Role of an Electrostatic Network of Residues in the Enzymatic Action of the *Rhizomucor miehei* Lipase Family†

Sanna Herrgård,‡,§,|| Cynthia J. Gibas,‡,⊥ and Shankar Subramaniam*,‡,||

Departments of Molecular and Integrative Physiology, Biochemistry, and Chemical Engineering, Center for Biophysics and Computational Biology, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, and Helsinki University of Technology, Laboratory of Bioprocess Engineering, P.O. Box 6100, 02015 HUT, Finland

Received July 22, 1999; Revised Manuscript Received November 29, 1999

ABSTRACT: We have used continuum electrostatic methods to investigate the role of electrostatic interactions in the structure, function, and pH-dependent stability of the fungal *Rhizomucor miehei* lipase (RmL) family. We identify a functionally important electrostatic network which includes residues S144, D203, H257, Y260, H143, Y28, R80, and D91 (residue numbering is from RmL). This network consists of residues belonging to the catalytic triad (S144, D203, H257), residues located in proximity to the active site (Y260), residues stabilizing the geometry of the active site (Y28, H143), and residues located in the lid (D91) or close to the first hinge (R80). The lid and the first hinge are associated with the interfacial activation of lipases, where an α-helical lid opens up by rotating around two hinge regions. All network residues are well conserved in a set of 12 lipase homologues, and 6 of the network residues are located in sequence motifs. We observe that the effects of modeled mutations R86L, D91N, and H257F on the pH-dependent electrostatic free energies differ significantly in the closed and open conformations of RmL. Mutation R86L is especially interesting since it stabilizes the closed conformation but destabilizes the open one. Site−site electrostatic interaction energies reveal that interactions between R86 and D61, D113, and E117 stabilize the open conformation.

Lipases (glycerol-ester-hydrolase, EC 3.1.1.3) are enzymes which catalyze the hydrolysis of acylglycerols. Microbial lipases are industrially important enzymes, and they are used in washing powders, in detergents, and in the dairy industry (*1*). Lipases can also be used to produce optically pure compounds for the pharmaceutical industry.

A unique feature of lipases is the increase in their catalytic activity at the water−lipid interface, known as interfacial activation (*2*). Crystallographic studies have shown that interfacial activation is associated with a conformational change, in which a lid consisting of at least one α-helix opens up by rotating around its hinge regions (*3*−*5*). In the inactive, closed conformation, the lid covers the active site so that the hydrophobic side of the amphipathic lid is buried in the active site and the polar side is solvent-exposed (*6*). At the interface, the lid opens up and reveals the hydrophobic side of the lid to the solvent. The hydrophilic side is buried in a polar cavity behind the lid. This is the open state in which the substrate is allowed to bind to the active site. Figure 1 shows the closed and open conformations of the fungal *Rhizomucor miehei* lipase (RmL).[1]

The lipase structure−function relationship has been the subject of both theoretical and experimental investigations during the past few years. Theoretical studies including molecular and Brownian dynamics simulations, essential dynamics, and normal-mode analysis have provided information about specificity, loop and domain motions of lipases, key residues, and interactions (*7*−*15*). Experimental studies include mutations of residues which affect lipase kinetics, activity, or specificity (*16*−*24*). Several key residues identified so far are either charged or polar, which indicates that electrostatic interactions may play an important role in the function or stability of lipases.

The aim of this study is to characterize in detail the key electrostatic interactions in the fungal RmL family, and to study the pH dependence of the conformational stability of these proteins. pK_a calculations are used to investigate the electrostatic interactions in the RmL family. The pK_a value is the pH at which an ionizable group is half-ionized, and it also describes the proton affinity of the ionizing group. The

† S.H. was supported by grants from the Foundation of Technology (Finland) and the Helsinki University of Technology Foundation for Financial Aid, and S.S. by NSF Grant DBI 94-04223, NSF Grant KDI 98-73384, NIH Grant R01 GM46535, and an NSF Metacenter Computer Allocation at NCSA.

* To whom correspondence should be addressed. Tel: +1-858-822 0986. Fax: +1-858-822 3752. Email: shankar@ucsd.edu.

‡ University of Illinois at Urbana−Champaign.

§ Helsinki University of Technology.

|| Present address: Department of Bioengineering, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0419.

⊥ Present address: Fralin Biotechnology Center, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

[1] Abbreviations: CaLB, *Candida antarctica* lipase B; HlL, *Humicola lanuginosa* lipase; HMM, hidden Markov model; NMR, nuclear magnetic resonance; PcL, *Penicillium camembertii* lipase; PDB, Protein Data Bank; RdL, *Rhizopus delemar* lipase; RmL, *Rhizomucor miehei* lipase; RMSD, root-mean-square atomic positional displacement; RnL, *Rhizopus niveus* lipase; RoL, *Rhizopus oryzae* lipase; SCOP, Structural Classification of Proteins; WT, wild-type.

10.1021/bi9916980 CCC: $19.00
Published on Web 02/23/2000

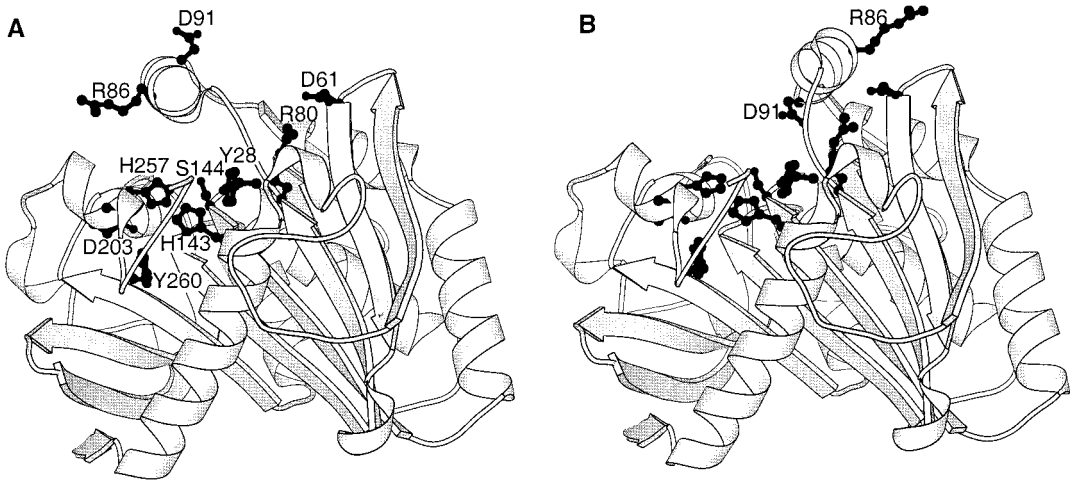

FIGURE 1: Schematic representation of the (A) closed and (B) open conformations of RmL. Residues forming the electrostatic network (S144-D203-H257-Y260-H143-Y28-R80-D91) and a few other key residues are labeled. The figure was generated with Molscript (*79*).

pK_a values of the titrating sites in a protein can be experimentally determined by nuclear magnetic resonance (NMR) spectroscopy (*25*−*28*). Standard NMR methods are currently limited to fairly small proteins with molecular masses of ∼30 kDa or less (*29*), although with new methods NMR spectroscopy studies of proteins as large as 110 kDa have been reported (*30*). An alternative to NMR spectroscopy is to use theoretical methods for determining pK_a values of the ionizable groups in a protein (*31*−*34*). Theoretical methods yield a description of the electrostatic interactions in a protein in terms of pK_a values and electrostatic interaction energies between titrating sites. It has been shown that it is possible to obtain a fairly good agreement between theoretically and experimentally determined pK_a values (*34*−*40*). Several studies have demonstrated that pK_a and titration calculations are useful tools for investigating structure−function relationships in proteins (*41*−*43*) and pH-dependent stability or association of biomolecules (*34*, *40*, *44*−*50*).

According to the SCOP database (Structural Classification of Proteins) (*51*), the fungal lipase family can be divided into two subfamilies: the triacylglycerol lipase family and the type-B carboxylesterase/lipase family. Here we focus on the triacylglycerol lipase family, and call it the RmL family. According to SCOP, this family consists of the following lipases: *Humicola lanuginosa* lipase (HlL), *Penicillium camembertii* lipase (PcL), *Rhizomucor miehei* lipase (RmL), *Rhizopus delemar* lipase (RdL), *Rhizopus niveus* lipase (RnL), and *Candida antarctica* lipase B (CaLB). We exclude CaLB from our analysis since this lipase does not demonstrate the interfacial activation and the conformational change associated with it (*52*, *53*). We include *Rhizopus oryzae* lipase (RoL) in the RmL family since this lipase has a 100% sequence identity with RdL and RnL. Sequence identity between RmL and the other RmL family members varies between 29% and 56%, and the RMSD of Cα atoms is 1.1−1.9 Å, according to FSSP (*54*).

We model several mutations in order to investigate the electrostatic interactions of some key residues as well as to understand in atomic detail the available experimental data. Electrostatic interactions are investigated by comparing pK_a values and site−site electrostatic interaction energies of the wild-type (WT) and mutant enzymes. We also study the relative effect of mutations on the pH-dependent stability of the closed and open conformations of RmL by comparing electrostatic free energies. Bioinformatics methods are used to identify sequence motifs, to find sequences with high similarity to the RmL family, and to characterize conserved electrostatic interactions. We will use the sequential numbering from RmL throughout the text, unless noted otherwise.

METHODS

Bioinformatics Methods. We used the program Meta-MEME 2.0.1 (*55*) for discovering sequence motifs (conserved regions) in the RmL family and for searching GenPept for homologous sequences. Meta-MEME uses the program MEME (Multiple Expectation-maximization for Motif Elicitation; *56*) for motif discovery. MEME uses expectation maximization to identify motifs in a set of DNA or protein sequences (*55*). Meta-MEME combines the models into a motif-based linear hidden Markov model (HMM). Nonmotif regions are not modeled rigorously. This reduces the number of free parameters, and the HMM can be effectively trained even with a small training set. Meta-MEME is able to produce a multiple alignment of motif regions, and search databases for statistically significant homologues.

Sequences of RmL (3tgl; *57*), RdL (1tic; *58*), HlL (1tib; *58*), and PcL (1tia; *59*) were obtained from the Brookhaven Protein Data Bank (PDB). These sequences were used to search for motifs with MEME. The two other family members, RnL and RoL, have amino acid sequences identical to that of RdL, and were left out. The motifs identified with MEME were used as an input for Meta-MEME for searching the GenPept database. A full sequence alignment was carried out using HMMer 2.1.1 (*60*; SR Eddy Group, Department of Genetics, Washington University). A structural alignment of RmL, RdL, HlL, and PcL from FSSP (*54*) was used to build a profile HMM, and other sequences were aligned to the profile. Positions of sequence motifs were included in the model as consensus secondary structure data. This method yields a sequence alignment which is based on a structural alignment, and in which sequence motifs are aligned correctly.

Preparation of Models. The following crystal structures were chosen for this study from the PDB: closed RmL (3tgl; *57*; resolution 1.9 Å), open RmL (4tgl; *6*; resolution 2.6 Å),

closed conformation of HlL (1tib; *58*; resolution 1.8 Å), and an intermediate conformation of RnL (1lgy; *61*; resolution 2.2 Å). In RmL and RnL, the coordinates of the first four residues are not resolved in the crystal structures. Crystal waters, alternate atom positions, and ligands were deleted. Polar hydrogens were added using HBUILD (*62*) in CHARMm through QUANTA (Molecular Simulations Inc., San Diego, CA) assuming protonated forms of His, Asp, and Glu. Histidine tautomers were assigned by selecting the lower-energy position for the hydrogen. The positions of the hydrogens were not optimized since HBUILD finds the energetically optimal position for the hydrogen atom.

Mutant structures were prepared by editing the desired side chain using QUANTA, and performing a rotamer search for the Karplus rotamer (*63*) with the lowest number of close contacts. Total potential energies of the wild type (WT) and the mutant were computed with CHARMm, to ensure that the modeled mutation caused no steric clashes. To ensure that the results obtained from the pK_a calculations for the mutants would be comparable with those obtained for the WT, the mutant proteins were not energy-minimized. All structures were superimposed to preserve a common orientation throughout the electrostatic calculations.

Solvent accessibilities of the lipase structures were computed with the program Naccess (version 2.1.1; *64*) using a solvent probe radius of 1.4 Å. Hydrogen bonds were analyzed using HBPLUS (*65*).

Electrostatics and pK_a Calculations. The University of Houston Brownian Dynamics (UHBD) program, version 5.1 (*66*, *67*), together with some supplementary programs (*33*, *34*) was used for the electrostatic and pK_a calculations. Each titrating site has a model pK_a value, which is its experimentally determined pK_a value in a solution. The pK_a shift from this model value for the same titrating group in a folded protein consists of three energy components. The energy penalty of moving the group from the solution into the protein interior is accounted for by the Born, or desolvation, term. Desolvation causes pK_a values to be shifted so that the residue remains in the neutral state. However, the effect of desolvation is partially compensated by the Coulombic interactions between the charged titrating group and the dipolar groups in a neutral protein. Third, the interaction of the titrating group with other ionized groups is calculated. To calculate the Born and Coulombic terms, we solve the linearized Poisson−Boltzmann equation using a finite difference method (*67*, *68*). We compute interactions between ionizing groups with the cluster method (*33*, *34*). The pK_a calculations also yield site−site electrostatic interaction energies, ΔG_{ij} (kcal/mol), and the electrostatic component of the free energy, G_{elec} (kcal/mol).

We employed the single-site titration model (*34*) which includes partial atomic charges of each ionizable group, and models the ionization by adding a +1 or −1 point charge to one central atom of the ionizing group. We used a dielectric constant of 20 to represent the protein interior, because it has been observed to produce the best agreement with experimental results, when this model is used (*34*). Partial atomic charges for the protein were taken from CHARMm (Molecular Simulations Inc., San Diego, CA) and radii from OPLS (*69*). The solvent was assigned a dielectric constant of 80 and ionic strength of 100 mM, and the ion exclusion radius was set to 2.0 Å. The ionic strength was chosen based on the observation that the pK_a values are fairly insensitive to ionic concentrations over 100 mM (*34*). A probe with a radius of 1.4 Å was used to define the boundary between the protein and solvent dielectric regions (*70*). Electrostatic potentials were calculated by focusing on one titratable residue at a time (*35*, *71*). Focusing runs with grid spacings of 1.5, 1.2, 0.75, and 0.25 Å and extents of 75, 18, 11.25, and 5 Å were used. The temperature was set at 298 K.

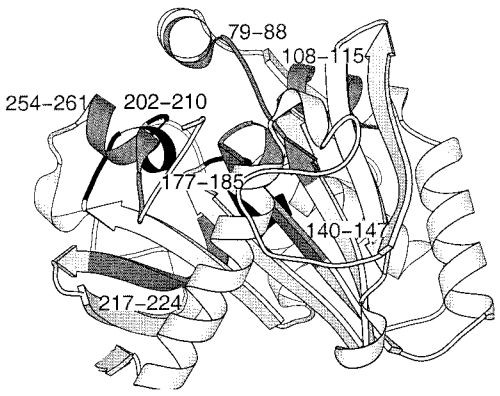

FIGURE 2: Schematic representation of the closed conformation of RmL. Sequence motifs discovered in the RmL family are shaded in gray and black. The figure was generated with Molscript (*79*).

RESULTS AND DISCUSSION

Sequence Motifs in the RmL Family. We identify sequence motifs (highly conserved regions) in the RmL family and search GenPept to find distant family members in order to be able to investigate whether the key residues identified by the electrostatic calculations in this study are generally conserved in lipases. We also study whether the key electrostatic residues are located in sequence motifs, since motifs often describe regions which are necessary for the functionality or stability of the particular protein, and have been conserved throughout evolution. Seven conserved motifs in the RmL family are identified by MEME. These motifs are also found when only three out of four sequences are used for identifying the motifs. The location of the motifs in the closed RmL structure is illustrated in Figure 2.

Searching GenPept with Meta-MEME results in eight statistically significant lipase or esterase sequences with Viterbi scores equal to or higher than 9.7. Meta-MEME uses a threshold of 8.5 as a statistically significant score. An alignment of the motif regions in the RmL family and its homologues is shown in Figure 3. The first sequence motif (residues 79−88) covers the first hinge (residues 83−84) and part of the lid, and contains two fully conserved residues, R80 and G81. The buried R80 is stabilized by the side chain of D61 which lines the polar cavity behind the lid (Figure 1). D61 is conserved in 10 out of 12 lipase homologues. The function of this conserved interaction could possibly be to "lock" the position of the amino acid chain before the mobile first hinge and the lid. The other fully conserved residue in the first motif, G81, has backbone torsion angles ($\varphi = -85°$ and $\psi = -162°$ for the closed RmL) which are allowed only for Gly in the Ramachandran plot. Next to G81 is located S82, which forms the oxyanion hole during the conformational change and stabilizes the tetrahedral intermediates formed in the reaction (*57*). In RmL, PcL, and in

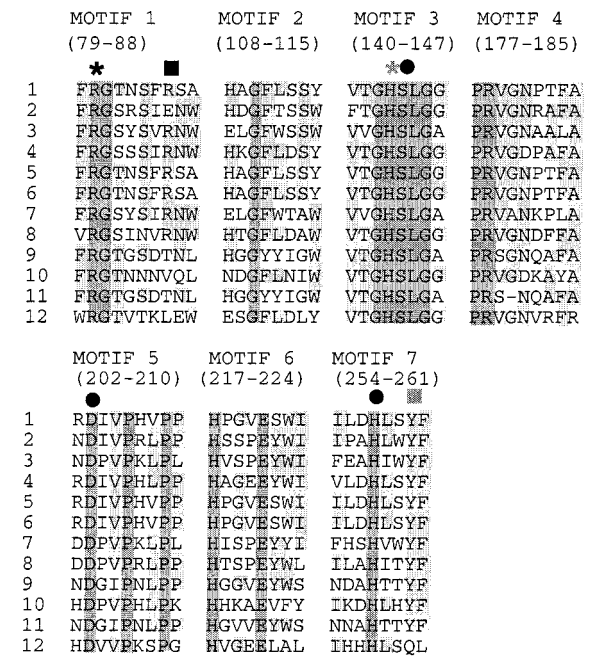

FIGURE 3: Alignment of the motif regions in 12 lipase homologues. Sequence numbering is from RmL. The following residues are marked with symbols: catalytic residues (black circle); R80 (black star); R86 (black square); H143 (gray star); and Y260 (gray square). The alignment contains the following sequences: (1) *Rhizopus delemar* lipase; (2) *Humicola lanuginosa* lipase; (3) *Penicillium camembertii* lipase; (4) *Rhizomucor miehei* lipase; (5) *Rhizopus niveus* lipase; (6) *Rhizopus oryzae* lipase; (7) *Aspergillus oryzae* diacylglycerol lipase; (8) *Fusarium heterosporum* lipase; (9) *Aspergillus niger* ferulic acid esterase; (10) *Caenorhabditis elegans* lipase homologue; (11) *Aspergillus tubingensis* ferulic acid esterase; (12) *Arabidopsis thaliana* lipase isolog. Alignment was generated with Meta-MEME. The gap was introduced in motif 4 on the basis of the HMMer alignment.

HlL, the oxyanion hole is formed by a Ser residue, and in RdL and RnL by Thr (*57*, *72*). It can be seen from Figure 3 that position 82 is occupied invariantly either by a Ser or by a Thr residue. Motif one contains also the positively charged residue in the lid α-helix, R86 (Figure 1), which is conserved in 7 out of 12 sequences. The second motif (residues 108−115) forms the N-terminal of the long α(3) helix behind the lid and partially surrounds the polar cavity, where the polar side of the lid is buried in the open conformation. The fully conserved residue in this motif, G110, is located in a position which has steric constraints caused by the open form of the lid.

The third motif (residues 140−147) contains the catalytic S144 (Figure 1). In the closed RmL, the catalytic triad (S144, D203, H257) is completely buried from solvent under the lid. In the open RmL, the active site is solvent-exposed, and the solvent accessibility increases to 18% for the catalytic S144. S144 is located in the consensus pentapeptide (G-X-S-X-G) characteristic to lipases (*73*). In RmL, the pentapeptide forms a *β*-*ϵ*Ser-α motif in which the catalytic Ser is located in a tight turn between a *β*-strand and an α-helix. In the closed RmL, in HlL, and in RnL, the hydroxyl group of the catalytic S144 forms a hydrogen bond with $N^{\epsilon 2}$ of the catalytic H257, and $N^{\delta 1}$ of H257 is a proton donor to the catalytic D203. The D203−H257 couple enhances the nucleophilicity of the catalytic S144, as has been found in all serine proteases (*57*). D203 accepts a hydrogen bond also from O^{η} of Y260 and from the main chain nitrogen of V205 in both the closed and open RmL, in HlL, and in RnL.

The fully conserved P177 in motif four (residues 177−185) forms partially the bottom of the hydrophobic crevice where the scissile fatty acyl chain is likely to bind in the lipase−substrate complex (*15*, *74*). The fully conserved R178 is buried in RmL, and may play a role in anchoring the P177-containing loop in the right place. R178 is stabilized by hydrogen bonds with E240 and P209. In 9 lipase homologues out of 12, position 240 accommodates either an Asp or a Glu. Motif five (residues 202−210) includes the fully conserved catalytic D203. Residues V205, L208, and P209 of motif five belong to the fatty acyl binding hydrophobic crevice, and residue I204 is part of the hydrophobic dent which binds the *sn*-2 substituent (*15*). The two conserved prolines (P206 and P209) in motif five are likely to reduce the flexibility of loop 201−219. Motif six consists of residues 217−224, and forms most of *β*(7). The two conserved residues in this motif, H217 and E221, are not located in the immediate proximity to the active site. Motif seven (residues 254−261) contains the fully conserved catalytic residue H257, and Y260, which is located close to the active site (Figure 1) and is conserved in 11 out of 12 lipase homologues.

Quantitative Assessment of the Electrostatic Interactions in the RmL Family. We evaluate the electrostatic effect of the mutations on lipase function by comparing pK_a values of the titrating sites of WT and mutants. The following residues were subject to modeled mutations (numbering is from RmL): the catalytic residues (S144, D203, and H257); a few residues which stabilize the active site geometry (H143 and Y28); two charged lid residues (R86 and D91); and one residue close to the first hinge of the lid (R80). The location of these residues is illustrated in Figure 1. We model each mutation for the closed and open conformations of RmL, and for the lipases for which we have relevant experimental data.

Generally, pK_a shifts are caused by changes in Born energies, Coulombic interactions with dipolar groups, or ionic interactions. Born energy is dependent on the solvent-accessible surface; buried charged residues have high Born energies which destabilize the protein. The ionization equilibrium of buried charged residues is often shifted toward the neutral state, unless they are stabilized by oppositely charged residues. Mutations which break existing hydrogen bonds cause large pK_a shifts. We analyze the source of the observed pK_a shifts by comparing solvent-accessible surface areas of each residue, changes in hydrogen bonding patterns, and site−site electrostatic interaction energies. Site−site electrostatic interaction energies ΔG_{ij} (kcal/mol) reveal whether the interaction between two titrating sites is attractive (negative energies) or repulsive (positive energies). Strong interactions can be of the order of magnitude of several kilocalories per mole. Site−site interaction energies also give an idea how much the interaction affects the pK_a value of the particular titrating site, because the pK_a shift of a residue from its model value is a sum of all of its site−site interaction energies.

Catalytic Residues. To identify the key interactions of the catalytic residues, we have modeled the following mutations in RmL, RnL, and HlL: Ser and Asp were mutated into Ala, and His into Phe. Figure 4 shows the calculated pK_a shifts

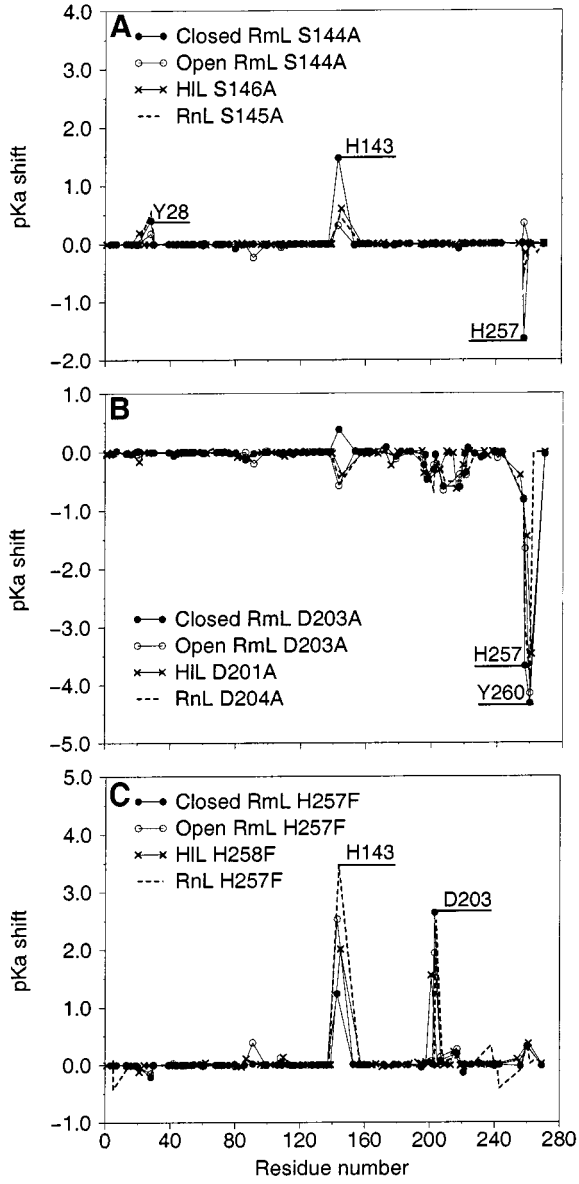

FIGURE 4: pK_a shifts caused by mutations (A) S144A, (B) D203A, and (C) H257F. A pK_a shift depicts the difference between the pK_a values of the corresponding residues in the mutant and wild-type protein.

of the titrating sites in the three lipases due to these mutations. We observe that each of these mutations causes pK_a shifts in only a few residues.

Mutation S144A in motif 3 causes pK_a shifts for residues Y28, H143 (motif 3), and H257 (motif 7) in all three lipases (Figure 4A). The location of Y28, H143, and H257 is represented in Figure 1. These residues are important for the function or stability of the lipase. Figure 4B shows that mutating the catalytic D203 in motif 5 to Ala causes large downward pK_a shifts for H257 and Y260 in motif 7 in all three lipases. The pK_a shift of the catalytic H257 from 3.9 to 0.2 in the closed RmL and from 6.3 to 3.9 in the open RmL is caused by loss of the stabilizing hydrogen bond between catalytic D203 and H257 in the mutant structure. H257 is fully buried in the closed RmL, and has a side chain accessibility of 13% in the open RmL. Buried charged residues with no stabilizing electrostatic interactions tend to have their pK_a values shifted toward the neutral state due to increased penalty in the Born desolvation energy. This explains why the mutation D203A shifts the ionization equilibrium of H257 toward the neutral state, i.e., lower pK_a value. O^{η} of Y260 is a hydrogen bond donor to D203, which shifts the ionization equilibrium of Y260 toward the neutral state, i.e., higher pK_a value. When D203 is mutated to Ala, the pK_a value of Y260 decreases from 16.3 to 12.0 in the closed RmL and from 15.4 to 11.3 in the open one.

Mutation H257F in motif 7 leads to upward pK_a shifts in D203 (motif 5) and H143 (motif 3) in all three lipases (Figure 4C). The upward pK_a shifts of D203 from −1.1 to 1.5 in the closed RmL and from −0.8 to 1.2 in the open RmL are caused by loss of the hydrogen bond between H257 and D203 in the mutant structure. This shifts the ionization equilibrium of D203 toward the neutral state, i.e., higher pK_a value. The mutation increases the pK_a value of H143 from 4.1 to 5.3 in the closed RmL and from 1.4 to 3.9 in the open one. H143 is located in the immediate proximity to H257, the distance between the two imidazole rings being ∼4 Å in RmL, and the two form a repulsive interaction with each other (2.2 kcal/mol in open RmL). The mutation removes this interaction and causes an increase in the pK_a value of H143.

Charged Hinge and Lid Residues. The lid of RmL consists of an α-helix (residues 85−91) and two hinge regions [first hinge 83−84 and second hinge 91−95, according to Derewenda et al. (*6*)]. In the RmL family, the lid has two charged residues, R86 and D91. HlL has two charged residues in the lid, R84 and E87, which are located in structurally different positions in the lid than in the other RmL family members. A third charged residue, R80, is located just before the first hinge in all RmL family members. We elaborate the role of the charged residues in the lid region by modeling mutations R80L, R86L, and D91N for the closed and open RmL. Mutation D92N is also modeled for RnL. For HlL, we model mutations R81L, R84L, and E87A.

Figure 5A shows that mutation R80L in motif 1 causes three pK_a shifts in the closed and open conformations of RmL and in HlL: D61, Y28, and H143 (motif 3). D61 lines the polar cavity behind the lid (Figure 1). In the closed RmL and HlL, D61 is hydrogen bonded to R80, and the residues have a strong attractive electrostatic interaction energy (−1.18 kcal/mol for the closed RmL). When R80 is mutated, the stabilizing interaction between D61 and R80 is lost, leading to an increase in the pK_a value of D61 from 2.8 to 3.9. Y28 has an attractive electrostatic interaction with R80, and loss of this interaction shifts the ionization equilibrium of Y28 toward its neutral state, i.e., higher pK_a value. The mutation also increases the pK_a value of H143 from 4.1 to 5.3 in the closed RmL. H143 is located ∼10 Å away from R80, and is separated from R80 by Y28. The site−site electrostatic interaction energies show that H143 has a fairly weak repulsive interaction with R80 (0.56 kcal/mol for the closed RmL) and a strong attractive interaction with Y28 (−4.88 kcal/mol for the closed RmL) due to a hydrogen bond. It is likely that the pK_a shift of H143 is caused by the loss of the weak interaction with R80, and partially by the pK_a shift of Y28 which affects the attractive interaction between Y28 and H143. This is an example of a significant long-range electrostatic effect caused by a combination of a direct long-range effect and a network of electrostatic interactions. A pK_a shift is observed for R30 in the closed

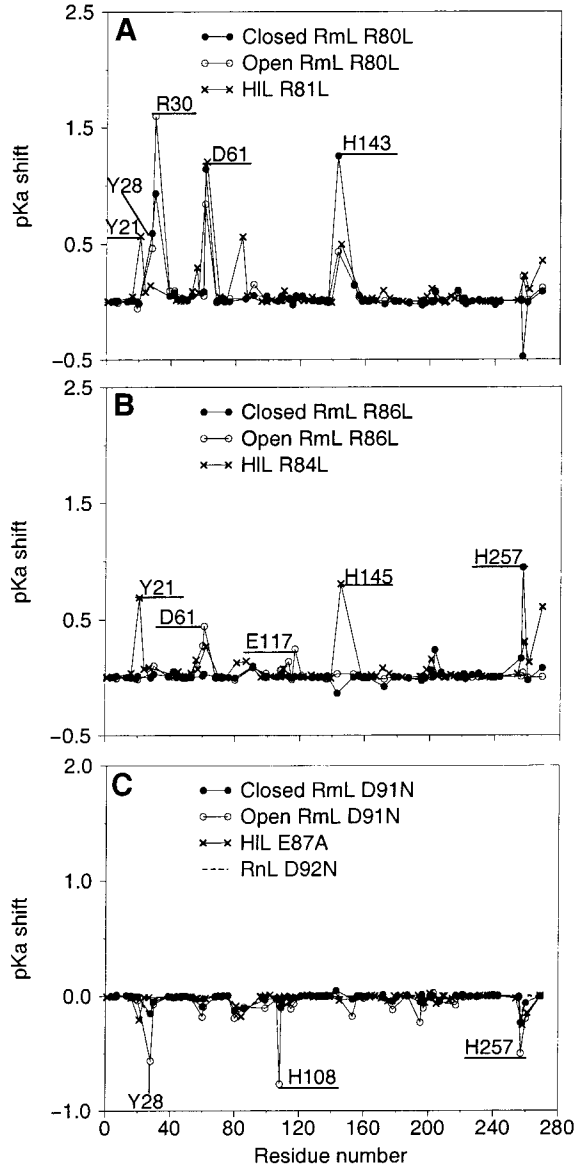

FIGURE 5: pK_a shifts caused by mutations (A) R80L, (B) R86L, and (C) D91N. Note in (A) that Y21 in HlL corresponds to Y28 in RmL.

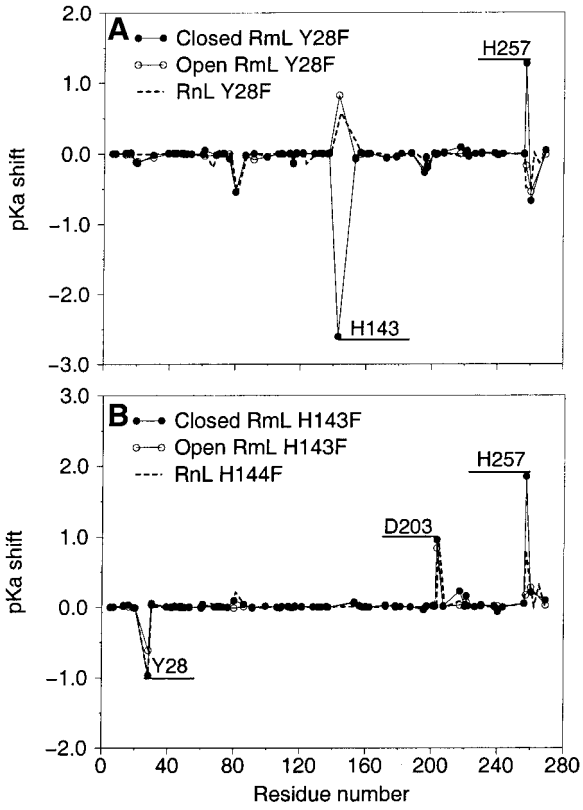

FIGURE 6: pK_a shifts caused by mutations (A) Y28F and (B) H143F.

and open RmL. R30 is located close to the first hinge, and it forms a repulsive interaction with R80 (1.03 kcal/mol in the closed RmL). Mutation of R80 leads to a loss of this repulsive interaction and to an increase in the pK_a value of R30 from 12.1 to 13.7 in the closed RmL. In HlL, a pK_a shift is also observed for the positively charged residue in the lid, R84. Mutation removes the repulsive electrostatic interaction between R81 and R84 (0.97 kcal/mol), and leads to an increase in the pK_a value for R84 from 14.5 to 15.0.

It can be seen from Figure 5B that mutation R86L in motif 1 causes only one significant pK_a shift in the closed RmL: the pK_a value of the catalytic H257 (motif 7) increases from 3.9 to 4.9. In the closed RmL, R86 and H257 have a fairly small repulsive interaction of 0.33 kcal/mol. In the open RmL, small upward pK_a shifts are observed for D61 and E117. D61 and E117 both line the polar cavity behind the lid, and, thus, the interactions between R86 and D61 and E117 stabilize the open conformation. Chemical modification of arginines in RmL has been observed to lead to a partial loss of activity, residual activity being 66/46% depending on substrate (*16*). This could be caused by the pK_a shift in H257 in the closed conformation or by loss of the stabilizing interactions between R86 and D61 and E117 in the open conformation. In HlL, mutation R84L affects pK_a values of Y21 and H145 (Figure 5B), suggesting that R84 may play a different role in HlL than R86 in RmL.

Mutation D91N causes three significant pK_a shifts in the open RmL: Y28, H108 (motif 2), and H257 (motif 7). The reason for the conformation-specific interaction of D91 can be easily understood by comparing the closed and open conformations of RmL (Figure 1). In the closed conformation, the side chain of D91 is over 90% solvent-exposed and is directed away from the active site, whereas in the open conformation it is only 21% solvent-exposed and is directed toward the active site. The distances between D91 and Y28, H108, and H257 are under 8 Å in the open conformation.

Y28F and H143F. Residues Y28 and H143 are located close to the catalytic residues (Figure 1). H143 is located next to the catalytic S144 in the tight β-ϵSer-α turn in motif 3, and Y28 is conserved in 11 out of 12 lipase homologues. $N^{\delta 1}$ of H143 forms a hydrogen bond with the hydroxyl group of Y28. According to experimental data, mutating either Y28 or H143 to Phe leads to a complete loss of catalytic activity in RoL (*20*). Beer et al. (*20*) have proposed that in RoL this is caused by loss of the hydrogen bond between Y28 and H143, which stabilizes the active site geometry. We investigate the effect of mutations Y28F and H143F on the electrostatics in RmL and RnL. Figure 6A shows that mutation Y28F leads to a 2.6 unit decrease in the pK_a value of H143 in the closed RmL; the pK_a value of H143 decreases from 4.1 to 1.5. H143 is fully buried, and loss of this stabilizing interaction drives its ionization equilibrium toward

the neutral state. Mutation H143F causes a pK_a shift in Y28 in RnL and in both conformations of RmL (Figure 6B). The pK_a shifts observed for Y28 and H143 could lead to a destabilization of the active site geometry and a loss of activity. However, the mutation Y28F also causes a pK_a shift in H257 in the closed RmL and RnL, and the mutation H143F causes pK_a shifts in catalytic D203 and H257 in RmL and in RnL.

Network of Electrostatic Interactions in the RmL Family. Our pK_a calculations of the WT and mutated lipases belonging to the RmL family indicate an existence of a structurally and functionally important network of electrostatic interactions. The location of residues belonging to the network, S144, D203, H257, Y260, H143, Y28, R80, and D91, is represented in Figure 1. These residues interact with each other, which is observed in the modeled mutant structures in the form of pK_a shifts (Figures 4−6) and changes in site−site electrostatic interaction energies. For instance, we observe that mutation S144A causes changes in site−site electrostatic interaction energies for Y28−R80, Y28−H143, Y28−H257, R80−H143, H143−H257, and D203−H257 in the open RmL, in HlL, and in RnL. In conclusion, mutating residues belonging to the electrostatic network cause pK_a shifts and changes in site−site electrostatic interaction energies mainly in the network itself.

The functional importance of mutating some residues belonging to this network has been tested experimentally (*20*, *23*). Mutations S146A (catalytic Ser) for HlL (*23*), and Y28F, H143F, and D91N for RoL (*20*) have been observed to affect lipase activity significantly. In addition, the fact that the network S144-D203-H257-Y260-H143-Y28-R80-D91 involves the three catalytic residues makes it clear that these interactions are functionally important.

Alignment of the sequence motifs in the six RmL family members with six homologues (Figure 3) shows that residues belonging to the network are either fully or very well conserved and that most of them are located in sequence motifs discovered by MEME. Y28 is conserved in 11 out of 12 sequences (data not shown), R80 is fully conserved, D91 is conserved in 9 sequences (data not shown), H143, S144, D203, and H257 are fully conserved, and Y260 is conserved in 11 sequences. The network residues R80, H143, S144, D203, H257, and Y260 are located in sequence motifs.

Electrostatic Free Energies of the Two Conformations. Figure 7 shows that the mutations R86L, D91N, and H257F affect the pH-dependent electrostatic free energies (G_{elec}, kcal/mol) of the closed and open RmL differently. For the other mutations modeled, no significant differences were observed in electrostatic free energies between the closed and open RmL. The optimum pH range for the hydrolytic activity of RmL has been observed to be between 7 and 8, depending on the substrate (*75*, *76*). Figure 7A shows that for the mutation R86L the difference in electrostatic free energy between R86L and WT is negative in the closed conformation at pH 7 but positive in the open one, indicating that R86 destabilizes the closed conformation but stabilizes the open one. The solvent-accessible surface area of the side chain of R86 is 65% in the closed conformation and 64% in the open one, suggesting that it is not the change in accessibility which causes the difference in the ΔG_{elec}. We compare the site−site electrostatic interaction energies of the closed and open conformations to find possible Coulombic

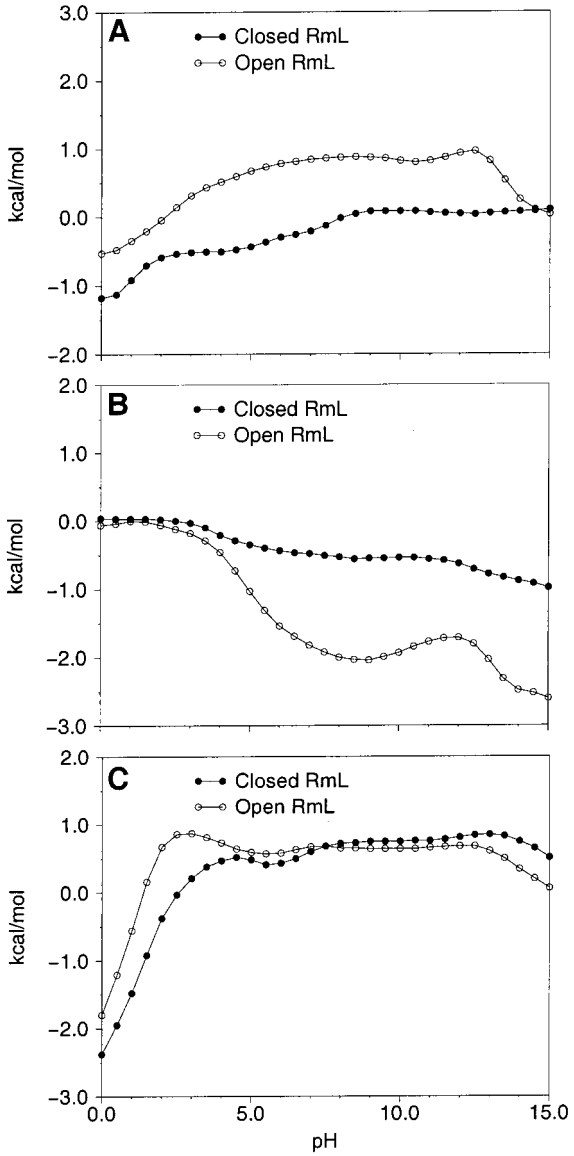

FIGURE 7: Difference in electrostatic free energies (kcal/mol) between mutants (A) R86L, (B) D91N, and (C) H257F and the wild type for the closed and open conformations of RmL.

interactions of R86 which could stabilize the open RmL. It is evident from Figure 8 that interactions between R86 and D61, D113, and E117 are more attractive in the open conformation than in the closed one. Residues D61, D113, and E117 surround the polar cavity behind the lid and stabilize the open conformation. Further, interactions between R86 and H143 and H257 are less repulsive in the open conformation than in the closed one. These Coulombic interactions contribute to the observed difference in the electrostatic free energy.

For the mutation D91N, the difference in the electrostatic free energy between mutant and WT is negative in both conformations (Figure 7B), suggesting that D91 destabilizes both conformations. However, the degree of destabilization caused by D91 is significantly higher in the open conformation. In the closed conformation, D91 is over 90% solvent accessible, whereas in the open conformation the solvent accessibility of its side chain is only 21%. Thus, in the open conformation, the energy penalty of desolvating this residue (Born energy) is significantly higher than in the closed

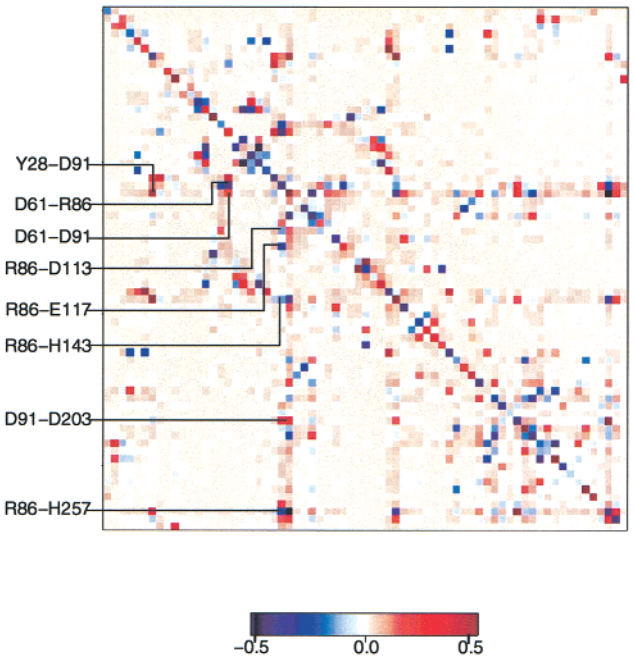

FIGURE 8: Site−site electrostatic interaction energy map. The map shows the difference in the interaction energy between open and closed RmL for every pair of titrating sites. The blue color indicates a negative difference and red positive. Energies are in kcal/mol.

conformation, and mutating D91 to Asn reduces the high Born energy. The site−site electrostatic interaction energy map (Figure 8) shows that interactions between D91 and Y28, D61, and D203 are more repulsive in the open conformation than in the closed one, and make it relatively more destabilized.

For the mutation H257F, a difference in the free energies can be observed in low pH ranges, 0−4.0 (Figure 7C). Because the difference is negligible at higher pH than 4.0, it is probable that the observed difference in free energies is caused by changes in the ionization states of some residues. Histidine has a positive charge at a low pH which increases its Born desolvation energy. Since H257 is fully buried in the closed RmL, but has a 13% accessibility in the open one, its Born desolvation energy is higher in the closed one. Thus, mutating H257 into Phe has a greater effect on the Born energy of H257 in the closed conformation.

CONCLUSIONS

We have investigated the role of electrostatic interactions in the fungal RmL family by identifying conserved sequence motifs and by performing pK_a calculations. We use a protein dielectric of 20 in the electrostatic calculations since it has been observed to yield the best agreement with experimental data (*34*). This value reflects the presence of polar amide and other polarizable groups in the solvent-inaccessible protein interior. When a protein dielectric of 20 is used, an accuracy of approximately 1 pH unit is reached (*38*). We emphasize that the aim of this study is more on characterizing the structurally and functionally important electrostatic interactions in lipases rather than estimating the exact pK_a value of each titrating group.

We observe that each of the mutations modeled causes only a few significant pK_a shifts. These pK_a shifts vary between ±0.5 and ±4.5 pH units, being way above of the observed 0.01−0.05 pH unit accuracy for experimentally determined pK_a values (*77*, *78*). This suggests that the pK_a shifts reported in this study can be verified by experiments. In most cases, the residue undergoing a significant pK_a shift is located in proximity to the mutated residue. However, we also observe pK_a shifts in residues as far as 10 Å away from the mutated residue. Site−site electrostatic interaction energies show that charged residues can have long-range electrostatic effects, especially in the low dielectric protein interior, and that electrostatic networks can cause pK_a shifts in residues located far away from the mutated residue.

The pK_a shifts and site−site electrostatic interaction energies indicate an existence of a functionally important network of electrostatic interactions. This network includes residues from the active site to the lid region, and, thus, connects the catalytic triad with the mobile lid associated with the conformational change. The network residues are fully or very well conserved in a set of 12 lipase homologues, and 6 out of 8 network residues are located in sequence motifs.

In this study, we explore the effect of the modeled mutations on the pH-dependent stability of the closed and open conformations of RmL in terms of electrostatic free energies. We find that electrostatic interactions of the positively charged arginine in the lid with negatively charged residues around the polar cavity behind the lid stabilize the open conformation. These residues are potential candidates for mutational studies to investigate their roles in the stability of RmL.

In conclusion, we have used sequence-derived information and structure-based biophysical computations to characterize the structure−function relationship in a family of fungal lipases. By combining information about the conserved sequence motifs together with the information obtained from the electrostatic calculations, we are able to characterize a conserved network of structurally and functionally important electrostatic interactions in this lipase family. The existence of this network suggests that polar and charged residues not only have pairwise interactions in proteins but also form structurally and functionally important networks which can connect distant parts in proteins.

ACKNOWLEDGMENT

We are grateful for Per Jambeck for help with bioinformatics methods and critical reading of the manuscript, and Markus Herrgård for help with computations. Dennis Livesay is acknowledged for useful discussions. We thank Professor Andrew McCammon for providing us the University of Houston Brownian Dynamics suite of programs.

REFERENCES

1. Vulfson, E. N. (1994) in *Lipases—Their Structure, Biochemistry and Application* (Woolley, P., and Petersen, S. B., Eds.) pp 271−288, Cambridge University Press, Cambridge.
2. Sarda, L., and Desnuelle, P. (1958) *Biochim. Biophys. Acta 30*, 513−521.
3. Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L., and Menge, U. (1990) *Nature 343*, 767−770.
4. Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A., and Thim, L. (1991) *Nature 351*, 491−494.

5. Derewenda, Z. S. (1994) *Adv. Protein Chem. 45*, 1−52.
6. Derewenda, U., Brzozowski, A. M., Lawson, D. M., and Derewenda, Z. S. (1992) *Biochemistry 31*, 1532−1541.
7. Norin, M., Olsen, O., Svendsen, A., Edholm, O., and Hult, K. (1993) *Protein Eng. 6*, 855−863.
8. Norin, M., Haeffner, F., Hult, K., and Edholm, O. (1994) *Biophys. J. 67*, 548−559.
9. Peters, G. H., Olsen, O. H., Svendsen, A., and Wade, R. C. (1996) *Biophys. J. 71*, 119−129.
10. Peters, G. H., van Aalten, D. M. F., Edholm, O., Toxvaerd, S., and Bywater, R. (1996) *Biophys. J. 71*, 2245−2255.
11. Peters, G. H., Toxvaerd, S., Olsen, O. H., and Svendsen, A. (1997) *Protein Eng. 10*, 137−147.
12. Peters, G. H., van Aalten, D. M. F., Svendsen, A., and Bywater, R. (1997) *Protein Eng. 10*, 149−158.
13. Orrenius, C., van Heusden, C., van Ruiten, J., Overbeeke, P. L. A., Kierkels, H., Duine, J. A., and Jongejan, J. A. (1998) *Protein Eng. 11*, 1147−1153.
14. Jääskeläinen, S., Verma, C. S., Hubbard, R. E., Linko, P., and Caves, L. S. D. (1998) *Protein Sci. 7*, 1359−1367.
15. Scheib, H., Pleiss, J., Kovac, A., Paltauf, F., and Schmid, R. D. (1999) *Protein Sci. 8*, 215−221.
16. Holmquist, M., Norin, M., and Hult, K. (1993) *Lipids 28*, 721−726.
17. Holmquist, M., Martinelle, M., Berglund, P., Clausen, I. G., Patkar, S., Svendsen, A., and Hult, K. (1993) *J. Protein Chem. 12*, 749−757.
18. Joerger, R. D., and Haas, M. J. (1994) *Lipids 29*, 377−384.
19. Holmquist, M., Clausen, I. G., Patkar, S., Svendsen, A., and Hult, K. (1995) *J. Protein Chem. 14*, 217−224.
20. Beer, H. D., Wohlfahrt, G., McCarthy, J. E. G., Schomburg, D., and Schmid, R. D. (1996) *Protein Eng. 9*, 507−517.
21. Martinelle, M., Holmquist, M., Clausen, I. G., Patkar, S., Svendsen, A., and Hult, K. (1996) *Protein Eng. 9*, 519−524.
22. Svendsen, A., Clausen, I. G., Patkar, S. A., Borch, K., and Thellersen, M. (1997) *Methods Enzymol. 284*, 317−340.
23. Berg, O. G., Cajal, Y., Butterfoss, G. L., Grey, R. L., Alsina, M. A., Yu, B.-Z., and Jain, M. K. (1998) *Biochemistry 37*, 6615−6627.
24. Peters, G. H., Svendsen, A., Langberg, H., Vind, J., Patkar, S. A., Toxvaerd, S., and Kinnunen, P. K. J. (1998) *Biochemistry 37*, 12375−12383.
25. Mandel, M. (1964) *Proc. Natl. Acad. Sci. U.S.A. 52*, 736−741.
26. Mandel, M. (1965) *J. Biol. Chem. 240*, 1586−1592.
27. Meadows, D. H., Markley, J. L., Cohen, J. S., and Jardetzky, O. (1967) *Proc. Natl. Acad. Sci. U.S.A. 58*, 1307−1313.
28. Richarz, R., and Wüthrich, K. (1978) *Biochemistry 17*, 2263−2269.
29. Clore, G. M., and Gronenborn, A. M. (1997) *Nat. Struct. Biol. 4*, 849−853.
30. Riek, R., Wider, G., Pervushin, K., and Wüthrich, K. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 4918−4923.
31. Antosiewicz, J., and Porschke, D. (1989) *Biochemistry 28*, 10072−10078.
32. Beroza, P., Fredkin, D. R., Okamura, M. Y., and Feher, G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5804−5808.
33. Gilson, M. K. (1993) *Proteins: Struct., Funct., Genet. 15*, 266−282.
34. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994) *J. Mol. Biol. 238*, 415−436.
35. Yang, A.-S., Gunner, M. R., Sampogna, R., Sharp, K., and Honig, B. (1993) *Proteins: Struct., Funct., Genet. 15*, 252−265.
36. Bashford, D., Case, D. A., Dalvit, C., Tennant, L., and Wright, P. E. (1993) *Biochemistry 32*, 8045−8056.
37. You, T. J., and Bashford, D. (1995) *Biophys. J. 69*, 1721−1733.
38. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1996) *Biochemistry 35*, 7819−7833.
39. Gibas, C. J., and Subramaniam, S. (1996) *Biophys. J. 71*, 138−147.
40. Trylska, J., Antosiewicz, J., Geller, M., Hodge, C. N., Klabe, R. M., Head, M. S., and Gilson, M. K. (1999) *Protein Sci. 8*, 180−195.
41. Raquet, X., Lounnas, V., Lamotte-Brasseur, J., Frère, J. M., and Wade, R. C. (1997) *Biophys. J. 73*, 2416−2426.
42. Dillet, V., Dyson, H. J., and Bashford, D. (1998) *Biochemistry 37*, 10298−10306.
43. Lamotte-Brasseur, J., Lounnas, V., Raquet, X., and Wade, R. C. (1999) *Protein Sci. 8*, 404−409.
44. Warwicker, J. (1992) *J. Mol. Biol. 223*, 247−257.
45. Yang, A.-S., and Honig, B. (1993) *J. Mol. Biol. 231*, 459−474.
46. Yang, A.-S., and Honig, B. (1994) *J. Mol. Biol. 237*, 602−614.
47. Gibas, C. J., Subramaniam, S., McCammon, J. A., Braden, B. C., and Poljak, R. J. (1997) *Biochemistry 36*, 15599−15614.
48. Schaefer, M., Sommer, M., and Karplus, M. (1997) *J. Phys. Chem., Ser. B 101*, 1663−1683.
49. Elcock, A. H., and McCammon, J. A. (1998) *J. Mol. Biol. 280*, 731−748.
50. van Vlijmen, H. W. T., Curry, S., Schaefer, M., and Karplus, M. (1998) *J. Mol. Biol. 275*, 295−308.
51. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol. 247*, 536−540.
52. Martinelle, M., Holmquist, M., and Hult, K. (1995) *Biochim. Biophys. Acta 1258*, 272−276.
53. Uppenberg, J., Öhrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., and Jones, T. A. (1995) *Biochemistry 34*, 16838−16851.
54. Holm, L., and Sander, C. (1996) *Science 273*, 595−602.
55. Grundy, W. N., Bailey, T. L., Elkan, C. P., and Baker, M. E. (1997) *Comput. Appl. Biosci. 13*, 397−406.
56. Bailey, T. L., and Elkan, C. (1994) in *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp 28−36, AAAI Press, Menlo Park, CA.
57. Derewenda, Z. S., Derewenda, U., and Dodson, G. G. (1992) *J. Mol. Biol. 227*, 818−839.
58. Derewenda, U., Swenson, L., Wei, Y., Green, R., Kobos, P. M., Joerger, R., Haas, M. J., and Derewenda, Z. S. (1994) *J. Lipid Res. 35*, 524−534.
59. Derewenda, U., Swenson, L., Green, R., Wei, Y., Dodson, G. G., Yamaguchi, S., Haas, M. J., and Derewenda, Z. S. (1994) *Nat. Struct. Biol. 1*, 36−46.
60. Eddy, S. R. (1995) in *Proceedings of the Third International Conference on Intelligent Systems for Molecular Biology*, pp 114−120, AAAI Press, Menlo Park, CA.
61. Kohno, M., Funatsu, J., Mikami, B., Kugimiya, W., Matsuo, T., and Morita, Y. (1996) *J. Biochem. 120*, 505−510.
62. Brünger, A. T., and Karplus, M. (1988) *Proteins: Struct., Funct., Genet. 4*, 148−156.
63. Dunbrack, R. L., Jr., and Karplus, M. (1993) *J. Mol. Biol. 230*, 543−574.
64. Hubbard, S. J., Campbell, S. F., and Thornton, J. M. (1991) *J. Mol. Biol. 220*, 507−530.
65. McDonald, I. K., and Thornton, J. M. (1994) *J. Mol. Biol. 238*, 777−793.
66. Madura, J. D., Davis, M. E., Gilson, M. K., Wade, R. C., Luty, B. A., and McCammon, J. A. (1994) in *Reviews in Computational Chemistry, Volume V* (Lipkowitz, K. B., and Boyd, D. B., Eds.) pp 229−267, VCH Publishers, Inc., New York.
67. Madura, J. D., Briggs, J. M., Wade, R. C., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J., Gilson, M. K., Bagheri, B., Scott, L. R., and McCammon, J. A. (1995) *Comput. Phys. Commun. 91*, 57−95.
68. Davis, M. E., Madura, J. D., Luty, B. A., and McCammon, J. A. (1991) *Comput. Phys. Commun. 62*, 187−197.
69. Jorgensen, W. L., and Tirado-Rives, J. (1988) *J. Am. Chem. Soc. 110*, 1657−1666.
70. Shrake, A., and Rupley, J. A. (1973) *J. Mol. Biol. 79*, 351−371.
71. Gilson, M. K., Sharp, K. A., and Honig, B. H. (1987) *J. Comput. Chem. 9*, 327−335.

72. Derewenda, U., Swenson, L., Green, R., Wei, Y., Yamaguchi, S., Joerger, R., Haas, M. J., and Derewenda, Z. S. (1994) *Protein Eng. 7*, 551−557.
73. Derewenda, Z. S., and Derewenda, U. (1991) *Biochem. Cell Biol. 69*, 842−851.
74. Lawson, D. M., Brzozowski, A. M., Rety, S., Verma, C., and Dodson, G. G. (1994) *Protein Eng. 7*, 543−550.
75. Huge-Jensen, B., Galluzzo, D. R., and Jensen, R. G. (1987) *Lipids 22*, 559−565.
76. Wu, X. Y., Jääskeläinen, S., and Linko, Y.-Y. (1996) *Appl. Biochem. Biotechnol. 59*, 145−158.
77. Inagaki, F., Kawano, Y., Shimada, I., Takahashi, K., and Miyazawa, T. (1981) *J. Biochem. 89*, 1185−1195.
78. Lockhart, D. J., and Kim, P. S. (1993) *Science 260*, 198−202.
79. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI9916980