ser at this position causing a destabilization by ∼2.2 kcal/mol relative to Pro. This destabilization primarily consists of entropic penalties and burial of the less hydrophobic Ser and is consistent with known tendencies for proline stabilization in loops[76] and turns.[12,77] Thus, stabilization by Pro may have been traded for a function via the more polar and flexible Ser at this position, although this hypothesis requires investigation.

Position 356 is the last residue in a α-helical segment and has a solvent-exposed backbone carbonyl oxygen. In Lccα′ the side chain of Lys472 on a neighboring α-helical segment hydrogen bonds with the carbonyl, accounting for ∼0.8 kcal/mol of the stability due to this site in Lccα′.

Position 373 occurs in an exposed β-sheet, and relative to Pro373 in Lccα, significant destabilization is found for Ser373 in Lccβ (∼2 kcal/mol), Thr373 in Lccγ (∼1.7 kcal/mol), and Ala373 in Lccδ (∼1.9 kcal/mol). Backbone hydrogen bonds contribute 0.5−1 kcal/mol to this relative destabilization. Hydrophobic solvation also destabilizes, particularly for Lccβ (0.9 kcal/mol) and Lccδ (0.8 kcal/mol).

Position 422 is located in a nonpolar region at the beginning of the channel leading to the T2 site. Arg422 (Lccγ, Lccδ) is destabilizing by ∼2 kcal/mol relative to Phe422 (Lccα, Lccβ) with a significant part (∼1.4 kcal/mol) being due to hydrophobic desolvation. Mutagenesis experiments may reveal if the charged Arg422 also contributes to the relatively higher activity of (Lccγ, Lccδ).

Position 478 is located at the beginning of the loop connecting the two final α-helices, relevant due to reports on C-terminal modifications affecting fungal laccase function and/or stability.[78−81] Lccδ′ has Thr at this position causing destabilization by ∼2 kcal/mol relative to Asn478 in the three remaining isoenzymes. Side chain hydrogen bonds and

intraresidue torsional repulsion both contribute with ∼1 kcal/mol to destabilization, because the side chain of Thr compared to Asn is less suited to form hydrogen bonds with the acceptors in the protein.

Position 494 terminates the large α-helix in the last part of the isoenzymes. Arg at this position is unique to Lccβ, and Leu is found in the other isoenzymes. The principal contributors to the relative destabilization (∼2 kcal/mol) of Arg494 in Lccβ are hydrophobic solvation (∼1.3 kcal/mol) reflecting the better hydrophobic packing of Leu and polar solvation (∼1.1 kcal/mol) reflecting the penalty for burial of Arg. Despite a net destabilization, FoldX also predicts a stabilizing contribution of 1.2 kcal/mol from side chain hydrogen bonds involving Arg.

Comparison with Other Fungal Laccases. To validate further our findings on the drivers of differential stability across the four isozymes, the sequences of Lccα, Lccβ, Lccγ, and Lccδ were aligned with 16 other fungal laccase sequences with experimentally available stability data, as seen in Table 8 (for full alignment and details in Supporting Information, Figure S4).

Different stability proxies were reported (T_{opt}, $t_{1/2}$, $T_{1/2}$) as well as different experimental conditions (different substrates, buffers, etc.). Still, the data do allow a consistent classification of laccases into high- and low-stability groups. Upon examining sequence differences, we focus on positions 27, 79, 265, 346, and 422 (Table 7) where our protocol has predicted substantial destabilization of Lccγ and/or Lccδ relative to Lccα and Lccβ. The interesting case of Thr478 in Lccδ is also discussed.

At position 27, Ala is present in half of the sequences and Val is present in the other half. The sequences with lowest stabilities, Lccδ (T_{opt} = 45 °C), Rlac2, the *Pleuroteus ostreatus* laccases POX2 (T_{opt} = 50 °C), and POX3 ("unusually instable"[100]), and 1V10 (T_{opt} = 40 °C) all have Ala at this position. In contrast, high T_{opt} (75−80 °C) correlates with Val at this position, except for 3T6V. This indicates that the destabilizing role for Ala at position 27 as predicted by our modified FoldX protocol is consistent with a similar correlation across most other laccase wild types.

Position 79 had a large destabilizing contribution to ΔΔ*G* for Lccγ′, Lccδ′ and was discussed above with a focus on its location close to one of the T3 copper atoms. The sequence alignment shows that only Lccγ, Lccδ (T_{opt} = 45 °C), and *Coprinus cinerus* laccase (T_{opt} = 60−70 °C) have Ala at this position. All remaining laccases have Pro at this position, which we predicted to contribute to protein stability.

The alignment shows considerable variation at position 265 found in the substrate binding region, a potential source of the variable K_M data for the isoenzymes. Phe is found in *Trametes hirsuta* laccase (unknown stability), Lccα, Lccβ, and *Trametes pubescens* laccase (T_{opt} = 50−60 °C). Ser occurs in six laccases of variable stability. In contrast, Ala and Leu occur in five laccases of generally low stability. In particular, Ala only occurs in the least stable TvL, Lccδ, and in the unstable POX3.

Pro is present at position 346 in 16 sequences, suggesting that the remaining four laccases may be penalized for having alternative residues at this position. Confirming this result, Lccδ (T_{opt} = 45 °C), POX3 ("unusually instable"), and *Coprinus cinerus* laccase (T_{opt} = 60−70 °C) all with Ser at this position fall within the "low" to "moderate" stability group. The influence of position 346 on POX1 could not be inferred due to the absence of experimental stability data for this laccase.[89]

Substantial destabilization was predicted with FoldX for R422 in Lccγ′ and Lccδ′ relative to F422 in Lccα′ and Lccβ′. The sequence alignment offers support for this finding by showing that Arg at this position is also found in the less stable POX isoforms and in 1V10 (T_{opt} = 40 °C), while Phe is found in the more stable laccases of *Trametes sanguinea* (40% activity retained after 3 h at 60 °C) and *Trametes pubescens* (T_{opt} = 50−60 °C). Thus, we suggest that this site is a relevant target for laccase optimization.

Evolutionary Implications. The several isoforms known for many fungal laccases are not widely considered, although it is often unclear from the literature which isoform has been studied in a particular work. Recent data indicate that laccase isoforms vary substantially in stability and functional proficiency (in terms of k_{cat} and K_M for particular substrates) and that these properties tend to anticorrelate,[38] consistent with the generally encountered problem of stability-function trade-off in protein engineering,[16,101] which would be highly relevant to solve.

The use of several isoforms is evolutionary in its origin and probably implies adaptation of proficiency for different substrates, where some substrates impose large constraints on stability due to changes required in the substrate binding sites. They may also arise from trade-offs between proficiency and stability, since stable proteins may be more selected for in variable environments, as the enzymes are often excreted. We have identified several sites near copper or substrate that illustrate such a trade-off. Thus, it is conceivable that, under rough conditions in the environment, the more stable isoforms may be used despite their lower proficiency. These questions remain unsolved but should be of major importance in protein engineering, in particular, in the search for proteins that are simultaneously optimized for proficiency and stability. The variable sites close to copper or substrate that have been identified in this work are key contenders in understanding these evolutionary choices.

CONCLUSIONS

In this work, a previously developed protocol[26] for describing relative stabilities of laccase mutants with multiple mutated sites was extended to enable studies of protein isoforms with substantial sequence variation, while correcting previously identified issues of structure sampling, template bias, and hysteresis. As an application of the protocol, four *Trametes versicolor* laccase (TvL) isoenzymes were investigated that display considerable variation in sequence, proficiency, and stability.[27] The modified protocol required editing the original TvL sequences (Lccα, Lccβ, Lccγ, and Lccδ) as well as an unbiased crystal structure template (3FPX) to produce the MD-validated homology models used for calculating the relative stabilities and their residual and energy term components. We emphasize that the new method is not a black-box and is only applicable when a high-resolution crystal structure with no strong similarity to any of the studied forms can be used to generate a structurally unbiased, threaded homology model as input for FoldX. Also, the method does not give quantitative accuracy but provides stability trends in good qualitative agreement with experimental data.

In this process, previous observations of template bias in FoldX and the problems of the repair protocol were confirmed, and the hysteresis correction (average of forward and backward mutations) was again found to provide stability rankings in good agreement with the experimental clustering of the four isozymes. It was found that deleting insertions in loops relative to the minimal model but substituting a single position in a β-

sheet with glycine throughout all TvL sequences to maintain backbone structure produced more consistent rankings of the proteins than when deletions were performed indiscriminately in all sequence areas considered as insertions. The experimentally observed stability ranking of Lccα' ~ Lccβ' ≫ Lccγ' ~ Lccδ' was obtained with several conditions and changes in protocol, implying that the main physics of the stability calculation are realistic. With the most realistic conditions, computed rankings were consistent with all experimental T_{opt} values.

A relatively small number of sites are responsible for the majority of the stability variation, and the stability effects were approximately Gaussian-distributed, in good agreement with experimental distributions of stability effects of mutations. Critical sites contributing most to the substantial differences in stability of the isozymes were then identified (Table 7) which should be of relevance in the optimization of highly robust laccases. These sites were divided into three sets, those close to copper, those close to the substrate binding site, and the remaining sites.

Of particular interest to the widely discussed stability-function trade-off were the findings of several variable sites within 7 Å of copper sites or substrate. Four variable sites out of 42 within 7 Å of Cu (68, 79, 241, and 425) contribute markedly to differential stability. The largest destabilization (~2.4 kcal/mol) among these functionally important sites was seen for Ala79 in Lccγ' and Lccδ' relative to Pro79 in Lccα' and Lccβ'; also, Thr425 in Lccγ' was 1.5 kcal/mol destabilizing relative to Val in the other isoforms; these effects correlate well with the overall stability trends of the isoforms and are thus suggested to be of importance in the experimentally observed stability-k_{cat} trade-off.

Of the variable sites in the substrate-binding region, position 265 had the largest impact on stability, with Phe in Lccα' being ~2 kcal/mol more stable than Ala (Lccδ') and Thr (Lccγ'). Both these stabilization effects and those of sites 79 and 425 close to Cu were mainly due to favorable desolvation of hydrophobic residues. The identified effects correlated with the experimental stability trends of the isoenzymes, and with K_M spanning 2 orders of magnitude for the isoforms, site 265 is a likely main determinant of stability-K_M trade-off, as desolvation of a large hydrophobic group at position 265 increases stability but sterically impairs binding affinity. The identified sites and mechanisms of stability-function trade-off should be of relevance to future laboratory optimization of laccases.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Figures S1–S4, Tables S1–S7, and associated text provide additional data and details of sequence aligment, sensitivity analysis, MD-simulations, and FoldX calculations for further detailed analysis. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kpj@kemi.dtu.dk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank the The Danish Council for Independent Research | Technology and Production Sciences (FTP) (Grant No. 10-082488) and the Danish Center for Scientific Computing (Grant No. 2012-02-23) for support.

■ REFERENCES

(1) Robertson, A. D.; Murphy, K. P. *Chem. Rev.* **1997**, *97*, 1251–1268.
(2) Vogt, G.; Argos, P. *Folding Des.* **1997**, *2* (Supplement1), S40–S46.
(3) Kumar, S.; Tsai, C.-J.; Nussinov, R. *Protein Eng.* **2000**, *13*, 179–191.
(4) Jaenicke, R.; Böhm, G. *Curr. Opin. Struct. Biol.* **1998**, *8*, 738–748.
(5) Vieille, C.; Zeikus, G. J. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1–43.
(6) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. *Science* **2003**, *299*, 1694–1697.
(7) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001**, *409*, 258–268.
(8) Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A. *Trends Biochem. Sci.* **2001**, *26*, 100–106.
(9) Romero, P. A.; Arnold, F. H. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866–876.
(10) Matsui, I.; Harata, K. *FEBS J.* **2007**, *274*, 4012–4022.
(11) Chen, J.; Stites, W. E. *Biochemistry* **2001**, *40*, 15280–15289.
(12) Trevino, S. R.; Schaefer, S.; Scholtz, J. M.; Pace, C. N. *J. Mol. Biol.* **2007**, *373*, 211–218.
(13) Fu, H.; Grimsley, G. R.; Razvi, A.; Scholtz, J. M.; Pace, C. N. *Proteins* **2009**, *77*, 491–498.
(14) Solá, R.; Rodríguez-Martínez, J.; Griebenow, K. *Cell. Mol. Life Sci.* **2007**, *64*, 2133–2152.
(15) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 7342–7372.
(16) Tokuriki, N.; Stricher, F.; Serrano, L.; Tawfik, D. S. *PLoS. Comput. Biol.* **2008**, *4* (2), e1000002 DOI: 10.1371/journal.-pcbi.1000002.
(17) Bornscheuer, U. T.; Pohl, M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 137–143.
(18) Eijsink, V. G.; Bjørk, A.; Gåseidnes, S.; Sirevåg, R.; Synstad, B.; van den Burg, B.; Vriend, G. *J. Biotechnol.* **2004**, *113*, 105–20.
(19) Mark, A. E.; Van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167–176.
(20) Patrick Brady, G.; Sharp, K. A. *J. Mol. Biol.* **1995**, *254*, 77–85.
(21) Patrick Brady, G.; Szabo, A.; Sharp, K. A. *J. Mol. Biol.* **1996**, *263*, 123–125.
(22) Noskov, S. Y.; Lim, C. *Biophys. J.* **2001**, *81*, 737–750.
(23) Potapov, V.; Cohen, M.; Schreiber, G. *Protein Engin. Des. Select.* **2009**, *22*, 553–560.
(24) Guerois, R.; Nielsen, J. E.; Serrano, L. *J. Mol. Biol.* **2002**, *320*, 369–387.
(25) Khan, S.; Vihinen, M. *Hum. Mutat.* **2010**, *31*, 675–684.
(26) Christensen, N. J.; Kepp, K. P. *J. Chem. Inf. Model.* **2012**, *52*, 3028–3042.
(27) Koschorreck, K.; Richter, S. M.; Swierczek, A.; Beifuss, U.; Schmid, R. D.; Urlacher, V. B. *Arch. Biochem. Biophys.* **2008**, *474*, 213–219.
(28) Nakamura, K.; Go, N. *Cell. Mol. Life Sci.* **2005**, *62*, 2050–2066.
(29) Dwivedi, U. N.; Singh, P.; Pandey, V. P.; Kumar, A. *J. Mol. Catal. B* **2011**, *68*, 117–128.
(30) Mot, A. C.; Silaghi-Dumirescu, R. *Biochemistry* **2012**, *77*, 1395–1407.
(31) Quintanar, L.; Stoj, C.; Taylor, A. B.; Hart, P. J.; Kosman, D. J.; Solomon, E. I. *Acc. Chem. Res.* **2007**, *40*, 445–452.
(32) Sirim, D.; Wagner, F.; Wang, L.; Schmid, R. D.; Pleiss, J. *Database* **2011**, DOI: 10.1093/database/bar006 .
(33) Rodríguez, C. S.; Herrera, J. L. T. *Biotechnol. Adv.* **2006**, *24*, 500–513.
(34) Mayer, A. M.; Staples, R. C. *Phytochemistry* **2002**, *60*, 551–565.

(35) Piscitelli, A.; Del Vecchio, C.; Faraco, V.; Giardina, P.; Macellaro, G.; Miele, A.; Pezzella, C.; Sannia, G. *C. R. Biol.* **2011**, *334*, 789−794.
(36) Wesenberg, D.; Kyriakides, I.; Agathos, S. N. *Biotechnol. Adv.* **2003**, *22*, 161−187.
(37) Stoilova, I.; Krastanov, A.; Stanchev, V. *Adv. Biosci. Biotechnol.* **2010**, *1*, 208−215.
(38) Madhavi, V.; Lele, S. S. *Bioresources* **2009**, *4*, 1694−1717.
(39) Necochea, R.; Valderrama, B.; Díaz-Sandoval, S.; Folch-Mallol, J. L.; Vázquez-Duhalt, R.; Iturriaga, G. *FEMS Microbiol. Lett.* **2005**, *244*, 235−241.
(40) Mikuni, J.; Morohoshi, N. *FEMS Microbiol. Lett.* **1997**, *155*, 79−84.
(41) Jönsson, L.; Sjöström, K.; Häggström, I.; Nyman, P. O. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1995**, *1251*, 210−215.
(42) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(43) Khatun, J.; Khare, S. D.; Dokholyan, N. V. *J. Mol. Biol.* **2004**, *336*, 1223−1238.
(44) Benson, D. A.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Wheeler, D. L. *Nucleic Acids Res.* **2005**, *33*, 34−38.
(45) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.
(46) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(47) Piontek, K.; Antorini, M.; Choinowski, T. *J. Biol. Chem.* **2002**, *277*, 37663−37669.
(48) Bertrand, T.; Jolivalt, C.; Briozzo, P.; Caminade, E.; Joly, N.; Madzak, C.; Mougin, C. *Biochemistry* **2002**, *41*, 7325−7333.
(49) Polyakov, K. M.; Fedorova, T. V.; Stepanova, E. V.; Cherkashin, E. A.; Kurzeev, S. A.; Strokopytov, B. V.; Lamzin, V. S.; Koroleva, O. V. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 611−617.
(50) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(51) Pieper, U.; Eswar, N.; Braberg, H.; Madhusudhan, M. S.; Davis, F. P.; Stuart, A. C.; Mirkovic, N.; Rossi, A.; Marti-Renom, M. A.; Fiser, A.; Webb, B.; Greenblatt, D.; Huang, C. C.; Ferrin, T. E.; Sali, A. *Nucleic Acids Res.* **2004**, *32*, D217−D222.
(52) Russell, R. B.; Barton, G. J. *Proteins* **1992**, *14*, 309−323.
(53) Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. *BMC Bioinf.* **2006**, *7*, 382.
(54) Kelley, L. A.; Sternberg, M. J. E. *Nat. Protocols* **2009**, *4*, 363−371.
(55) *Maestro, version 9.2*; Schrödinger, LLC: New York, 2011.
(56) *Maestro-Desmond Interoperability Tools, version 3.0*; Schrödinger: New York, 2011.
(57) Shleev, S. V.; Morozova, O. V.; Nikitina, O. V.; Gorshina, E. S.; Rusinova, T. V.; Serezhenkov, V. A.; Burbaev, D. S.; Gazaryan, I. G.; Yaropolov, A. I. *Biochimie* **2004**, *86*, 693−703.
(58) *Desmond Molecular Dynamics System, version 3.0*; D. E. Shaw Research: New York, 2011.
(59) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(60) Nose, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(61) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(62) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(63) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(64) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(65) Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 2773−2781.
(66) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. *J. Comput. Chem.* **2005**, *26*, 1752−1780.
(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(68) Jensen, K. P.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2006**, *2*, 1499−1509.
(69) Sedlák, E.; Stagg, L.; Wittung-Stafshede, P. *Arch. Biochem. Biophys.* **2008**, *479*, 69−73.
(70) Tadeo, X.; López-Méndez, B.; Castaño, D.; Trigueros, T.; Millet, O. *Biophys. J.* **2009**, *97*, 2595−2603.
(71) Baldwin, R. L. *Biophys. J.* **1996**, *71*, 2056−2063.
(72) Vogel, R.; fan, G. B.; Sheves, M.; Siebert, F. *Biochemistry* **2001**, *40*, 483−493.
(73) Jensen, K. P. *J. Phys. Chem. B* **2008**, *112*, 1820−1827.
(74) Christensen, N. J.; Kepp, K. P. *PLoS ONE* **2013**, *8*, e61985.
(75) Kallio, J. P.; Gasparetti, C.; Andberg, M.; Boer, H.; Koivula, A.; Kruus, K.; Rouvinen, J.; Hakulinen, N. *FEBS J.* **2011**, *278*, 2283−2295.
(76) Watanabe, K.; Kitamura, K.; Suzuki, Y. *Appl. Environ. Microbiol.* **1996**, *62*, 2066−2073.
(77) Suzuki, Y.; Oishi, K.; Nakano, H.; Nagayama, T. *Appl. Microbiol. Biotechnol.* **1987**, *26*, 546−551.
(78) Bleve, G.; Lezzi, C.; Spagnolo, S.; Tasco, G.; Tufariello, M.; Casadio, R.; Mita, G.; Rampino, P.; Grieco, F. *Protein Eng., Des. Sel.* **2013**, *26*, 1−13.
(79) Hakulinen, N.; Kiiskinen, L. L.; Kruus, K.; Saloheimo, M.; Paananen, A.; Koivula, A.; Rouvinen, J. *Nat. Struct. Biol.* **2002**, *9*, 601−605.
(80) Gelo-Pujic, M.; Kim, H. H.; Butlin, N. G.; Palmore, G. T. *Appl. Environ. Microbiol.* **1999**, *65*, 5515−5521.
(81) Robert, V.; Mekmouche, Y.; Pailley, P.; Tron, T. *Curr. Genomics* **2011**, *12*, 123−129.
(82) Pegasova, T. V.; Zwart, P.; Koroleva, O. V.; Stepanova, E. V.; Rebrikov, D. V.; Lamzin, V. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1459−1461.
(83) Necochea, R.; Valderrama, B.; Díaz-Sandoval, S.; Folch-Mallol, J. L.; Vázquez-Duhalt, R.; Iturriaga, G. *FEMS Microbiol. Lett.* **2005**, *244*, 235−241.
(84) Mikuni, J.; Morohoshi, N. *FEMS Microbiol. Lett.* **1997**, *155*, 79−84.
(85) Jönsson, L.; Sjöström, K.; Häggström, I.; Nyman, P. O. *BBA, Biochim. Biophys. Acta, Protein. Struct. Mol. Enzymol.* **1995**, *1251*, 210−215.
(86) Hildén, K.; Mäkelä, M.; Lundell, T.; Kuuskeri, J.; Chernykh, A.; Golovleva, L.; Archer, D.; Hatakka, A. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1589−1599.
(87) Dedeyan, B.; Klonowska, A.; Tagger, S.; Tron, T.; Iacazio, G.; Gil, G.; Le Petit, J. *Appl. Environ. Microbiol.* **2000**, *66*, 925−929.
(88) Coll, P. M.; Tabernero, C.; Santamaría, R.; Pérez, P. *Appl. Environ. Microb.* **1993**, *59*, 4129−4135.
(89) Giardina, P.; Cannio, R.; Martirani, L.; Marzullo, L.; Palmieri, G.; Sannia, G. *Appl. Environ. Microbiol.* **1995**, *61*, 2408−2413.
(90) Giardina, P.; Aurilia, V.; Cannio, R.; Marzullo, L.; Amoresano, A.; Siciliano, R.; Pucci, P.; Sannia, G. *Eur. J. Biochem.* **1996**, *235*, 508−515.
(91) Pezzella, C.; Autore, F.; Giardina, P.; Piscitelli, A.; Sannia, G.; Faraco, V. *Curr. Genet.* **2009**, *55*, 45−57.
(92) Giardina, P.; Palmieri, G.; Scaloni, A.; Fontanella, B.; Faraco, V.; Cennamo, G.; Sannia, G. *Biochem. J.* **1999**, *341*, 655−663.
(93) Janusz, G.; Mazur, A.; Checinsks, A.; Malek, W.; Rogalski, J.; Ohga, S. *J. Fac. Agric., Kyushu Univ.* **2012**, *57*, 41−49.
(94) Ferraroni, M.; Matera, I.; Chernykh, A.; Kolomytseva, M.; Golovleva, L. A.; Scozzafava, A.; Briganti, F. *J. Inorg. Biochem.* **2012**, *111*, 203−209.
(95) Garavaglia, S.; Teresa Cambria, M.; Miglio, M.; Ragusa, S.; Iacobazzi, V.; Palmieri, F.; D'Ambrosio, C.; Scaloni, A.; Rizzi, M. *J. Mol. Biol.* **2004**, *342*, 1519−1531.
(96) Ducros, V.; Brzozowski, A. M.; Wilson, K. S.; Ostergaard, P.; Schneider, P.; Svendson, A.; Davies, G. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 333−336.

(97) De la Mora, E.; Lovett, J. E.; Blanford, C. F.; Garman, E. F.; Valderrama, B.; Rudino-Pinera, E. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 564−577.
(98) Galhaup, C.; Goller, S.; Peterbauer, C. K.; Strauss, J.; Haltrich, D. *Microbiology* **2002**, *148*, 2159−2169.
(99) Dantán-González, E.; Vite-Vallejo, O.; Martínez-Anaya, C.; Méndez-Sánchez, M.; González, M. C.; Palomares, L. A.; Folch-Mallol, J. *Int. Microbiol.* **2008**, *11*, 163−169.
(100) Pezzella, C.; Autore, F.; Giardina, P.; Piscitelli, A.; Sannia, G.; Faraco, V. *Curr. Genet.* **2009**, *55*, 45−57.
(101) Beadle, B. M.; Shoichet, B. K. *J. Mol. Biol.* **2002**, *321*, 285−296.