Exceptionally Stable Salt Bridges in Cytochrome P450cam Have Functional Roles[†]

Valère Lounnas and Rebecca C. Wade*

European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 10-2209, 69012 Heidelberg, Germany

Received September 11, 1996; Revised Manuscript Received December 23, 1996[⊗]

ABSTRACT: A long-standing puzzle in structure−function studies of cytochrome P450cam is how the substrate, camphor, reaches the buried active site. The crystal structure shows no channel from the surface to the active site large enough for substrate to pass through. Recent experiments indicate that access of the rather nonpolar substrate to the active site is controlled by electrostatic interactions and may involve rupture of the two salt links to Asp251 [Deprez, E., Gerber, N. C., Di Primo, C., Douzou, P., Sligar, S. G., & Hui Bon Hoa, G. (1994) *Biochemistry 33*, 14464−14468]. Consequently, we have computed the electrostatic strength of 53 ionic pairs, including 32 salt links, in cytochrome P450cam by numerical solution of the finite-difference linearized Poisson−Boltzmann equation. The calculated electrostatic free energies, ΔG_{tot}, of the salt links range from −9 to +6 kcal/mol with approximately 60% of the salt links being energetically favorable and 40% being unfavorable with respect to mutation to their uncharged, nonpolar isosteres. Strikingly, of the four most stable salt links in the protein ($\Delta G_{tot} < -6$ kcal/mol), two involve the propionate groups of the heme and the other two involve Asp251. In the modeled D251N mutant, for which electrostatic effects on substrate binding are diminished, the latter two salt links lose their stability ($\Delta G_{tot} > -2.4$ kcal/mol). Thus it appears that cytochrome P450cam has evolved four unusually strong salt bridges, stabilized by surrounding charged and polar groups in the protein, to keep its heme cofactor in place and to regulate substrate binding.

(A) Cytochrome P450. Cytochrome P450 enzymes are hemoprotein monooxygenases that play a critical role in the synthesis and degradation of many physiologically important compounds and xenobiotics [for reviews, see FASEB (1992, 1996)]. Because of their ability to catalyze stereo- and regioselective hydroxylation of hydrocarbons, cytochrome P450s are also potentially attractive catalysts for the degradation of environmental pollutants and for biotechnological drug production. The bacterial cytochrome P450cam (P450cam)[1] from *Pseudomonas putida* catalyzes the hydroxylation of its natural substrate, camphor, and also reactions on many other compounds. It has been more extensively characterized by biochemical and biophysical techniques than other cytochrome P450s and has thus served as a prototype for cytochrome P450 structure−function studies. Its three-dimensional atomic structure is known from X-ray crystallographic studies by Poulos and co-workers [for review, see Poulos and Raag (1992)]. The protein structure has been solved at high resolution in the uncomplexed form and in the presence of a number of substrates and inhibitors. The active site of P450cam is completely buried in the protein interior (see Figure 1), and there are no channels large enough for camphor to pass through connecting the active site to the protein surface. Thus the crystal structure shows that the protein must undergo significant motion in order to allow substrate access and product exit. However, there is

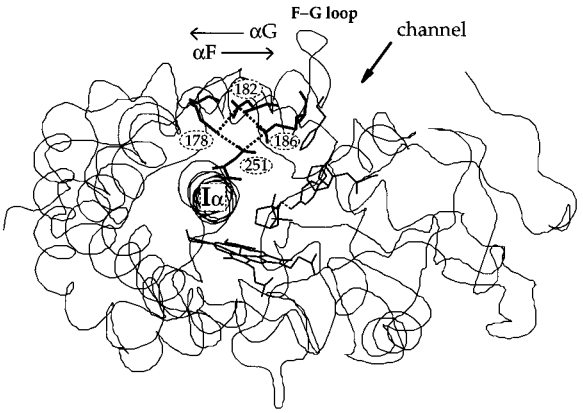


FIGURE 1: Ribbon diagram of cytochrome P450cam showing the location of the buried active site with camphor bound, the entrance of the proposed substrate access channel, the F−G loop, the F, G, and I α-helices, and the salt-link tetrad involving Asp251.

no direct evidence of the route by which ligands pass to and from the active site. Examination of the crystal structure shows two small pockets that might, with protein motion, enlarge to form channels to and from the active site: Poulos and co-workers (Poulos, 1995; Poulos et al., 1987) suggested that the more hydrophobic of these (lined by Phe87 and Phe98) opens up to allow substrate into the active site while the more hydrophilic channel could be the route by which the product leaves the enzyme. The recently solved structure of cytochrome P450BM-3 (Ravichandran et al., 1993) is very similar to that of P450cam but contains a wide channel connecting the active site to the exterior of the protein in the equivalent position to the proposed substrate access channel in P450cam. The F and G helices and the F−G loop, which line this channel, are much more flexible in cytochrome P450BM-3 (*B*-factors up to 120 Å^2) and also in cytochrome P450terp (where the F−G loop is disordered)

[†] This work was supported in part by the European Union Biotechnology Programme, Contract BIO2-CT94-2060.

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.

[1] Abbreviations: D251N, mutant of cytochrome P450cam in which aspartate 251 is replaced by an asparagine residue; NMR, nuclear magnetic resonance; NVE and NVT, values kept constant during molecular dynamics simulation (N, number of particles; V, volume; E, total energy; T, temperature); P450cam, cytochrome P450cam; rmsd, root mean square deviation; SASA, solvent-accessible surface area; wt, wild-type protein.

S0006-2960(96)02294-5 CCC: $14.00

(Hasemann et al., 1994) than in P450cam (backbone atoms *B*-factors of 20−30 Å^2). Indeed the F−G helix−loop−helix could be considered as a flap that controls the size and dynamics of the proposed active-site entrance channel.

Adjacent to this proposed entrance channel in P450cam, there is a tetrad of charged residues, Asp251, Lys178, Arg186, and Asp182, which form a ring of four salt links (see Figure 1). We will refer to these residues as the entrance channel salt-link tetrad. Two of these salt links, Asp251−Arg186 and Asp251−Lys178, connect the F and I helices. Recent experiments indicate that the passage of camphor to and from the active site of P450cam is modulated, if not regulated, by electrostatic interactions involving residue 251 (Deprez et al., 1994). The increase of solvent ionic strength was found to favor binding, whereas the reduction of the solvent dielectric constant by addition of glycerol had the opposite effect. Passage of camphor to the active site was enhanced for the D251N mutant compared to wild type P450cam. The effects of solvent ionic strength and dielectric constant were much smaller for the D251N mutant but still present, indicating that other electrostatic interactions play a role in the binding of substrate. These data indicate that the salt-link tetrad in P450cam may act as an anchor keeping the F−G flap in place over the I-helix and only allowing substrate access to the active site when the tetrad salt links are perturbed.

Residue 251 is conserved among cytochrome P450s as an Asp or Glu. In P450cam, it has a role not only in substrate binding but also in oxygen activation (Gerber & Sligar, 1994). According to the available three-dimensional structures of other cytochrome P450s whose structures have been solved, there are no salt-link tetrads equivalent in terms of location in sequence and three-dimensional space to the entrance channel salt-link tetrad in P450cam. However, in cytochrome P450eryf, the side chain of Arg 185 in the G helix makes several hydrogen bonds to the protein core that could influence motion of the F−G flap in a similar manner to the salt links to Asp251 in P450cam. This suggests that the control of protein dynamics permitting substrate access to the active site is tuned in each cytochrome P450 to the particular substrate it acts upon. For large substrates, such as that of cytochrome P450BM-3, sufficient desolvation of the catalytic site may be achieved without the need to bury the active site from bulk solvent. On the other hand, for smaller substrates such as that of P450cam, the protein may play a more dominant role in isolating the catalytic site from bulk solvent, which in turn requires more protein dynamics to permit substrate into the active site and more control of protein dynamics to keep the active site buried when substrate is bound.

(B) Salt Links. More generally, the effect of salt links on protein stability has been investigated with both experimental and theoretical approaches (Anderson et al., 1990; Barlow & Thornton, 1983; Dao-pin et al., 1991; Fersht, 1972; Gandini et al., 1996; Hendsch & Tidor, 1994; Horowitz et al., 1990; Lumb & Kim, 1995; Lyu et al., 1992; Musafia et al., 1995; Perutz, 1978; Schueler & Margalit, 1995; Waldburger et al., 1995; Warshel & Russel, 1984; Wimley et al., 1996). Experimentally, the contribution of salt links to stability has been studied with NMR measurements as a function of pH and by mutation of the salt-link residues to alanine or to residues that are nonpolar but approximately isosteric. These experiments use different reference states from which the salt link is absent to assess the contribution of salt links to protein stability. The results indicate that, relative to the chosen reference state, salt links can either stabilize or destabilize a protein's folded state.

An example of a single salt link making a sizable contribution to protein stability is Asp70−His31 in T4 lysozyme, which was found, in NMR studies of pH-induced denaturation, to stabilize the protein by 3−5 kcal/mol (Anderson et al., 1990). However, attempts to create salt bridges on the surface of T4 lyzozyme by engineering six mutants resulted in, at most, 0.5 kcal/mol increase in stability (Dao-pin et al., 1991). Moreover, Lumb and Kim (1995) found that native salt links could be unfavorable relative to neutral-charge interactions for bZIP dimerization via leucine zippers. Some of the factors identified from experiments as important for salt-link stability are (i) exposure to solvent; (ii) correct alignment of the residues participating in the salt link by the folded protein (Dao-pin et al., 1991); (iii) cooperative interactions between adjacent salt links, such as those observed in a salt-link triad (two salt links involving three residues) in barnase (Horowitz et al., 1990); and (iv) steric packing complementarity. This was recently observed by Sauer and colleagues (Waldburger et al., 1995), who performed both combinatorial mutagenesis and alanine screening on a buried salt-link triad in the Arc repressor and assessed stability with respect to urea denaturation. On one hand, active mutants, in which all three residues of the triad were simultaneously mutated to approximately isosteric and nonpolar residues, were more stable than the wild-type protein by 1.5−4.5 kcal/mol, indicating that the electrostatic interactions of the triad are globally unfavorable to the overall protein stability. On the other hand, single, double, and triple mutations of the triad residues to alanines resulted in systematically less stable proteins than the wild type as they introduced destabilizing packing defects in the protein interior.

The latter study shows how different contributions to stability can be investigated in different experiments. Protein engineering experiments usually probe, to a greater or lesser extent, several physical contributions to stability at the same time: electrostatic, steric, and entropic. It is difficult to design mutations to investigate each of these contributions individually. An advantage of the theoretical electrostatic continuum approach is that point mutations of charged residues to their hypothetical hydrophobic nonpolar isosteres can be made and thus purely electrostatic contributions to salt-link stabilities can be evaluated. In addition, electrostatic continuum calculations employing the linearized Poisson−Boltzmann equation allow the dissection of electrostatic free energies into contributions from the solvent and from individual interactions within the protein.

Hendsch and Tidor (1994) performed electrostatic continuum calculations that show that salt links usually destabilize proteins with respect to mutation to hydrophobic isosteric residues. This is a direct consequence of the fact that ionic groups are thermodynamically more stable when immersed in a high-dielectric medium such as water than when embedded in a lower dielectric medium such as that of the protein matrix. Conversely, uncharged apolar groups have higher affinities for the protein interior than for water. However, important stabilizing roles for salt links are suggested by analyses showing statistically significant clusterings of charged residues as salt-link networks (Barlow &

Thornton, 1983; Musafia et al., 1995) and a much greater occurrence of salt-link networks in some thermostable proteins (Yip et al., 1995).

Thus, at present, a full understanding of how and when salt links can contribute to protein stability and function is lacking. Although the underlying physical principles governing the energetics of salt links are known, simple rules for reliably estimating the net stability resulting from different competing contributions are difficult to establish. Electrostatic continuum calculations enable the strengths of different salt links and the interactions contributing to their stability to be compared quantitatively. Here, by performing a detailed characterization of selected examples, we identify simple indicators that can be used to assess the stability of a salt link in a protein (or, reciprocally, its contribution to protein stability).

Can salt links either isolated or in networks contribute to protein stability, and how do they contribute to their function? It is our goal to address these two questions via a detailed study of the salt links in cytochrome P450cam. In the next section, we describe the methods used to compute the electrostatic free energies of ionic pairs. In the Results section, we give the computed energies. In the Discussion section, we first consider the factors contributing to salt-link stability and then relate the calculated salt-link strengths to experimental data on substrate access to the active site of P450cam.

MATERIALS AND METHODS

(A) Protocol To Calculate Electrostatic Contributions to Ionic Pair Stability. An ionic pair consists of two residues with oppositely charged side chains an unspecified distance apart. We will define a salt link or salt bridge as an ionic pair formed by two residues with oppositely charged side chains within hydrogen-bonding distance of each other. The electrostatic contribution to the stability of an ionic pair, and thus the folded protein, was determined following the method of Hendsch and Tidor (1994), who evaluated the electrostatic contribution to the "free energy of folding" for 21 salt bridges in nine different proteins with the Delphi program (Gilson et al., 1987). The computed electrostatic free energy difference results from simultaneously mutating the side chains of each ionic pair to their respective uncharged nonpolar hydrophobic isosteres, i.e., hypothetical hydrophobic residues of the same size and shape as the original residues. The mutation allows the determination of the total electrostatic free energy contribution of an ionic pair to protein stability, although, as mentioned in the introduction, this quantity cannot be measured experimentally.

The total electrostatic free energy change, ΔG_{iso}, on mutating the side chains of an ionic pair to their hydrophobic isosteres corresponds to the reversible work associated with the process of removing all partial atomic charges from each side chain forming the ionic pair in the protein and taking the side chains to infinite separation in the solvent medium. Note that if experimental measurements could be made of the stability of a protein with hydrophobic isosteres substituted for an ionic pair, then ΔG_{iso} would be obtained as the difference between two ΔG measurements of unfolding, one for the protein with charged side chains and the other with them replaced by hydrophobic isosteres. Thus ΔG_{iso} would be a $\Delta\Delta G$. Here, results will be presented in terms of $\Delta G_{tot} = -\Delta G_{iso}$ rather than ΔG_{iso} since our purpose is to obtain the stability of the ionic pairs with respect to their isosteres.

From a computational viewpoint, ΔG_{tot} can be decomposed into three distinct terms:

$$\Delta G_{tot} = -\Delta G_{sol} + \Delta G_{brd} + \Delta G_{prt} \quad (1)$$

ΔG_{sol} is due to the reaction field, i.e., the effect of the solvent dielectric response and mobile ions in solution on the partial atomic charges born by the mutated side chains. In our work, we will refer to, and compute, this contribution as the desolvation term, $-\Delta G_{sol}$, resulting from the process of bringing the side-chain charges from the isolated side chain in the solvent medium to the protein interior. The second and third terms simply arise from the sum of the electrostatic interactions of the ionic pair side chains with one another (the bridging term, ΔG_{brd}) and with all other partial atomic charges in the protein (the protein interaction term, ΔG_{prt}). As depicted in Chart 1, each term of ΔG_{tot} is obtained from the difference between grid energies, ΔG^{grid}. These are evaluated on superimposed focused grids. Finally, an additional term, referred to as the electrostatic free energy of association, ΔG_{ass}, was also evaluated for each ionic pair. This term represents the free energy of an ionic pair, relative to mutation to its hydrophobic isosteres, in the absence of the rest of the protein, i.e., when isolated in a homogeneous solvent medium.

The procedure described above was extended to calculate the electrostatic contribution to the stability of salt-link triads, tetrads, pentads, and hexads (involving three, four, five, and six residues, respectively) found in P450cam. It was also used to evaluate the electrostatic free energies of the salt-link tetrad at the entrance to the proposed access channel of the modeled structures of the D251N mutant. The definitions of ΔG_{tot}, ΔG_{sol}, ΔG_{brd}, and ΔG_{prt} for these networks of salt-links are identical to those for a single salt link. As for a single salt link, ΔG_{tot} is the electrostatic free energy change on bringing all the partial atomic charges on the side chains involved in the network from the side chains at infinite separation to their positions in the folded protein. It corresponds to the stability of the salt-link network with respect to simultaneous mutation of all the side chains involved to their hydrophobic isosteres.

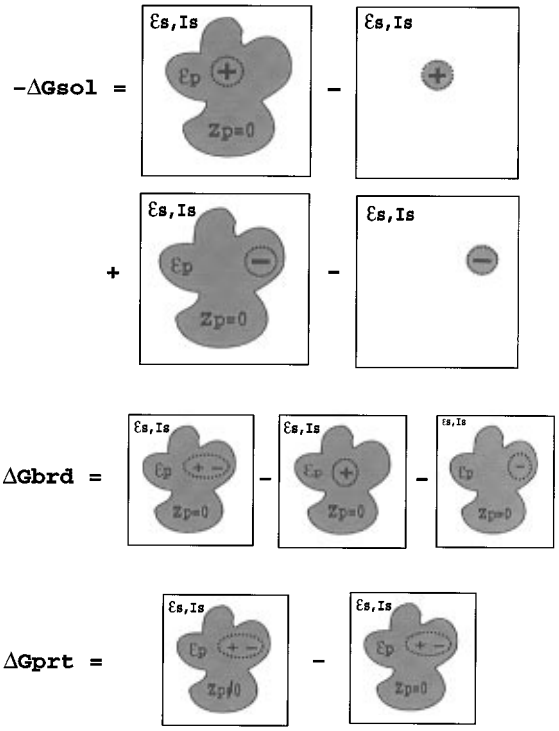
The ΔG_{prt} term was analyzed in detail to evaluate the contributions of different parts of the protein. It is given by

$$\Delta G_{prt} = \sum q_i \phi_i \quad (2)$$

where i runs over all the partial atomic charges excluding those in the titratable groups of the salt link and ϕ_i is the potential created at the location of charge q_i by the charges of the salt link. ΔG_{prt} was decomposed in two ways: (1) by selectively switching off partial atomic charges contributing to dipoles (-C═O and -N−H) in the protein matrix involved in hydrogen bonds to the salt link side chains and (2) by switching off partial atomic charges at which the potential ϕ_i was less than a cutoff value ($|\phi_i| < \phi_{cut}$).

(B) System Setup. Cartesian coordinates of cytochrome P450cam with camphor bound were obtained from the 1.6 Å resolution crystal structure of Poulos et al. (1987) (Brookhaven Protein Data Bank, file 2cpp). The coordinates of the D251N mutant were derived by modeling as described in Appendix A.

Chart 1: Schematic Diagram Showing How the Three Components of ΔG_{tot} Are Computed for an Ionic Pair[a]



[a] The finite difference Poisson−Boltzmann equation is solved in order to evaluate electrostatic free energies in inhomogeneous dielectric environments for charges distributed on grids symbolized by square areas. The charged titratable groups forming the salt link are symbolized by the net charge they bear, (+) and (−). The solvent medium is characterized by its high dielectric response (ϵs) and ionic strength (Is) and the protein core is assigned a lower dielectric response (ϵp). The three components of ΔG_{tot}, given in eq 1, are as follows: (1) The desolvation term, $-\Delta G_{sol}$, calculated as the sum of the grid energy differences obtained on transferring each titratable group in the ionic pair with its charge distribution from the solvent medium to the protein's dielectric environment [all other protein charges being switched off ($Z_p = 0$)]. The self-energies of the charged groups are automatically removed by the difference operation. (2) The bridging term, ΔG_{brd}, consisting of the electrostatic interaction free energy between the two oppositely charged titratable groups embedded in the protein's dielectric environment. The substracted quantities allow the removal of the self-energies of the two titratable groups. (3) The protein interaction term, ΔG_{prt}, corresponding to the amount of reversible work resulting from bringing the protein charges from infinite separation to their loci in the protein core with the salt-link charges already embedded in the protein. In this case also, the self-energies are removed by taking the difference between two grid energies. Note that there is cancellation of quantities when ΔG_{tot} is computed by summing the three terms.

To model the protein at pH 7, residues Glu366 and Asp197 were treated as neutral and residues His80, His270, His308, His352, and His355 as positively charged, as pK_a calculations using the Poisson−Boltzmann model indicated that their pK_as were raised from their standard values in solution (data not shown). These five doubly protonated histidines are in close proximity to the carboxylic acid groups of, respectively, Asp77, Glu269, Glu73, Glu76, and the D propionic group of the heme and thus form salt links whose strengths are computed in the present study. Other titratable residues were assigned their standard acidobasic states at neutral pH. Polar hydrogen atoms were added to the crystal structure using the Whatif program (Vriend, 1990). This program assigns hydrogen atom positions by optimizing hydrogen-bond networks in the protein interior and at the protein surface (Hooft et al., 1996). Thus the protein was assumed to be in one protonation state for the calculations. In addition, each salt link side chain was assumed to be in the same ionization state when in the folded protein and when isolated in solvent, representing an unfolded state of the protein. Thus the contribution to the salt-link energy due to a pK_a shift of any of the salt-link residues on protein folding was neglected. Although this term would be expected to reduce the stabilization of the protein by salt links containing residues with shifted pK_as, it should have little impact on the conclusions of our study.

A potassium ion with unit positive charge was inserted at the position occupied by crystallographic water #515 which is located in an octahedrally coordinated binding site (Poulos et al., 1987). All crystallographically observed water molecules were removed.

(C) Ionic Pair Assignment. Fifty-three ionic pairs were defined by visual inspection of the structure of P450cam. All possible salt links were identified on the basis of the distance between the oppositely charged groups and constitute 32 of the 53 ionic pairs. The distance for defining salt links was measured between the centers of the charged groups, defined as follows: Cγ for Asp, Cδ for Glu, Nζ for Lys, Cζ for Arg, and Nδ or Nϵ for doubly protonated His (whichever is closer to the negative residue in the ionic pair). For an ionic pair to be considered a salt link, this distance was required to be <4.8 Å for ionic pairs involving Arg and <4.3 Å for ionic pairs involving Lys and His. In P450cam, all ionic pairs satisfying these criteria were making hydrogen bonds with favorably aligned dipoles. The 21 other ionic pairs have separation distances of up to 12 Å and were selected to investigate long-range ionic pair interactions. This purposely weak criterion led to the selection of a wide variety of ionic pair types, including those illustrating the effect of long-range coulombic interactions on protein stability. In addition to individual ionic pairs, networks of salt links were characterized as clusters of ionic pairs sharing common charged side chains with one another. Thus, clusters of three, four, five, and six charged residues, referred to in this work as triads, tetrads, pentads, and hexads, were detected and their stability was studied.

(D) Electrostatic Calculations. Continuum electrostatic calculations were performed with the UHBD program (Madura et al., 1994, 1995). This was used to numerically solve the finite difference linearized Poisson−Boltzmann equation with an incomplete Cholesky conjugate gradient method (Davis & McCammon, 1989). Partial atomic charges and atomic radii from the Charmm19 parameter set (Brooks et al., 1983) were used. All radii were scaled by 1.122 to correspond to the minimum in the Lennard−Jones potential. The molecular surface was defined using a solvent probe with a radius of 1.4 Å and the protein/solvent dielectric interface was smoothed over the grid points adjacent to the molecular surface (Davis & McCammon, 1991). A 2 Å Stern layer surrounded the molecule. Mobile ions were treated as distributed in the solvent with a Boltzmann temperature of 300 K. An ionic strength of 150 mM and solvent and protein dielectric constants of, respectively, 78.5 and 4 were used for all calculations unless stated otherwise. The dependence of the stability of the salt links to residue 251 on ionic strength and solvent dielectric was also computed.

In order to make calculation of salt-link strength computationally feasible for this rather large molecular system

Table 1: Computed Electrostatic Free Energy Contributions[a] for All Cytochrome P450cam Salt Links[b]

salt links	no.[c]	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}	ΔG_{tot20}	% SASA	dis[d] (Å)
Lys178−Asp251	4	−7.34	29.13	−14.45	−22.03	1.73	−3.45	7	3.72
Arg186−Asp251	4	−6.03	23.23	−16.19	−13.08	0.32	−3.08	5	4.05
Asp182−Arg186	4	−2.18	16.91	−5.96	−13.14	0.54	−1.37	22	4.61
Lys178−Asp182	4	3.27	22.97	−11.39	−8.30	0.15	−0.86	23	3.68
Arg112−HemCGD	3	−8.77	20.57	−10.69	−18.65	−0.85	−2.40	1	3.36
His355−HemCGD	3	−6.84	24.88	−15.69	−16.04	−2.06	−3.33	1	3.50
Glu287−Arg342	3	−4.12	23.23	−9.37	−17.87	0.47	−2.25	12	4.61
Asp97−Lys197	2	−4.02	13.87	−7.46	−10.42	0.09	−2.30	23	3.43
Arg211−Asp218	2	−2.84	18.67	−10.06	−11.45	−0.37	−1.65	19	4.07
Asp77−His80	2	−2.14	10.29	−8.78	−3.66	−1.83	−1.48	34	3.28
Arg277−Glu279	2	−1.80	5.36	−3.24	−3.92	0.40	−1.31	64	4.72
Glu373−Arg377	3	−1.44	13.74	−6.47	−8.71	0.96	−2.13	5	4.74
Glu76−His352	2	−1.27	5.91	−4.39	−2.79	−0.77	0.94	46	3.60
Arg67−Asp328	3	−1.26	22.44	−9.66	−14.04	−0.22	−1.91	2	4.57
Arg67−Glu329	3	−1.05	14.95	−6.64	−9.36	0.47	−1.40	29	4.42
Lys372−Glu373	3	−0.79	10.48	−4.71	−6.56	0.99	−1.21	34	4.01
Glu73−His308	2	−0.47	7.72	−8.26	0.07	−2.51	−0.98	45	3.31
Arg231−Glu237	2	−0.39	13.72	−9.62	−4.49	−0.23	−1.27	30	4.20
Asp236−Arg240	2	−0.34	8.90	−3.35	−5.88	1.11	−1.52	44	4.78
Asp25−Arg57	2	−0.28	10.58	−7.70	−3.16	−0.37	−0.90	29	4.10
Arg299−HemCGA	2	0.16	26.67	−18.91	−7.60	−0.99	−1.58	1	4.38
Glu273−Arg280	2	0.63	6.87	−5.06	−1.17	−1.05	−0.59	35	4.44
Glu287−Arg290	3	0.85	28.86	−16.07	−11.93	−0.25	−1.38	0	4.02
Glu140−Arg143	2	1.71	11.32	−8.64	−0.97	−1.18	−0.60	29	4.30
Arg271−Asp380	2	2.01	8.35	−2.24	−4.09	0.64	−0.47	40	4.01
Glu262−Lys266	2	2.21	10.14	−3.33	−4.60	1.03	−0.44	35	4.00
Arg130−Glu133	2	2.28	10.04	−7.21	−0.55	1.45	−0.53	62	3.90
Glu128−Arg365	2	2.55	8.01	−5.45	−0.01	−0.58	−0.08	38	4.60
Asp52−Arg330	2	2.70	8.27	−4.16	−1.42	0.84	−0.36	49	3.99
Glu40−Arg291	2	2.78	14.09	−5.76	−0.55	0.45	−0.44	22	4.47
Glu286−Arg364	2	2.94	13.99	−8.21	−2.83	−0.63	−0.22	30	4.10
Arg72−Glu331	2	5.44	12.91	−8.10	0.62	−1.22	0.40	23	4.26

[a] Given in kilocalories per mole. [b] All salt links with oppositely charged side chains within hydrogen-bonding distance of each other are listed. The tetrad of salt links at the entrance to the proposed substrate access channel is listed separately at the beginning of the table. The other salt links are listed in order of decreasing stability with respect to their hydrophobic isosteres, i.e., in order of increasing ΔG_{tot}. [c] Number of charged residues in the salt-link network. [d] Distance between charge centers as defined in the Materials and Methods section.

(P450cam has approximately the shape of a prism with triangular edge 60 Å and thickness 30 Å), we used the technique of successively focused grids. The electrostatic potential was first calculated with a cubic grid of 50 × 50 × 50 points with a 7.5 Å spacing centered at the center of geometry of the protein. This initial grid was large enough to ensure a vanishing potential at the box edges, and boundary conditions were assigned according to the Debye−Hückel model with each protein atom treated as a sphere. Then, for each salt link studied, two cubic grids with spacings of 3.6 and 0.4 Å, respectively, were successively focused on the salt-link center. Boundary conditions for the 3.6 and 0.4 Å grids were interpolated from the potential obtained with the initial coarse grid and the 3.6 Å grid, respectively. For calculations involving isolated salt links, both grids had 50 × 50 × 50 points, while for calculations for groups of salt links, the dimensions of the third grid were increased up to 75 × 75 × 75 or 100 × 100 × 100 points to accommodate their spatial extent. It should be noted that the first two grids both include all the protein atoms whereas the third grid is embedded in the protein to an extent depending on the degree of burial of each salt link. This method allows reasonably fast calculations: less than 18 min of cpu time per salt link on a single R8000 processor of a Silicon Graphics Power Challenge computer. Sensitivity to numerical parameters was investigated as described in Appendix B and the method was found to be robust.

(F) Solvent-Accessible Surface Areas. Percentages of solvent-accessible surface areas (SASA) for the side chains of the titratable residues were calculated with the NACCESS program (S. Hubbard, 1994) using a solvent probe of radius 1.4 Å and default atomic radii as given by Lee and Richards (Lee & Richards, 1971; Richards, 1979).

RESULTS

Ionic Pair Stabilities

Electrostatic Free Energy, ΔG_{tot}. The electrostatic free energies, ΔG_{tot}, of the ionic pairs studied relative to their corresponding hydrophobic isosteres are shown, along with values for the terms contributing to ΔG_{tot}, in Tables 1 and 2. The salt links are listed in order of increasing ΔG_{tot} in Table 1, while ionic pairs with longer separation distances than satisfy the criteria for salt links are listed in Table 2. The distribution of ΔG_{tot} for the salt links is shown in Figure 2. Nineteen salt links have $\Delta G_{tot} < 0$ kcal/mol (stabilizing the protein's folded state with respect to mutation to hydrophobic isosteres) and 13 have $\Delta G_{tot} > 0$ kcal/mol.

The assignment of a higher dielectric constant than 4 to the protein has been found to give better agreement with experiment in pK_a calculations (Antosiewicz et al., 1994; Demchuk & Wade, 1996), with a dielectric constant of 20 giving reasonable results for the CHARMm parameter set (Antosiewicz et al., 1994). Salt-link strengths recomputed with a protein dielectric constant of 20 are given in the eighth column of Table 1. As would be expected, the magnitude of the salt-link strengths is generally decreased. As the solvation penalty is reduced, all but two of the 32 salt links

Table 2: Computed Electrostatic Free Energy Contributions[a] for a Subset of the Ionic Pairs with Side-Chain Charge Centers Further than Hydrogen-Bonding Distance Apart[b]

ionic pairs	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}	% SASA	dis[c] (Å)
Glu237−Arg240	−7.24	12.52	−1.38	−18.38	0.36	19	6.6
Asp97−Arg240	−6.85	12.95	−2.24	−17.56	1.09	28	5.11
Glu279−Arg280	−6.18	5.86	−0.72	−11.31	0.03	55	6.39
Arg212−Glu237	−3.95	14.66	−0.95	−17.67	−0.22	23	9.00
Lys314−Asp316	−2.89	8.34	−0.19	−11.03	−0.07	43	8.49
Glu269−His270	−2.49	8.92	−1.77	−9.64	0.42	53	5.50
Asp339−Arg342	−1.76	14.16	−1.76	−14.16	0.24	27	5.44
Glu276−Arg277	−1.72	2.50	−0.40	−3.82	−0.21	64	7.27
Asp304−Lys314	−1.66	3.18	−0.32	−4.53	0.43	67	8.50
Glu20−Arg57	−1.48	3.43	−0.39	−4.53	−0.23	23	9.00
Asp236−Lys239	−1.17	5.39	−1.90	−14.16	0.12	42	4.68
Glu329−Arg330	−0.11	8.55	−0.87	−7.78	−0.04	55	6.56
Asp304−Lys313	−0.03	4.06	−0.72	−3.34	0.58	57	6.30
Lys197−Glu198	0.09	9.58	−1.95	−7.54	0.08	38	5.28
Arg377−Valc414	0.86	9.61	−1.14	−7.60	0.19	30	5.06
Glu209−Arg212	1.24	8.59	−0.27	−7.78	−0.10	50	9.50
Asp125−Lys126	1.28	2.90	−0.14	−1.47	−0.07	80	11.4
Arg90−Glu94	1.33	7.10	−0.82	−4.95	0.14	55	6.40
Glu171−Arg161	1.66	3.58	−1.12	−0.79	0.61	47	5.02
Asp104−His355	1.87	16.64	−4.28	−10.49	−0.06	16	5.67
Glu198−Lys239	2.38	5.24	−1.13	−1.73	0.40	43	6.35

[a] Given in kilocalories per mole. [b] The ionic pairs are listed in order of decreasing stability with respect to their hydrophobic isosteres. [c] Distance between charge centers as defined in the Materials and Methods section.

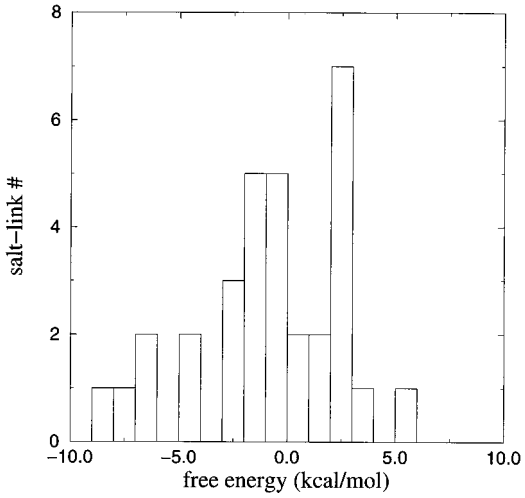


FIGURE 2: Distribution of electrostatic free energies, ΔG_{tot}, of the 32 salt links (defined according to distance criteria described in the text) in cytochrome P450cam. Negative and positive values of ΔG_{tot} indicate salt links that are, respectively, more stable or less stable than their hydrophobic isosteres.

become electrostatically stable with free energies ranging down to −3.45 kcal/mol. Strikingly, at both protein dielectric constants of 4 and 20, the same four salt links are the most stable. Thus, while quantitative numbers for free energies may be sensitive to protein dielectric constant assignment, the relative energies and the distribution of salt-link strengths seem to be rather robust to variation of the protein dielectric constant. However, the ratio of stable to unstable salt links when the protein dielectric constant is set to 4 is more consistent with experimental data (Waldburger et al., 1995; Wimley et al., 1996), which indicate that the presence of salt links frequently destabilizes the protein fold with respect to mutation to hydrophobic isosteres. We have therefore used a protein dielectric constant of 4 for all further analysis.

The influence of the presence of camphor on the strength of the salt links to Asp 251 was investigated by recomputing the strength of all the salt links with camphor removed from the active site. In the crystal structure of the substrate-free enzyme (Poulos et al., 1986), one ordered water molecule is observed liganded to the iron atom of the heme and a further five water molecules are assigned to occupy a large blob of density as disordered water molecules. In the electrostatic calculations, these water molecules are represented by continuum high-dielectric solvent even though their dielectric properties may be different from those of pure bulk water. For the majority of salt links the difference in ΔG_{tot} was negligible. The largest differences in ΔG_{tot} were obtained for the salt links closest to the active site, but these did not exceed 0.6 kcal/mol. The energetic ranking of the salt links remained the same.

Contributions to ΔG_{tot}. All of the salt links are characterized by a significant negative ΔG_{brd} ranging from −2 to −19 kcal/mol (with an average of about −7 kcal/mol). The ΔG_{brd} values of non-salt-link ionic pairs (Table 2) are mostly smaller in magnitude than for the salt links, lying in the range 0 to −4 kcal/mol and making little contribution to the overall ΔG_{tot}. Thus, it appears that the very simple hydrogen-bond length criterion has a clear correspondence to ΔG_{brd} and can discriminate between ionic pairs that are hydrogen-bonded and those that are not. However, as is clear from the similar ranges of ΔG_{tot} in both tables, this criterion would not be useful for predicting the electrostatic stability of any salt link. ΔG_{sol} and ΔG_{brd} are opposite in sign and ΔG_{brd} is usually smaller in magnitude than ΔG_{sol}. Therefore, the electrostatic stabilization of a salt link generally requires a substantial favorable ΔG_{prt} term. The ΔG_{prt} values reported in Tables 1 and 2 range from 1 to −22 kcal/mol. In most but not all cases for which there is a large negative ΔG_{prt} term (< -10 kcal/mol), the resultant ΔG_{tot} is negative. However, exceptions are also found for which the ΔG_{sol} value is comparable in magnitude to ΔG_{brd} and thus ΔG_{tot} becomes negative with a small ΔG_{prt} value, as for instance in the case of the Asp77−His80 and Arg277−Glu279 salt links.

Contributions to the ΔG_{prt} *Term.* Because of the importance of the protein interaction term, ΔG_{prt}, for salt-link stability, we analyzed the different types of contributions to

Table 3: Electrostatic Free Energies[a] Computed for (i) an "Isolated" Salt Link and (ii) a Loose Pentad Salt-Link Cluster[b]

salt link	charges neutralized[c]	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}
(i) Arg211−Asp218	none	−2.84	18.67	−10.06	−11.45	−0.37
	G168 C═O	−1.02	*d*	*d*	−9.64	*d*
	G168 C═O and A167 C═O	0.68	*d*	*d*	−7.39	*d*
	Asp218 N−H	1.35	*d*	*d*	−6.72	*d*
	other titratable groups	2.09	*d*	*d*	−5.98	*d*
	$\Phi i < 10.0$				−4.56	
	$\Phi i < 1.0$				−10.57	
	$\Phi i < 0.1$				−11.28	
(ii) pentad from 276 to 273	none	−2.30	13.10	−8.44	−6.93	−0.28
	Glu279 N−H	−1.10	*d*	*d*	−5.73	*d*
	Glu279 N−H and Gln272 $-NH_2$	0.37	*d*	*d*	−4.27	*d*
	other titratable groups	1.49	*d*	*d*	−3.16	*d*

[a] Given in kilocalories per mole. [b] These examples illustrate how the protein charge distribution surrounding a salt link or salt-link cluster contributes to the protein interaction term ΔG_{prt}. [c] Partial charges were set to zero on the specified atoms or on all atoms at which the electrostatic potential (Φi) due to the salt-link residues was smaller in magnitude than the specified cutoff value given in kilocalories per mole per electron. [d] Same as above.

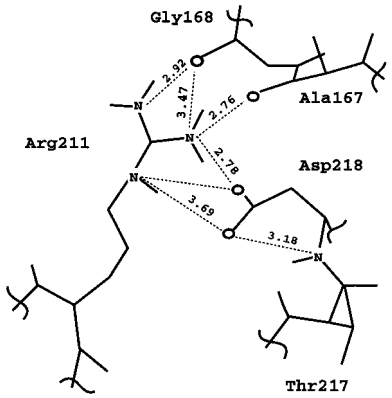


FIGURE 3: Detail of the electrostatically stable "isolated" Arg211−Asp218 salt link and its immediate surroundings. Hydrogen bonds are represented by dashed lines with distances given in angstroms.

it. Four types can be identified: (1) short-range (hydrogen-bond length) interactions with other charged titratable groups present in the vicinity of the studied salt link, (2) long-range interactions with other charged titratable groups, (3) hydrogen-bond interactions with dipolar groups close to the salt link, and (4) interactions with dipolar groups further from the salt link. Interactions of the first two types result in networks of salt links: tight ones via interactions of type 1 only and looser ones via interactions of types 1 and 2. The number of possible ways to cluster the 126 titratable groups in cytochrome P450cam into loose salt-link networks is very large. To carry out a systematic and detailed analysis of all the salt-link networks and clusters in P450cam and of the different contributions to the ΔG_{prt} term is not the aim of this work. We have instead investigated how salt links are stabilized by the surrounding protein environment in a few cases which illustrate the different mechanisms by which stabilization can be obtained. These are as follows:

(a) Arg211−Asp218: A Stable "Isolated" Salt Link. This salt link has no titratable residues in its immediate vicinity but is stabilized ($\Delta G_{tot} < 0$, see Tables 1 and 3) by a network of hydrogen bonds from surrounding neutral residues (see Figure 3).

The values of ΔG_{prt} when the electrostatic dipoles on neighboring residues were removed are reported in Table 3. As the hydrogen bonds are switched off, ΔG_{prt} increases from −11.5 to −6.7 kcal/mol, resulting in a positive ΔG_{tot} (+1.35 kcal/mol) and therefore an unstable salt link. The hydrogen bonds between the Arg211 guanidinium nitrogen atoms and the carbonyl oxygen atoms of Gly168 and Ala167 are strong (length 2.8−2.9 Å) and each provides a net stabilizing energy of about −2 kcal/mol. The hydrogen bond between the Asp 218 carboxylate group and its backbone nitrogen atom is longer and provides only about −0.7 kcal/mol stabilization.

When all other titratable groups in the protein core are neutralized, ΔG_{prt} is reduced in magnitude by about 0.7 kcal/mol, so for the Arg211−Asp218 salt link, the contribution of the net charges of other charged residues is stabilizing but rather small. Thus, the remaining approximately −6 kcal/mol of the stabilizing energy of ΔG_{prt} is provided neither by short-range dipole interactions nor by long-range net charge interactions but rather by other (long-range) protein dipoles.

In order to estimate how far the influence of the protein charge distribution on the ΔG_{prt} term extends, we decomposed ΔG_{prt} into partial sums which are a function of the magnitude of the electrostatic potential generated by the salt-link charge distribution (see Table 3). With an electrostatic potential cutoff of 10 kcal/mol/e, ΔG_{prt} is −4.6 kcal/mol and is due to partial atomic charges in the protein at an average distance of 4.1 Å, and a maximum distance of 4.5 Å, from the center of the salt link (equidistant between the centers of the titratable groups). These are the partial charges of atoms making hydrogen bonds to the residues in the salt link. With an electrostatic potential cutoff of 1 kcal/mol/e, ΔG_{prt} is −10.8 kcal/mol, and with a cutoff of 0.1 kcal/mol/e, ΔG_{prt} is −11.3 kcal/mol. The average and maximum distances to protein partial atomic charges within these electrostatic potential cutoffs are, respectively, 5.3 and 8.8 Å and 10.2 and 16.9 Å. These two partial sums show that about 95% of the ΔG_{prt} term arises from interactions with partial atomic charges located within about 10 Å of the center of the salt link. Thus, stabilization of a salt link by the surrounding protein is a primarily local effect (within 10 Å).

(b) Salt-Link Hexad of Glu198−Lys197−Asp97−Arg240−Asp236−Lys239: An Unstable Loosely Connected Salt-Link Network. This network is unstable with respect to simultaneous mutation of all residues to their hydrophobic isosteres (see Table 4). It is composed of two hydrogen-bonded salt links, Lys197−Asp97 and Arg240−Asp236, and four more separated ionic pairs with center-to-center distances of 5.1, 4.7, 5.3, and 6.4 Å for Asp97−Arg240, Asp236−Lys239, Glu198−Lys197, and Glu198−Lys239, respectively. Cooperativity operates between neighboring titratable residues

Table 4: Computed Electrostatic Free Energy Contributions[a] for Three Clusters of Salt Links[b]

network of salt links	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}
tetrad 178−182−186−251	3.24	46.24	−33.55	−9.44	2.57
triad 178−251−186					
wild type, X-ray structure	−4.27	39.72	−24.34	−19.66	3.02
wild type, MDS average	−8.65	39.52	−24.95	−23.23	1.47
D251N, MDS average	−2.09	22.32	−1.11	−23.29	2.11
hexad 198−197−97−240−236−239	0.48	27.11	15.42	−11.21	1.69
pentad 197−97−240−236−239	−2.20	25.08	−11.93	−13.81	2.51
tetrad 197−97−240−236	−3.18	22.54	−11.93	−13.80	2.30
triad 197−97−240	−3.52	19.92	−9.02	−14.41	−0.10
pentad 276−277−279−280−273	−2.30	13.10	−8.44	−6.93	−0.28
tetrad 276−277−279−280	−5.02	8.37	−3.89	−9.49	0.68
triad 276−277−279	−2.07	6.27	−3.44	−4.90	−0.06

[a] Given in kilocalories per mole. [b] Results for the tetrad at the entrance to the proposed substrate access channel, a hexad, and a pentad are shown. Results are also shown for subclusters of these clusters.

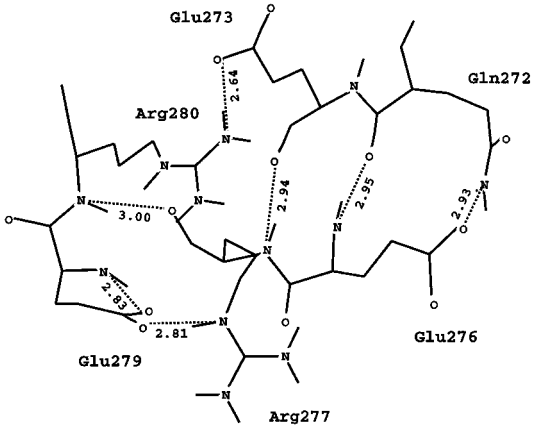


FIGURE 4: Detail of the electrostatically stable loosely connected Glu273−Arg280−GLu279−Arg277−Glu276 network of ionic pairs. Hydrogen bonds are represented with dashed lines with distances given in angstroms.

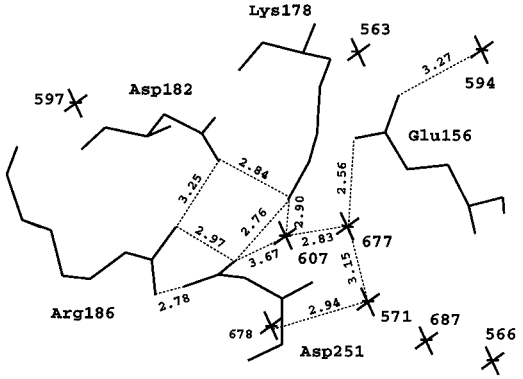


FIGURE 5: Detail of the tetrad of salt links near the entrance of the proposed substrate access channel and the adjacent water molecules as determined in the 1.6 Å resolution crystal structure from Poulos *et al.* (1987). Hydrogen bonds are represented by dashed lines with distances given in angstroms.

and contributes to the stability of the Asp97−Lys197, Asp97−Arg240, and Asp236−Lys239 ionic pairs, which all have $\Delta G_{tot} < -1$ kcal/mol due to large negative ΔG_{prt} terms (<-10 kcal/mol), cf. Tables 1 and 2. Conversely, the Arg240−Asp236, Glu198−Lys197, and Glu198−Lys239 ionic pairs have $\Delta G_{tot} \gtrsim 0$ and concomitantly smaller ΔG_{prt} terms (> -10 kcal/mol). The subclusters inside this network are similarly cooperatively stabilized. Thus, as shown in Table 4, the triad 197−97−240, the tetrad 197−97−240−236, and the pentad 197−97−240−236−239 prove to be more stable than their hydrophobic isosteres ($\Delta G_{tot} < -2$ kcal/mol). However, the whole hexad network is less stable than its corresponding hydrophobic isosteres. This is because of the weak stability of the loose interactions of Glu198 with Lys197 and Lys239.

(c) Salt-Link Pentad of Glu276−Arg277−Glu279−Arg280−Glu273: A Stable Loosely Connected Salt-Link Network. This network is stable with respect to simultaneous mutation of all residues to their hydrophobic isosteres largely because of hydrogen-bonding interactions with adjacent backbone and neutral side-chain atoms (see Tables 3 and 4).

As a whole, $\Delta G_{tot} = -2.30$ kcal/mol for this salt-link pentad, which is a chain of four ionic pairs. A detailed view of this network is given in Figure 4. There are two hydrogen-bonded salt links, Glu273−Arg280 and Glu279−Arg277, and two more separated ion pairs, Arg280−Glu279 and Arg277−Glu276. Geometrically, this association is representative of common arrangements of titratable groups at the protein surface. Examination of its surroundings shows that it is connected to the protein matrix via two quite strong hydrogen bonds: one from the amide group of Gln272 to the Glu276 carboxylic group and the other from the Glu279 backbone amide group to the Glu 279 carboxylic group. As shown in Table 3, these hydrogen bonds contribute 2.7 kcal/mol to the network stability, ΔG_{tot}. Other titratable groups in the protein contribute about 1 kcal/mol to the network stability.

(d) Salt-Link Tetrad of Lys178−Asp182−Arg186−Asp251: The Entrance Channel Tetrad. As well as containing Asp251 and thus being of interest for its functional role, this tetrad provides an example of a tightly connected salt-link network which is unstable with respect to simultaneous mutation of all residues to their hydrophobic isosteres.

The residues involved in the entrance channel tetrad form a particularly tight network of four salt links arranged in an approximately square geometry as depicted in Figure 5. As reported in Table 1, the two salt links to Asp251 are among the four strongest in the whole protein. They have ΔG_{tot} values of −7.3 and −6.0 kcal/mol, which are at the extreme of the distribution shown in Figure 2 and reported in other work (Hendsch & Tidor, 1994). These two salt links have large negative ΔG_{brd} and ΔG_{prt} terms. The large ΔG_{brd} term reflects the good geometry of the hydrogen bonds between the titratable groups, and the large ΔG_{prt} term arises predominantly from the presence of the other charged groups in the tetrad. Indeed, the ΔG_{prt} term for the Lys178−Asp251 salt link includes, among other protein contributions, the

Table 5: Computed Electrostatic Free Energy Contributions[a] for the Tetrad of Salt Links at the Entrance to the Proposed Substrate Access Channel[b]

system[c]	residue pairs	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}	% SASA	dis (Å)
wt (−)	Lys178−Asp251	−12.77	19.18	−4.41	−27.54	−0.16	8	5.42
	Arg186−Asp251	−2.88	23.67	−17.66	−8.90	−1.30	5	3.84
	Asp182−Arg186	−11.30	17.57	−7.13	−21.73	0.00	21	4.57
	Lys178−Asp182	−3.64	12.94	−11.29	−5.29	−1.89	23	3.12
wt (H)	Lys178−Asp251	−9.71	28.12	−14.70	−23.12	1.46	8	3.49
	Arg186−Asp251	−8.41	23.60	−17.40	−14.61	−1.15	6	3.94
	Asp182−Arg186	−9.73	16.86	−6.67	−19.91	0.34	25	4.48
	Lys178−Asp182	−3.02	21.37	−12.08	−12.32	−0.71	24	3.70
D251N (H)	Lys178−Asn251	1.87	14.31	−5.05	−7.39	−0.01	10	3.64
	Arg186−Asn251	−2.36	10.91	0.11	−13.38	0.13	7	6.43
	Asp182−Arg186	−2.88	14.56	−9.46	−7.97	−0.97	17	4.02
	Lys178−Asp182	−3.64	17.58	−10.34	−10.88	−0.44	21	3.54
D251N (−)	Lys178−Asn251	−4.72	7.86	−0.47	−12.11	0.46	11	5.29
	Arg186−Asn251	−1.66	11.14	0.86	−13.65	0.51	7	6.07
	Asp182−Arg186	0.80	14.40	−9.70	−3.90	−1.09	18	3.23
	Lys178−Asp182	−10.80	10.96	−3.26	−18.49	−0.15	22	4.77
model made by hand	Lys178−Asn251	5.25	18.52	−4.34	−8.93	1.07	7	3.71
	Arg186−Asn251	0.56	11.93	−0.71	−10.65	0.82	6	5.53
	Asp182−Arg186	−0.92	16.01	−9.92	−7.01	−0.73	20	4.36
	Lys178−Asp182	1.02	22.56	−11.45	−10.09	−0.48	21	3.97

[a] Given in kilocalories per mole. [b] Computed for average structures obtained from simulations of the wild-type and D251N mutant protein with Glu156 both protonated and unprotonated and for a "hand-built" model of the D251N mutant. [c] The protonation state of Glu156 is denoted by − (unprotonated) and H (protonated).

ΔG_{brd} of both the Lys178−Asp182 salt link ($\Delta G_{brd} = -11.4$ kcal/mol) and the Asp251−Arg186 salt link ($\Delta G_{brd} = -16.2$ kcal/mol). Similar terms contribute to the other two stabilizing salt links of the tetrad: Arg186−Asp251 and Asp182−Arg186. The Asp182−Arg186 salt link is, however, less stable because its bridging term is less favorable ($\Delta G_{brd} = -6.0$ kcal/mol) as the hydrogen bond between the residues is relatively long (3.25 Å). The relative weakness of the bridging term of the Asp182−Arg186 salt link is the direct cause of the instability of the Lys178−Asp182 salt link (+3.3 kcal/mol) and the somewhat reduced stability of the Arg186−Asp251 salt link compared to the Lys178−Asp251 salt link. Even though imperfect, this tetrad is nonetheless a remarkable case where cooperativity results in the abnormally strong stability (−4.3 kcal/mol) of a triad of residues, 178, 251, and 186, which, in the case of P450cam, apparently plays a functional role in substrate binding. Note, though, that as a whole this tetrad does not stabilize the protein with respect to the mutation of all four residues to their hydrophobic isosteres; see Table 4.

D251N Mutant

The results of computations of salt-link stability for the four average structures derived from simulations (wild-type P450cam with and without Glu156 protonated and the D251N mutant with and without Glu156 protonated, as described in Appendix A) are presented in Table 5 along with data for the hand-built model of the D251N mutant. In general, the ΔG_{brd} and ΔG_{prt} terms evaluated for the structures from the simulations tend to be somewhat more favorable than those obtained for the X-ray structure because the X-ray structure was not energy-minimized before electrostatic calculations were carried out. The free energies computed for the D251N mutant are also more favorable for the simulated structures than for the hand-built model.

For the average structure of wild-type P450cam with Glu156 protonated, all but two values of ΔG_{sol}, ΔG_{brd}, and ΔG_{prt} are reasonably consistent with the calculation performed on the X-ray structure itself (compare values in Tables 1 and 5). These two values, which are the ΔG_{prt} for the Asp182−Arg186 and Asp182−Lys178 salt links, are substantially larger in magnitude than expected relative to differences for the other reported values. Analysis of the terms involved for the Asp182−Lys178 salt link shows that the main contribution to the difference of about 4 kcal/mol is the hydrogen bond between the Leu250 carbonyl group and the amino group of Lys178, which is shorter than in the X-ray structure (2.6 instead of 3.0 Å). The rest of the stabilization is the result of more subtle energetic optimizations of the tetrad structure. Similarly, for the Asp182−Arg186 salt link, the difference of about 7 kcal/mol results largely from improved ΔG_{brd} and ΔG_{prt} terms (∼2 kcal/mol) for the Lys178−Asp182 and Asp251−Arg186 salt links, and a shorter (2.7 instead of 3.0 Å) hydrogen bond between the carbonyl group of Val396 and the NH1 amine of Arg186 (∼3.5 kcal/mol).

Solvent Dielectric and Ionic Strength Dependence of Substrate binding

The computed dependence on ionic strength and dielectric constant of the strengths of the salt links to residue 251 is shown in Table 6. Computations were performed at the ionic strengths and dielectric constants at which substrate binding was measured experimentally (Deprez et al., 1994). The stability differences due to changes in ionic strength and dielectric constant are small but of similar magnitude to the experimentally measured differences in the equilibrium constants for camphor entrance to the active site. However, as will be discussed in the next section, these numbers are not directly comparable.

DISCUSSION

Salt-Link Stability

While the ΔG_{brd} term is a direct reflection of the geometry of the salt link and thus the strength of the interactions

Table 6: Ionic Strength and Solvent Dielectric Constant Dependence of the Stability of Salt Links of Importance in Substrate Binding[a]

protein	structure[a]	residue pairs/triads	ΔG_{tot}[b]	$\Delta\Delta G_{IS}$[c]	$\Delta\Delta G_{IS}^{ex}$[d]	$\Delta\Delta G_{\epsilon}$[e]	$\Delta\Delta G_{\epsilon}^{ex}$[d]
wt	X-ray	Lys178−Asp251	−6.15	0.25		−0.73	
		Arg186−Asp251	−6.15	0.24		−0.73	
		Arg186−Asp251	−6.15	0.24		−0.39	
		Lys178−Asn251−Arg186	−2.00	0.98		−0.94	
		Asp97−Lys197	−4.03	0.26		−0.42	
wt	MDS	Lys178−Asp251	−9.56	0.26		−0.53	
		Arg186−Asp251	−8.44	0.25		−0.44	
		Lys178−Asp251−Arg186	−8.80	0.98		−1.00	
		Asp97−Lys197	−4.32	0.25		−0.44	
					−0.42 ± 0.13		1.01 ± 0.08
D251N	MDS	Lys178−Asn251	1.78	0.43		−0.14	
		Arg186−Asn251	−2.50	0.43		−0.04	
		Lys178−Asn251−Arg186	−2.34	1.09		−0.56	
		Asp97−Lys197	−4.18	0.24		−0.45	
					−0.17 ± 0.08		0.95 ± 0.07

[a] Stabilities, ΔG_{tot}, and relative stabilities, $\Delta\Delta G$, are given in kilocalories per mole and were computed with the protein dielectric constant set to 4 and Glu 156 assumed to be protonated. Control calculations in which Glu 156 was unprotonated resulted in $\Delta\Delta G$ values within 0.1 kcal/mol of those given in the table. [b] ΔG_{tot} computed at 100 mM ionic strength and a solvent dielectric constant of 80.4. [c] $\Delta\Delta G_{IS} = (\Delta G_{tot})_{IS=500} - (\Delta G_{tot})_{IS=100)}$, where IS is the ionic strength in millimolar. [d] Experimental measurements of Deprez et al. (1994). [e] $\Delta\Delta G_{\epsilon} = (\Delta G_{tot})_{\epsilon=64.5} - (\Delta G_{tot})_{\epsilon=80.4}$, where ϵ is the solvent relative dielectric constant.

between the residues of an ion pair, the electrostatic stability of an ionic pair is highly dependent on its environment, which can be described by the ΔG_{sol} and ΔG_{prt} terms. The desolvation energy, $-\Delta G_{sol}$, is always positive (unfavorable) while the protein interaction term, ΔG_{prt}, is nearly always negative. ΔG_{brd} and ΔG_{prt} must together overcome the desolvation penalty, $-\Delta G_{sol}$, for a salt link to be stable with respect to mutation to its hydrophobic isosteres.

The Arg211−Asp218 salt link, the first example given in the Results section, shows how stabilization of an isolated salt link (i.e., a salt link that is not in a salt-link network) can be achieved by surrounding it by hydrogen-bonding groups from neutral residues. Indeed, an examination of all isolated salt links that are more stable than their hydrophobic isosteres in P450cam shows that they are systematically sharing additional hydrogen bonds with the protein matrix. Isolated salt links that do not participate in additional hydrogen bonds with the rest of the protein are necessarily unstable with respect to their hydrophobic isosteres. Such salt links are found more frequently than stable isolated salt links. Thus, while isolated salt links do not usually contribute to protein stability when compared to substitution by isosteric hydrophobic residues, their formation can be important for defining protein structure and is usually preferable to the burial of a single isolated titratable residue because of the favorable contribution of ΔG_{brd}.

The next three examples in the Results section show the importance of hydrogen bonds to neutral residues in the surrounding protein for the stability of salt-link networks. A network will be unstable with respect to the simultaneous mutation of all of its residues to hydrophobic isosteres if hydrogen bonds between the residues of the network and the rest of the protein are lacking. However, these examples also show that subclusters (both loose and tight) can be stabilized with respect to mutation to their hydrophobic isosteres by the proximity of other titratable groups as well as neutral hydrogen-bonding moieties.

The factors that this analysis reveals as important for stability would vary in magnitude if a different reference state were chosen. We have computed the electrostatic free energy change on taking the side chains of the residues of a salt link to infinite separation from each other and the protein in solvent. This is the same as the free energy of the salt-link residues relative to uncharged isosteric residues under the assumption that the protein structure is unchanged by the mutations. The isolated side chain in solvent can be considered to represent a denatured state of the protein. Clearly this is an incomplete model of the structure and dynamics of the unfolded state; however, Hendsch and Tidor (1994) found that including local backbone atoms to make a more realistic unfolded state model produced similar results to those with the side chain alone. Mutation to isosteric polar, uncharged residues, e.g., Asp to Asn, assuming that the protein structure is unchanged by mutation, could be treated similarly. Mutations to alanine, often performed experimentally, require account to be taken of the change in protein−solvent dielectric boundary due to its different size. They thus require a qualitatively different computational treatment from that performed here (as would relaxation of the assumption of protein rigidity on mutation).

Functional Role of Exceptionally Stable Salt Links

Salt Links to the Heme Cofactor. Two of the three strongest salt links in the protein are made to the D propionate group of the heme cofactor. These are deeply buried salt links with very favorable ΔG_{brd} and ΔG_{prt} terms. One of them is to a histidine (His355) which is doubly protonated and has a raised pK_a.

In general, apoproteins are less stable than holoproteins and this is also the case for cytochrome P450cam (Pfeil et al., 1993). Apo-P450cam occurs earlier during *in vivo* folding than the holoprotein. It is weakly stable (~2 kcal/mol at 25 °C) and is not fully structured (Pfeil et al., 1993). The apoprotein must bind its prosthetic group in a specific manner that leads to structuring of the protein and greater stability so that the reverse transition back to the apoprotein does not take place readily. In P450cam, specific and tight binding of the heme appears to be achieved by the Cys357 fifth ligand, nearby aromatic side chains, and the unusually strong salt links to the D propionate group. The importance of these salt links for heme stabilization is supported by (a) sequence alignment of cytochrome P450s [see, e.g., Hasemann et al. (1995)] showing that, in other P450s, His355 is usually replaced by an arginine and (b) mutagenesis studies

showing that mutation of Arg112 leads to decreased affinity for putidaredoxin and reduced electron transfer and oxygen consumption rates (Koga et al., 1993). Further mutagenesis studies, e.g., of His355, aimed at examining heme binding could define the energetic and stabilizing roles of these residues better.

Salt Links To Regulate Channel Opening. In this discussion, we will first review pertinent experimental data and then discuss the results of the salt-link calculations in the context of the experimental data.

The binding of camphor to P450cam can be considered as a two-step process in which camphor first diffuses from outside the protein to the binding site and there is then a low- to high-spin transition at the heme Fe (Deprez et al., 1994; Sligar, 1976):

$$\text{P-450}^{\text{o}}{}_{\text{LS}} + \text{S} \overset{K_{1d}}{\leftrightarrow} \text{P-450}^{\text{os}}{}_{\text{LS}} \overset{K_2}{\leftrightarrow} \text{P-450}^{\text{os}}{}_{\text{HS}}$$

where o, LS, S, and HS refer to oxidized, low-spin, substrate, and high-spin, respectively. K_{1d} is the equilibrium constant for the diffusion step and K_2 is the spin equilibrium constant. Experimental measurements of K_d for the overall binding process and K_2 permit K_{1d} to be determined. Measurements for wt and D251N P450cam as a function of ionic strength, solvent dielectric constant, and temperature have been made (Deprez et al., 1994). It was observed that the D251N mutant has a slightly greater overall binding affinity for camphor than wt P450cam. Thermodynamic parameters for the second step are almost identical for wt and the D251N mutant. ΔG for the first step slightly favors binding to the D251N mutant over wt at low ionic strength, but at high ionic strength, the thermodynamic parameters are essentially the same for the first step. The difference in ΔG for the first step at low salt is small (∼0.5 kcal/mol) due to compensation of larger differences (∼3 kcal/mol) between mutant and wt proteins in the unfavorable enthalpy and favorable entropy terms. The van't Hoff plot for wt P450cam shows a deflection at 11 °C that is not present for the mutant (Deprez et al., 1994; Griffen & Peterson, 1972).

Experimental measurements of the free energy barrier heights for the two steps have not been reported. However, photoacoustic calorimetry experiments have been performed that provide enthalpy and volume changes on relaxation of the protein after CO photodissociation (Di Primo et al., 1997). The enthalpy change in the wt protein is dependent on ionic strength, while in the D251N mutant, it is independent of ionic strength and similar to the value for wt at high ionic strength. The measured enthalpy changes may thus relate to re-formation of salt links perturbed on opening the substrate access channel. Fast substrate on-rates (for the complete binding event—steps 1 and 2) of $\sim 2 \times 10^7$ M^{-1} s^{-1} were measured for both wt and the D251N mutant (Gerber & Sligar, 1992). The D251N mutation was, however, observed to have a large effect on the activation enthalpy of the camphor on-rate. Camphor off-rates of about 20 s^{-1} were measured for wt and the D251N mutant and the mutation had no effect on the activation enthalpy for dissociation of substrate (Gerber & Sligar, 1992) or for protein relaxation after CO dissociation (Deprez et al., 1994).

These observations indicate that the two salt links, Asp251−Lys178 and Asp251−Arg186, participate in the regulation of substrate binding. There are no differences in structure between the substrate-free and substrate-bound crystal structures of P450cam that show how this regulation may occur (Poulos et al., 1986, 1987). However, increased temperature factors in the substrate-free form for Asp251 and residues in the region of Thr185 point to a role for the salt-link tetrad. Since the proposed entrance channel, as viewed in the crystal structures, is not large enough to allow camphor penetration, its opening to allow substrate access must involve substantial protein deformation, which would disrupt the salt links to Asp251. The opening and closing of the substrate entrance channel can be considered as a gate controlling binding and subsequent reaction. It is unclear whether the opening of the gate is an activated process or whether the protein is constantly undergoing fluctuations in which the channel periodically opens and closes. This applies as much to the entrance of camphor as to its exit from the binding site and also to the exit of the product, 5-hydroxycamphor, which might leave via the same or another channel. Recent studies of the binding of ligands in hydrophobic cavities in T4 lysozyme created by site-directed mutagenesis (Feher et al., 1996), for which protein backbone motions of 1−2 Å are necessary for ligand access to the cavities, show ligand on-rates of 10^6−10^7 M^{-1} s^{-1} with a corresponding activation barrier to binding of 2−5 kcal/mol. NMR experiments indicated that T4 lysozyme is undergoing substantial fluctuations on the microsecond to millisecond time scale, which permit the ligands to enter the protein core. The comparable on-rate for camphor to P450cam indicates that it may have a similarly low activation barrier to binding. This would suggest that, in the absence of substrate, P450cam is undergoing fluctuations that allow the substrate access channel to open and close but that the dynamics is altered by the presence of salt links to D251.

The computed stability of the salt links is an indicator of the electrostatic energy required to perturb or break them. The structure of the state in which the substrate access channel is open is unknown but, considering electrostatic interactions only, could be approximately modeled by the isolated side chains of the salt links in solvent. At least, the energy change on breaking a salt link and taking the side chains to infinite separation in solvent can provide an upper limit on the electrostatic free energy change on perturbing that salt link during opening of the substrate access channel. Opening will of course also involve significant changes in packing energetics.

The barrier for the first step is clearly influenced by residue 251. It may arise in part due to the perturbation of one or both salt links to Asp251 in wt P450cam. In the mutant, the barrier may be determined by perturbation of hydrogen bonds to Asn251 or by perturbation of other interactions that become rate-limiting in the D251N mutant. Clearly, the electrostatic continuum approach used here is limited by its static nature and does not provide complete free energies or enthalpies. Molecular dynamics simulations of the whole protein could provide a more detailed description, but treatment of the large structural changes on substrate access to/exit from the active site requires use of nonstandard techniques (Lüdemann and Wade, personal communication). Thus, we cannot make definite statements as to which interactions are broken on substrate channel opening and which determine the barrier height. However, the calculations do provide the following insights:

(1) The two salt links to Asp 251 are unusually stable electrostatically. Given the experimental evidence of their role in substrate binding, this indicates that the protein has evolved them for the purpose of regulating substrate access to the active site. Protection of the catalytic site from bulk solvent is important to prevent side reactions and uncoupling. Thus the enzyme requires a way to bury and desolvate the active site and at the same time allow the substrates and products to diffuse in and out of it. This requires protein flexibility. Consider, hypothetically, that P450cam were a totally hydrophobic protein with a buried active site. It would be very difficult to open a channel in the protein because of the high cost of solvating the hydrophobic residues in the protein. The same problem might arise, though to a lesser extent, if the salt-link tetrad residues were replaced by their hydrophobic isosteres. Thus some polar and charged residues are required in the protein to permit large motions sufficient for substrate penetration to the active site. P450cam appears to have acquired particularly stable salt links for this purpose.

It has been suggested that, in general, functional residues in enzymes are not optimized for stability (Shoichet et al., 1995). Studies on T4 lysozyme show that functionally important residues that bind substrate in the active site can be mutated such that enzymatic activity is reduced but protein stability is increased. The residues making very stable salt links to the heme might be seen to follow this rule if the reference state was taken to be the apoprotein without heme rather than the heme-bound form used here. However, mutation of Asp251 to another residue is unlikely to be stabilizing. It does not fit the "rule" because its intramolecular, rather than intermolecular, interactions are important to the protein's function. Note, however, that the electrostatic calculations indicate that replacement of the complete entrance channel tetrad by hydrophobic isosteres would, while resulting in loss of function, be stabilizing.

(2) If the two salt links to residue 251 are perturbed on channel opening, allowing release of the F−G flap, the stability of the 178−251−186 triad should be considered. As shown in Table 4, the triad is less stable in the D251N mutant than in wt. The ΔG_{tot} values for each triad are smaller in magnitude than the sum of the ΔG_{tot} values for the individual salt links because of the cooperativity of interactions between residues.

(3) The energetic barrier to channel opening may, however, correspond to the perturbation of a single salt link. Its perturbation would then weaken the second salt link to Asp251 and other interactions in the protein. Perturbation of one of the salt links to Asp251 in wt P450cam would be less favorable than perturbing one of the Asn251 hydrogen bonds in the D251N mutant. However, the Asp97−Lys197 salt link, which connects the F−G loop to the short B′ helix and is the most stable isolated salt link, has a $\Delta G_{tot} = -4$ kcal/mol. This is more stable than the 178−251 and 186−251 interactions in the modeled D251N mutant. Thus its disruption could contribute to the barrier for the D251N mutant and would result in a smaller difference in barrier heights between wt and the D251N mutant. This salt link might also contribute to the ionic strength and solvent dielectric dependence observed (Deprez et al., 1994) in the D251N mutant. While the ionic strength dependence of the equilibrium constant for the first binding step is much weaker than for wt, the dependence on solvent dielectric constant is almost the same, showing that there are still important electrostatic interactions influencing substrate binding in the D251N mutant. This proposal could be investigated by site-directed mutagenesis of residues 97 and 197.

(4) As ionic strength and solvent dielectric constant increase, the binding of camphor during the first step becomes favored. Computed changes in stability of the salt links to residue 251 and the 178−251−186 triad as a function of ionic strength and dielectric constant (see Table 6) are qualitatively consistent with the experimental data. That is, increasing the dielectric constant or ionic strength destabilizes the salt links and therefore leads to a more favorable binding free energy when camphor binds and stabilizes the protein. The differences between wt and the D251N mutant reflect the dampening of electrostatic interactions by substituting a dipole for a unit charge. However, they are not quantitatively consistent in terms of the relative magnitudes of the stability differences for wt and the D251N mutant. This may be because of deficiencies in the modeling of the D251N mutant or because other factors, such as perturbation of the Asp97−Lys197 salt link, become rate-limiting in the D251N mutant.

CONCLUSIONS

The main results ensuing from this work on the role of salt links in the stability and function of cytochrome P450cam are as follows: (1) Salt links may either stabilize or destabilize a protein (i.e., be stable or unstable) with respect to mutation to their hydrophobic isosteres. For stabilization of a salt link, it is necessary for the salt link either to be part of a salt-link network making cooperative interactions or to be surrounded by polar groups that make hydrogen bonds to it. (2) The electrostatic stabilization of a salt link is almost completely due to favorable interactions with protein atoms within 10 Å of the center of the salt link. Removal of a stable salt link removes these favorable interactions. (3) Groups of charged residues can associate in loosely connected networks of salt links that, as a whole, are stable with respect to mutation to their hydrophobic isosteres because of favorable local interactions with the surrounding protein. (4) Cytochrome P450cam seems to have evolved unusually stable salt links to have functional roles: they bind the heme cofactor and control substrate access to the buried active site by regulating protein motions, necessary for creating a substrate access channel.

ACKNOWLEDGMENT

We thank Gaston Hui Bon Hoa for bringing the subject of electrostatic control of substrate access to the active site of cytochrome P450cam to our attention. We thank Razif Gabdoulline for many fruitful discussions on protein electrostatics, Carmelo di Primo and Gaston Hui Bon Hoa for sharing photoacoustic calorimetry data in advance of publication, Volkhard Helms for assistance with parameterization and valuable discussions, and Susanna Lüdemann for critical reading of the manuscript.

APPENDIX A: MODELING OF THE D251N MUTANT

(1) Modeling Methods. Modeling and molecular dynamics simulations were performed for the D251N mutant of P450cam with camphor bound, for which no crystal structure

is available. The X-ray structure of the substrate-bound wild-type P450cam (Poulos et al., 1987) was used as a template for the modeling. The polar hydrogen positions assigned to the wild-type P450cam were retained and nonpolar hydrogen atoms were built with the Sybyl program (Tripos Associates, St. Louis, MO, 1994). The carboxylic acid group of Asp251 was then manually mutated to an amide group. Modeling is dependent on which of the two carboxylic oxygens is substituted by an NH_2 group and we chose the substitution that resulted in the least disruption of hydrogen bonding. The carboxylic oxygen making hydrogen bonds to Lys178 and Arg186 (see Figure 5) was retained while the other, which makes one hydrogen bond to Arg186 and one to the hydroxyl group of Thr185, was substituted by an NH_2 group, thus removing the second hydrogen bond to Arg 186. This led to an unfavorable conformation due the proximity of the NH_2 group and the positively charged Arg186 side chain. Rotation of the Arg186 side-chain dihedrals, and then the Asp182 side-chain dihedral in order to maintain the hydrogen bond between Arg186 and Asp182, was performed and followed by a short minimization in order to locally relax the structure of the mutant. This resulted in a "hand-built" model of the D251N mutant.

This modeled structure was relaxed and refined by molecular dynamics simulation using the ARGOS program (Straatsma & McCammon, 1990) with the CHARMM22 all-atom force field (Brooks et al., 1983; QUANTA parameter-handbook, MSI, 1992) and the TIP3P water model (Jorgensen et al., 1983). Parameters for the heme were assigned from Kuczera et al. (1990), but the partial atomic charges were modified from those in the CHARMM22 force field to be suitable for a ferric heme with a thiolate ligand (Cys357). The additional +1.4e net charge was distributed over the porphyrin ring as follows: the partial charge of the four pyrrole nitrogens were each increased by +0.1e, whereas the 20 unsaturated carbons were assigned an additional +0.05e, leading to a ferric porphyrin with a net charge of −0.6e. In addition, the partial charges of the Cys357 side chain were zeroed except at the sulfur atom, which was given a partial charge of −0.4e. The covalent linkage between the Cys357 thiolate and the heme iron was defined following Helms and Wade (1995). Camphor was assigned CHARMM22 parameters with the following template partial atomic charges: all hydrogens, +0.05e; tetrasubstituted C, 0.0e; trisubstituted C, −0.15e; disubstituted C, −0.10e; monosubstituted C, −0.05e; keto C, +0.60e; and keto O, −0.55e.

In order to ensure electroneutrality of the simulated system, 13 positively charged counterions were positioned at energy minima below −15 kcal/mol in the sodium ion interaction map evaluated with the GRID program (Goodford, 1985). At this energy level, none of the counterions are located within 10 Å of any atoms in the entrance channel tetrad. In the energy interval from −15 to −10 kcal/mol, however, a spread of energy minima for a sodium ion exists in close vicinity to the entrance channel tetrad, including two near the carboxylic acid group of Glu156. Prior to simulation, the system was immersed in a sphere of water molecules of 20 Å radius centered on the Cγ of Asn251. Crystallographic water molecules were kept and a minimum distance criterion of 2.6 Å from any protein atom or crystal water was used for superposition of waters of solvation. Water molecules that were automatically positioned near the heme plane or in the active-site pocket (solely on the basis of steric criteria) were removed.

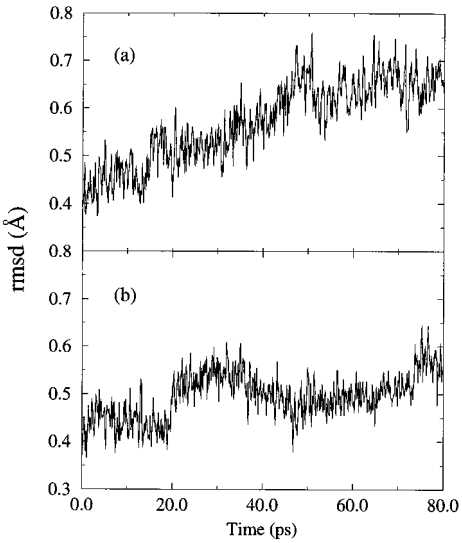


FIGURE 6: Rmsd for backbone non-hydrogen atoms from the crystal structure versus time (in picoseconds) during the molecular dynamics simulation of wild-type P450cam with Glu156 (a) unprotonated and (b) protonated. The sharp jump in the rmsd in panel b after 20 ps reflects a rotation of 92° around the χ^2 side chain torsion of Leu250.

The solvated region of the protein and all solvent molecules were then subjected to a maximum of 200 steps of steepest descent energy minimization to release bad contacts. This was followed by molecular dynamics simulation of a 10 Å radius sphere of dynamic atoms centered on Cγ Asn251. The surrounding atoms were kept fixed in their positions after minimization. The resulting system consisted of 540 dynamic solute atoms and 14 dynamic water molecules. A twin-range cutoff of 8 and 12 Å was used for short- and long-range interactions, respectively. A time step of 2 fs was used and bonds were constrained with the SHAKE algorithm (Ryckaert et al., 1977). Heating of the system was achieved in a two-step procedure involving 2 ps NVE runs at 150 and 300 K with velocity reassignment every 20 time steps. Then a 40−80 ps NVT run was performed.

Two such simulations were performed for the D251N mutant: one with Glu156 unprotonated and one with it protonated. Glu156 is the closest titratable residue to the entrance tetrad (distance E156 Cδ−K178 Nζ = 5.7 Å). In addition, two control simulations were performed for the wild-type P450cam using the same procedure as for the D251N mutant for both protonation states of Glu156.

(2) Resultant Models. Control simulations with the wild-type enzyme showed that the system equilibrated faster and with a smaller RMSD from the crystal structure when Glu156 was treated as protonated rather than unprotonated (see Figure 6). The same features were seen for the D251N mutant.

Detailed views of the average structures of the wild-type protein from these two simulations are shown in Figure 7. When Glu156 is unprotonated, the hydrogen bond between residues 178 and 251 is broken and the Lys178 side chain reorients and forms a new salt link with Glu156. Glu156 undergoes distortion due to the interaction with Lys178 and the carboxylic group of Glu156 rotates by 90° with respect to its crystallographically observed orientation. Only three

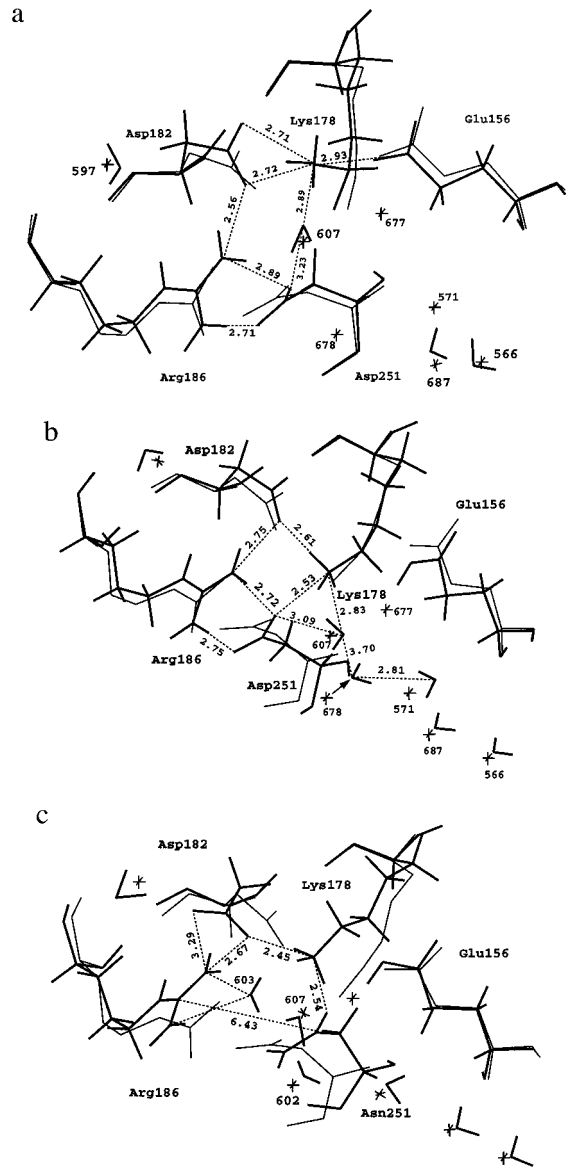


FIGURE 7: Detail of the tetrad of salt links at the entrance of the proposed substrate access channel showing the average structures resulting from the simulations. (a) Wild-type P450cam with Glu156 unprotonated; (b) wild-type P450cam with Glu156 protonated; and (c) the D251N mutant with Glu156 protonated. The averages are carried out over the last 10 ps of the trajectory (a) or over the last 20 ps (b, c). The simulation averages are superimposed on the X-ray structure (thin lines). Average positions of conserved water molecules are shown in thick lines. Hydrogen bonds are represented by dashed lines, and their lengths are reported in angstroms.

of the six crystallographically observed water molecules in the 10 Å radius sphere of fully dynamic atoms stay close to their experimentally observed positions in the simulations with Glu156 unprotonated, and the waters near Glu156 are particularly perturbed from their crystallographic positions.

When Glu156 is protonated, the overall structure of the entrance channel tetrad, including the network of hydrogen bonds, is maintained compared to the crystallographic data. The only significant difference is that the position of crytallographic water 677 hydrogen-bonded to Glu156 is not conserved (see Figure 7b) and the crystallographic water 678 moves 0.9 Å from its position in the crystal structure so as to partially compensate for the loss of water 677. Neutralization of the Glu156 negative charge thus results in stabilization of the tetrad in its native stage but not complete conservation of the solvent structure. Consistent with this, pK_a calculations (data not shown) indicate that Glu156 has a raised pK_a compared to the standard value for Glu but that it is not above 7. An alternative to protonation would be placement of a cation in one of the energetically favorable positions adjacent to Glu156. However, no cation was observed crystallographically so we have not pursued this possibility further.

On the basis of the results for wild-type P450cam, we expect that the structure derived from simulations of the D251N mutant with Glu156 protonated to be more realistic than those from simulations with Glu156 unprotonated. The simulation of the D251N mutant with Glu156 protonated shows that the four residues of the entrance tetrad are perturbed by the D251N mutation. They all undergo conformational reordering (see Figure 7c). The 186−251 salt link is broken, leaving an open groove between these residues. On the other hand, the hydrogen bonds from Lys178 to Asp182 and Asn251 are maintained. The deviation of Glu156 from the X-ray structure is minimal. The network of crystallographic waters undergoes the same changes as in the wild-type simulation. Strikingly, in all four simulations, water 607 is maintained in its X-ray location where it forms hydrogen bonds to the side-chain nitrogen of Lys178 and the side-chain oxygen of Asn251. The model does not point to any restrictions in the motion of Asn251 that might explain the role of D251 in a proposed distal charge relay (Gerber & Sligar, 1992) during the catalytic reaction.

APPENDIX B: ANALYSIS OF SENSITIVITY TO NUMERICAL PARAMETERS

Discretization Parameters. To evaluate sensitivity to grid extent and spacing, a set of calculations were performed for the Lys178−Asp251 salt link with different sets of grids. The grid sets (given in Table 7) were chosen to investigate (i) whether a grid spacing of 0.4 Å is sufficiently small and a dimension of 50 × 50 × 50 points is sufficiently large for the final focused grid and (ii) whether the chosen coarser grids provided sufficiently accurate potential values at the boundaries of the final focused grid.

Electrostatic free energies calculated for the Lys178−Asp251 salt link in tests of sensitivity to grid extent and spacing are given in Table 7. Rotational and translational averaging was performed for grid set 1 from 16 calculations, each with the molecule at a different position on the same initial grid. The positions were determined from combinations of 0.3 and 0.5 Å translations in the *x*, *y*, and *z* directions and 45° and 90° rotations around the *x*, *y*, and *z* axes. The small standard deviations (about 0.2 kcal/mol) indicate that a spacing of 0.4 Å for the final focused grid leads to reliable estimates of all the ΔG terms. This is confirmed by a comparison of the results for grid sets 2 and 3, which differ only in the grid spacing of the final focused grid (0.4 Å for set 2 and 0.3 Å for set 3). The differences in the ΔG terms are 0.1 kcal/mol or less. In addition, from comparison of sets 1 and 2, it can be seen that the use of finer outer grids to derive the boundary potential in set 2 has a minimal effect on the ΔG terms.

In set 4, the extent of the final focused grid was increased and four grids were used. This is the most accurate and computationally expensive calculation. The ΔG terms from

Table 7: Sensitivity of the Calculated Electrostatic Free Energies[a] of the Lys178−Asp251 Salt Link in Cytochrome P450cam to Grid Size and Spacing[a]

grid set	grid spacing (Å)	grid dimensions (no. of points)	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}
1	7.5	50 × 50 × 50	−7.34	29.14	−14.45	−22.03	1.73
	3.6	50 × 50 × 50	[−7.42]	[28.93]	[14.30]	[−22.05]	[1.74]
	0.4	50 × 50 × 50	(0.18)	(0.22)	(0.15)	(0.18)	(0.22)
2	2.5	150 × 150 × 150					
	1.2	150 × 150 × 150	−7.38	29.27	−14.57	−22.09	1.73
	0.4	50 × 50 × 50					
3	2.5	150 × 150 × 150					
	1.2	150 × 150 × 150	−7.41	29.16	−14.58	−21.98	1.71
	0.3	60 × 60 × 60					
4	2.5	150 × 150 × 150					
	1.2	150 × 150 × 150	−7.63	29.13	−14.56	−22.20	1.71
	0.6	150 × 150 × 150					
	0.3	150 × 150 × 150					

[a] Given in kilocalories per mole. [b] The values given in square brackets and parentheses are the average and standard deviation, respectively, from rotational and translational averaging calculations.

Table 8: Comparison of the Electrostatic Free Energies[a] Computed for Two Salt Bridges in T4 Lysozyme with Different Protocols

salt link	grid set[b]	boundary smoothing	program[c]	ΔG_{tot}	$-\Delta G_{sol}$	ΔG_{brd}	ΔG_{prt}	ΔG_{ass}
H31−D70	A^1	N	UHBD	3.53 (0.57)	12.34 (0.73)	−8.49 (0.25)	−0.31 (0.05)	−0.69 (0.48)
	A*	N	UHBD	3.56 (0.46)	11.95 (0.45)	−8.08 (0.20)	−0.31 (0.05)	−0.54 (0.31)
	A	N	Delphi	3.46 (0.26)	11.62 (0.35)	−7.91 (0.21)	−0.26 (0.09)	−0.86 (0.17)
	B	Y	HHBD	3.01	11.83	−8.61	−0.20	−1.06
	B	N	UHBD	3.36	11.81	−8.25	−0.21	−0.73
	C	Y	UHBD	2.75	11.80	−8.67	−0.37	−1.18
	C	N	UHBD	2.78	11.70	−8.58	−0.34	−1.08
E11−R145	A	N	UHBD	9.78 (1.02)	18.39 (1.05)	−8.21 (0.27)	−0.39 (0.13)	1.64 (1.15)
	A	N	Delphi	10.13 (1.03)	18.66 (1.47)	−7.72 (0.82)	−0.81 (0.15)	0.45 (0.54)
	B	Y	UHBD	7.61	16.70	−8.68	−0.41	0.28
	B	N	UHBD	8.81	17.81	−8.61	−0.39	1.05
	C	Y	UHBD	7.57	16.73	−8.69	−0.47	0.35
	C	N	UHBD	7.80	16.89	−8.56	−0.53	0.48

[a] Given in kilocalories per mole. [b] Grid sets are defined as follows. Grid set A: Two 65 × 65 × 65 grids with grid spacings of 1.6 and 0.8 Å. A* is identical to A^1 but with the His31 Nδ1−Hδ1 bond length = 0.98 Å instead of 1.03 Å. Results for grid set A are given as the average and standard deviation (in parentheses) from rotational and translational averaging calculations. Grid set B: Grid parameters generally used in this work (set 1 in Table 7). Grid set C: Two 150 × 150 × 150 grids with spacings of 2.5 and 1.2 Å and two 50 × 50 × 50 grids with grid spacings of 0.4 and 0.2 Å, respectively. [c] All numbers computed with the Delphi program are from Hendsch and Tidor (1994).

sets 1 and 4 differ by 0.3 kcal/mol at most. Thus the computationally much more efficient grid set 1 was used for calculations for all other salt links in P450cam.

Consistency Checks for Salt Links in T4 Lysozyme. To examine the consistency of our results with those of Hendsch and Tidor (1994), we applied our procedure to two salt bridges, His31−Asp70 and Glu11−Arg145, in T4 lysozyme that they studied. Hendsch and Tidor employed the same thermodynamic decomposition and physical parameters but they did not do dielectric boundary smoothing and they used different grid spacings and dimensions (two 65 × 65 × 65 grids with grid spacings of approximately 1.6 and 0.8 Å). When UHBD calculations were performed with these grid parameters (grid set A in Table 8), the results agree within error bars with those reported by Hendsch and Tidor using the Delphi program, although the computed ΔG_{sol}, ΔG_{brd}, and ΔG_{prt} values can differ by up to 0.7 kcal/mol. These differences arise from differences in the averaging procedure, grid spacing, and the placement of protons (see Table 8 for effect of altering the length of the Nδ−H bond in His31). However, the energy values for grid set A are not fully converged, as can be seen from the calculations with sets B (corresponding to those used in our P450cam calculations) and C (with a spacing of 0.2 Å for the final grid). Good agreement between calculations with and without smoothing is obtained for set C. The energies computed with grid set B with a final grid spacing of 0.4 Å and smoothing are close to those obtained with grid set C. This shows, as noted previously (Mohan et al., 1992), that convergence is better when dielectric boundary smoothing is performed. All other calculations were therefore performed with grid set B and smoothing.

REFERENCES

Anderson, D. E., Beckel, W. J., & Dahlquist, F. W. (1990) *Biochemistry 29*, 2403−2408.

Antosiewicz, J., McCammon, J. A., & Gilson, M. K. (1994) *J. Mol. Biol. 238*, 415−436.

Barlow, D. J., & Thornton, J. M. (1983) *J. Mol. Biol. 168*, 867−885.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Dao-pin, S., Nicholson, H., & Matthews, B. W. (1991) *Biochemistry 30*, 7142−7153.

Davis, M. E., & McCammon, J. A. (1989) *J. Comput. Chem. 10*, 386−391.

Davis, M. E., & McCammon, J. A. (1991) *J. Comput. Chem. 12*, 909−912.

Demchuk, E., & Wade, R. C. (1996) *J. Phys. Chem. 100*, 17373−17387.

Deprez, E., Gerber, N. C., Di Primo, C., Douzou, P., Sligar, S. G., & Hui Bon Hoa, G. (1994) *Biochemistry 33*, 14464−14468.

Di Primo, C., Deprez, E., Sligar, S. G., & Hui Bon Hoa, G. (1997) *Biochemistry 36*, 112−118.

FASEB (1992) *FASEB J. 6*, 667−792.
FASEB (1996) *FASEB J. 10*, 202−214.
Feher, V., Baldwin, E. P., & Dahlquist, F. W. (1996) *Nat. Struct. Biol. 3*, 516−521.
Fersht, A. R. (1972) *J. Mol. Biol. 64*, 497−509.
Gandini, D., Gogioso, L., Bolognesi, M., & Bordo, D. (1996) *Proteins: Struct., Funct., Genet. 24*, 439−449.
Gerber, N. C., & Sligar, S. G. (1992) *J. Am. Chem. Soc. 114*, 8742−8743.
Gerber, N. C., & Sligar, S. G. (1994) *J. Biol. Chem. 269*, 4260−4266.
Gilson, M. K., Sharp, K. A., & Honig, B. H. (1987) *J. Comput. Chem. 9*, 327−335.
Goodford, P. J. (1985) *J. Med. Chem. 28*, 849−857.
Griffin, B. W., & Peterson, J. A. (1972) *Biochemistry 11*, 4740−4746.
Hasemann, C. A., Ravichandran, K. G., Peterson, J. A., & Deisenhofer, J. (1994) *J. Mol. Biol. 236*, 1169−1185.
Hasemann, C. A., Kurumbail, R. G., Boddupalli, S., Peterson, J. A., & Deisenhofer, J. (1995) *Structure 3*, 41−62.
Helms, V., & Wade, R. C. (1995) *Biophys. J. 69*, 810−824.
Hendsch, Z. S., & Tidor, B. (1994) *Protein Sci. 3*, 211−226.
Hooft, R. W. W., Sander, C., & Vriend, G. (1996) *Proteins: Struct., Funct., Genet. 26*, 363−376.
Horowitz, A., Serrano, L., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *J. Mol. Biol. 216*, 1031−1044.
Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−935.
Koga, H., Sagara, Y., Yaoi, T., Tsujimara, M., Nakamura, K., Sekimizu, K., Nakino, R., Shimada, H., Ishimura, Y., Yura, K., Go, M., Ikeguchi, M., & Horiuchi, T. (1993) *FEBS Lett. 331*, 109−113.
Kuczera, K., Kuriyan, J., & Karplus, M. (1990) *J. Mol. Biol. 213*, 351−373.
Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
Lumb, K. J., & Kim, P. S. (1995) *Science 268*, 436−439.
Lyu, P. C., Gans, P. J., & Kallenbach, N. R. (1992) *J. Mol. Biol. 223*, 343−350.
Madura, J. D., Davis, M. E., Gilson, M., Wade, R. C., Luty, B. A., & McCammon, J. A. (1994) *Rev. Comput. Chem. 5*, Chapt. 4.
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., & McCammon, J. A. (1995) *Comput. Phys. Commun. 91*, 57−95.
Mohan, V., Davis, M. E., McCammon, J. A., & Pettitt, B. M. (1992) *J. Phys. Chem. 96*, 6328−6431.
Musafia, B., Burchner, V., & Arad, D. (1995) *J. Mol. Biol. 254*, 761−770.
Perutz, M. (1978) *Science 201*, 1187−1191.
Pfeil, W., Noelting, B. O., & Jung, C. (1993) *Biochemistry 32*, 8856−8862.
Poulos, T. L. (1995) *Curr. Opin. Struct. Biol. 6*, 767−774.
Poulos, T. L., & Raag, R. (1992) *FASEB J. 6*, 674−679.
Poulos, T. L., Finzel, B. C., & Howard, A. J. (1986) *Biochemistry 25*, 5314−5322.
Poulos, T. L., Finzel, B. C., & Howard, A. J. (1987) *J. Mol. Biol. 195*, 687−700.
Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A., & Deisenhofer, J. (1993) *Science 261*, 731−736.
Richards, F. M. (1979) *Carlsberg Res. Commun. 44*, 47−63.
Ryckaert, J.-P., Cicotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327−341.
Schueler, O., & Margalit, H. (1995) *J. Mol. Biol. 248*, 125−135.
Shoichet, B. K., Baasee, W. A., Kuroki, R., & Matthews, B. W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 452−456.
Sligar, S. G. (1976) *Biochemistry 24*, 5399−5406.
Straatsma, T. P., & McCammon, J. A. (1990) *J. Comput. Chem. 11*, 943−951.
Vriend, G. (1990) *J. Mol. Graph. 8*, 52−56.
Waldburger, C. D., Schilbach, J. E., & Sauer, R. T. (1995) *Struct. Biol. 2*, 122−128.
Warshel, A., & Russel, S. T. (1984) *Q. Rev. Biophys. 17*, 283−422.
Wimley, W. C., Gawrisch, K., Creamer, T. P., & White, S. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2985−2990.
Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnivoka, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R., & Rice, D. W. (1995) *Structure 3*, 1147−1158.

BI9622940