IUBMB *Life*, 53: 85–98, 2002

1521-6543/02 $12.00 + .00
DOI: 10.1080/10399710290038972

Review Article

Structural Basis of Perturbed pK_a Values of Catalytic Groups in Enzyme Active Sites

Thomas K. Harris[1] and George J. Turner[2]
[1]*Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida*
[2]*Department of Physiology and Biophysics and the Neurosciences Program, University of Miami School of Medicine, Miami, Florida*

Summary

In protein and RNA macromolecules, only a limited number of different side-chain chemical groups are available to function as catalysts. The myriad of enzyme-catalyzed reactions results from the ability of most of these groups to function either as nucleophilic, electrophilic, or general acid–base catalysts, and the key to their adapted chemical function lies in their states of protonation. Ionization is determined by the intrinsic pK_a of the group and the microenvironment created around the group by the protein or RNA structure, which perturbs its intrinsic pK_a to its functional or apparent pK_a. These pK_a shifts result from interactions of the catalytic group with other fully or partially charged groups as well as the polarity or dielectric of the medium that surrounds it. The electrostatic interactions between ionizable groups found on the surface of macromolecules are weak and cause only slight pK_a perturbations (<2 units). The sum of many of these weak electrostatic interactions helps contribute to the stability of native or folded macromolecules and their ligand complexes. However, the pK_a values of catalytic groups that are found in the active sites of numerous enzymes are significantly more perturbed (>2 units) and are the subject of this review. The magnitudes of these pK_a perturbations are analyzed with respect to the structural details of the active-site microenvironment and the energetics of the reactions that they catalyze.
IUBMB *Life*, 53: 85–98, 2002

Keywords Acetoacetate decarboxylase; bacteriorhodopsin; cysteine protease; glycosidase; serine protease; thioredoxin.

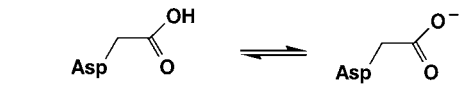
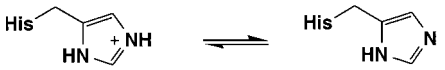
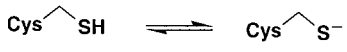
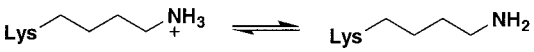
Received 26 November 2001; accepted 28 January 2002.
Address correspondence to Thomas K. Harris, University of Miami School of Medicine, Department of Biochemistry and Molecular Biology (R-629), P. O. Box 016129, Miami, FL 33101-6129, USA. Fax: 305-243-3955. E-mail: tkharris@miami.edu

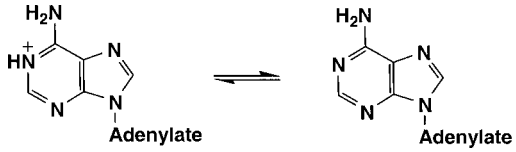
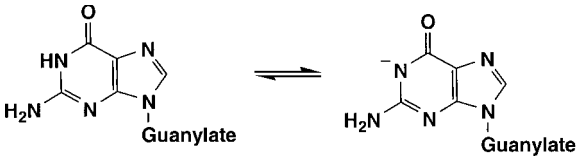
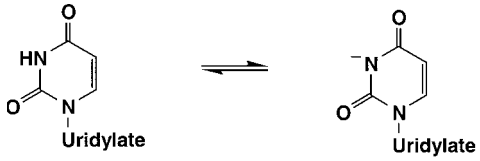
Abbreviations: AAD, acetoacetate decarboxylase; AbAld, antibody aldolase; Ala-Race, alanine racemase; hdvAR, hepatitis delta virus antigenomic ribozyme; ArsC, arsenate reductase; AspAT, aspartate aminotransferase; BCX, *Bacillus circulans* xylanase; BR, ground-state bacteriorhodopsin with all *trans* retinal, protonated D96, protonated Schiff base, and unprotonated D85; BR-M, excited M-state bacteriorhodopsin with 13-*cis* retinal, protonated D96, unprotonated Schiff base, and protonated D85; BR-N, excited N-state bacteriorhodopsin with 13-*cis* retinal, unprotonated D96, protonated Schiff base, and protonated D85; Chy-TFKs, chymotrypsin complexed with various peptidyl trifluoroketones; hmCK, human muscle creatine kinase; DsbA and DsbC, disulfide bond enzymes A and C in *E. coli*; GalE, UDP-galactose 4-epimerase; GRX, glutaredoxin; HB, hydrogen bond; HPE, hydrophobic environment; KSI, ketosteroid isomerase; LBHB, low-barrier hydrogen bond; NMR, nuclear magnetic resonance; NTα, *N*-terminus of an α-helix; Nuc/Lv, nucleophile/leaving group; 4-OT, 4-oxalocrotonate tautomerase; PDI, protein disulfide isomerase; PLP, pyridoxal 5′-phosphate; PMP, pyridoxamine 5′-phosphate; blmPTP, bovine liver low molecular weight protein tyrosine phosphatase; hPTP1, human protein tyrosine phosphatase; yPTP, *Yersenia* protein tyrosine phosphatase; rPTC, ribosomal peptidyl transferase center; RNaseH1, ribonuclease H1; TIM, triosephosphate isomerase; bTRX, bacterial thioredoxin; hTRX, human thioredoxin.

INTRODUCTION

The nature of enzymatic rate enhancement entails knowledge of the chemistry of the individual catalytic groups commonly found in enzymes as well as the ensemble of native protein or RNA structures that form a given substrate's binding site. The chemistry of the different enzymatic catalytic groups can be classified into four categories: nucleophiles, electrophiles, general-base catalysts, and general-acid catalysts (*1*). The common *titratable catalytic* groups in proteins include the *C*-terminal carboxyl, the carboxyl groups of aspartic and glutamic acids, the imidazole of histidine, the sulfhydryl of cysteine, the amino group of lysine, the hydroxyl of tyrosine, and the *N*-terminal amino group (Table 1A). The positively charged side-chain guanidinium group of arginine residues is often utilized as an electrophilic catalyst, but it is rarely titratable due to its very high pK_a of 12. In nucleic acids, recent structural and mechanistic studies now indicate that the ring nitrogens of the nucleoside bases can function as general acid-base catalysts (Table 1B). Yet, with this limited repertoire of chemical groups, proteins and RNA catalyze a myriad of biological reactions. In part, this is due to the ability of many of these groups to function either as

Table 1
Intrinsic pK_a values of ionizable side-chain groups of proteins and RNA in aqueous solution

(A) Amino-acid side-chain ionizations (*2*)	pK_a[a]
	4.0
	4.5
	6.4
	9.1
	9.3
	9.7
	10.4

(B) RNA side-chain ionizations (*4*)	pK_a
	3.5
	4.2
	9.4
	9.4

[a]Intrinsic values for side-chain functional groups in proteins and peptides present complex issues, because the values for the amino acids listed in this table are not necessarily appropriate. The best values would be for the side chains appearing in small peptides, so that the α-amino and carboxylate groups do not interact internally with the ionizing group. However, placement of an ionizable group within a model peptide that contains no nearby charged groups is often difficult due to the lack of solubility of these peptides. A more comprehensive listing of pK_a values of amino acid side-chains that have been measured in the context of either small peptides or as derivatives of amino acids has been described (*3*). For example, the pK_a value of 9.1 for the thiol group of cysteine is slightly higher than the p$K_a = 8.9$ of this group in glutathione. In addition, the pK_a value of 9.7 for the phenol group of tyrosine is slightly lower than the pK_a of 10.2 measured for *p*-methylphenol (*p*-cresol), which is a most reasonable model compound. Thus, the intrinsic pK_a value for a given side-chain has not been fully established, and may vary by $\leq \pm 0.5$ units, depending on the location of the side-chain within a particular peptide or compound.

nucleophiles, electrophiles, general-base catalysts, or general-acid catalysts. The key to their adapted role or function lies in their state of protonation, and the state of protonation is determined, at first approximation, by the intrinsic pK_a of the given group.

Table 1 lists the intrinsic pK_a values of the catalytic groups found in either protein or RNA enzymes. For these groups to be activated as nucleophiles, the ionizing proton needs to dissociate from either the reactive oxygen, nitrogen, or sulfur atom. This requires that the effective pK_a of the group be lower than the effective pH at the reaction site. At physiological pH, only aspartate, glutamate, and histidine would be fully active as nucleophiles. For groups to be activated as electrophiles, the group needs to be in its protonated form to function either as a hydrogen bond donor or to provide a localized partial or full positive charge. This requires that the effective pK_a of the group be higher than the effective pH at the reaction site. For a group to optimally function as a general acid or base, its intrinsic pK_a value must be near the pH at the reaction site so that its functional state of protonation can be regenerated for the next enzymatic turnover, indicating histidine and cysteine to be the only candidates. However, it has been demonstrated that most of the catalytic groups in Table 1 can function as nucleophiles, electrophiles, general acids, and general bases, depending on the given enzyme. Activation of these catalytic groups is a direct result of the microenvironment created around the group by the protein or RNA structure.

When a titratable catalytic group is transferred into a specific site in a macromolecular structure, its intrinsic pK_a value is driven to its functional or apparent value by interactions with other fully or partially charged groups as well as by the polarity or dielectric of the medium that surrounds it. The ionizable groups found on the surface of proteins, accessible to bulk solvent, play a vital role in electrostatic stabilization of the native or folded protein, primarily through salt-bridge formation or ion pairing. These electrostatic interactions cause slight perturbations of their intrinsic pK_a values (<2 units), and has been the subject of much experimental and theoretical study. However, many ionizable catalytic groups that are buried in enzyme active sites are significantly more perturbed (>2 units) and are listed in Table 2. In addition, Table 2 lists the function of each catalytic group and the likely structural basis for the pK_a perturbation.

CHARGE-CHARGE INTERACTIONS

Ionization of a catalytic group is affected through long-range as well as local electrostatic effects. The electric force experienced between two interacting charges (q_1 and q_2) depends on the distance (r) and the dielectric constant (ε) of the medium between them as given by Coulomb's Law ($F = q_1q_2/\varepsilon r^2$) (*38*). An ionizable catalytic group on an enzyme can interact with other neighboring charged groups. In such a case, if both groups are deprotonated and negative in their ionized forms, then one group will tend to increase the pK_a of the other, facilitating protonation and neutralization to avoid the unfavorable Coulombic energetics of like-charge repulsion. If both groups are protonated and positive in their ionized forms, then one group will tend to decrease the pK_a of the other, facilitating deprotonation and neutralization also to avoid like-charge repulsion. If one group is positive and one group is negative in their ionized forms, then the positively charged group will tend to lower the pK_a of the negatively charged group and vice versa, facilitating the favorable energetics of opposite-charge attraction. All of these effects are increased to a greater degree as the effective dielectric constant or solvent polarity is decreased. Thus, the energetics of charge-charge interactions become more significant if these interactions are removed from aqueous solvent of high polarity and transferred to the interior of proteins where the dielectric constant, ε_{prot}, is estimated to be much lower due to hydrophobic packing.

Examples of Proximal Negative Charges

Glycosidases. *Bacillus circulans* xylanase (BCX, Table 2) catalyzes the hydrolysis of the polysaccharide xylan, a major constituent of plant biomass, in a double-displacement reaction in which Glu-78 functions as a nucleophile, forming a covalent glycosyl-enzyme intermediate (Fig. 1). Glu-172 acts as a general-acid catalyst during glycosylation, protonating the hydroxyl group of the departing aglycone (HOR), and then as a general-base catalyst to activate the attacking water molecule during the deglycosylation reaction. The pK_a values of Glu-78 and Glu-172, in the free enzyme, were measured by direct ^{13}C NMR pH titrations to be 4.6 and 6.7, respectively (*12*). The increased pK_a value of 6.7 for Glu-172, necessary for its functioning as a general-acid catalyst, appears to be due to electrostatic repulsion from the negatively charged Glu-78. Mutation of Glu-78 to the neutral glutamine causes the pK_a of Glu-172 in the free enzyme to decrease to 4.2. In a trapped covalent glycosyl-enzyme intermediate, the pK_a of Glu-172 was also near 4.2. Thus, charge neutralization of Glu-78 as a result of glycosylation results in a decrease in the pK_a of Glu-172, facilitating a thermodynamically favorable loss of this proton to the departing hydroxyl. Loss of the proton enables its subsequent functioning as a general-base catalyst to deprotonate water, activating it for hydrolysis of the glycosyl-enzyme intermediate in the deglycosylation reaction. A pair of carboxylic acids have been shown to be the catalytic groups in essentially all glycosidases described to date, and such "pK_a cycling" of the carboxylic acid that functions as the general acid-base is likely to be common among this class of enzymes.

Ribonucleases, Aspartic Proteases, and Hen Egg White Lysozyme. Smaller perturbations of the pK_a values for a pair of proximal carboxyl groups have been reported for ribonucleases, aspartic proteases, and hen egg-white lysozyme. In ribonuclease H1 (RNaseH1, Table 2), Asp-10, Asp-70, Glu-48, and a divalent cation have been shown to be essential for catalytic activity. Using two-dimensional ^1H and ^{13}C NMR, Oda et al. (*11*) measured the pK_a values for all of the carboxyl groups in RNase H1 and found that the pK_a of Asp-10 (6.1) is unusually

Table 2
The function and structural basis of catalytic groups with perturbed pK_a values on enzymes

Protein[a]	Group	ΔpK_a[b]	Function	Structural basis
BR (*5*)	D85	−1.8	Proton transport	Unknown
BR-M (*5*)	D85	+2.9	Proton transport	HPE, anionic E204
BR (*6*)	D96	>+8.0	Proton transport	HPE
BR-N (*7*)	D96	+3.1	Proton transport	HPE, solvent accessibility
bTRX (*8*)	D26	+3.5	Tune C32 pK_a	HPE
hTRX (*9*)	D26	+4.1	Tune C32 pK_a	HB to S28 OH
KSI (*10*)	D99	>+5.0	HB donor	HPE, anionic D38
RNaseH1 (*11*)	D10	+2.1	HB donor	Anionic D70
BCX (*12*)	E172	+2.2	General acid/base	Anionic E78, HB network
Chy-TFKs (*13*)	H57	+(3.9–5.7)	General acid/base	LBHB
TIM (*14*)	H95	>−1.9	Electrophile	NTα, HB to S96 NH
bTRX (*15*)	C32	−1.6	Nuc/Lv	NTα, anionic D26
hTRX (*16*)	C32	−2.8	Nuc/Lv	NTα, HB to C35 NH
GRX (*17*)	C11	>−3.6	Nuc/Lv	HB to C14 SH
	C14	>+1.4	effect C11 pK_a	Anionic C11
PDI (*18*)	C36	−4.6	Nuc/Lv	NTα, cationic H38
DsbA (*19*)	C30	−5.6	Nuc/Lv	NTα, cationic H32
DsbC (*20*)	C98	nd	Nuc/Lv	NTα, HB to C35 NH
ArsC (*21*)	C12	−2.8	Nuc/Lv	Cationic H8
Papain (*22*)	C25	−5.8	Nuc/Lv	Cationic H159
Caricain (*22*)	C25[c]	−6.2	Nuc/Lv	Cationic H159*
Ficin (*22*)	C25[c]	−6.6	Nuc/Lv	Cationic H159*
yPTP (*23*)	C403	−4.4	Nuc/Lv	NTα, cationic H402
hPTP1 (*24*)	C403[d]	−3.5	Nuc/Lv	NTα, cationic H402
blwPTP (*25*)	C403[d]	−(1.6–2.4)	Nuc/Lv	NTα
hmCK (*26*)	C282	−3.5	Unknown	HB to S284 OH
4-OT (*27*)	P1	−3.0	General acid/base	HPE
KSI (*28*)	Y14	+1.9	HB donor	HPE
Ala-Race (*29*)	Y265	−2.5	General base	Cationic R219
GalE (*30*)	Y149	−3.6	General acid/base	Cationic NAD^+ and K153, HB to S124 OH
AspAT (*31*)	K258-PLP	−3.1	General base	Imine-pyridine torsion, cationic R292 and R386
BR (*32*)	K216-retinal	+2.5	Proton transport	Unknown
BR-M (*33*)	K216-retinal	−2.2	Proton transport	Unknown
AAD (*34*)	K115	−4.4	Schiff base	Cationic K116
AbAld (*35*)	K	−(4.4–4.9)	Schiff base	HPE
rPTC (*36*)	A2451 N1	+4.1	General acid/base	HB network
hdvAR (*37*)	C76 N3	+1.9	General base	Unknown

[a] AAD, acetoacetate decarboxylase; AbAld, antibody aldolase; Ala-Race, alanine racemase; hdvAR, hepatitis delta virus antigenomic ribozyme; ArsC, arsenate reductase; AspAT, aspartate aminotransferase; BCX, *Bacillus circulans* xylanase; BR, ground-state bacteriorhodopsin with all *trans* retinal, protonated D96, protonated Schiff base, and unprotonated D85; BR-M, excited M-state bacteriorhodopsin with 13-*cis* retinal, protonated D96, unprotonated Schiff base, and protonated D85; BR-N, excited N-state bacteriorhodopsin with 13-*cis* retinal, unprotonated D96, protonated Schiff base, and protonated D85; Chy-TFKs, chymotrypsin complexed with various peptidyl trifluoroketones; hmCK, human muscle creatine kinase; DsbA and DsbC, disulfide bond enzymes A and C in *E. coli*; GalE, UDP-galactose 4-epimerase; GRX, glutaredoxin; HB, hydrogen bond; HPE, hydrophobic environment; KSI, ketosteroid isomerase; LBHB, low-barrier hydrogen bond; NMR, nuclear magnetic resonance; NTα, *N*-terminus of an α-helix; Nuc/Lv, nucleophile/leaving group; 4-OT, 4-oxalocrotonate tautomerase; PDI, protein disulfide isomerase; PLP, pyridoxal 5′-phosphate; PMP, pyridoxamine 5′-phosphate; blmPTP, bovine liver low molecular weight protein tyrosine phosphatase; hPTP1, human protein tyrosine phosphatase; yPTP, *Yersenia* protein tyrosine phosphatase; rPTC, ribosomal peptidyl transferase center; RNaseH1, ribonuclease H1; TIM, triosephosphate isomerase; bTRX, bacterial thioredoxin; hTRX, human thioredoxin.

[b] These values were calculated from the difference between the measured pK_a and the intrinsic pK_a listed for that group in Table 1, which are not fully established. Thus, caution should be observed when comparing ΔpK_a values between different side-chains, whereas comparison of ΔpK_a values of enzymes that contain the same side-chain ionizable group will be most useful since they are referenced to the same intrinsic pK_a value.

[c] Papain numbering system.

[d] yPTP numbering system.

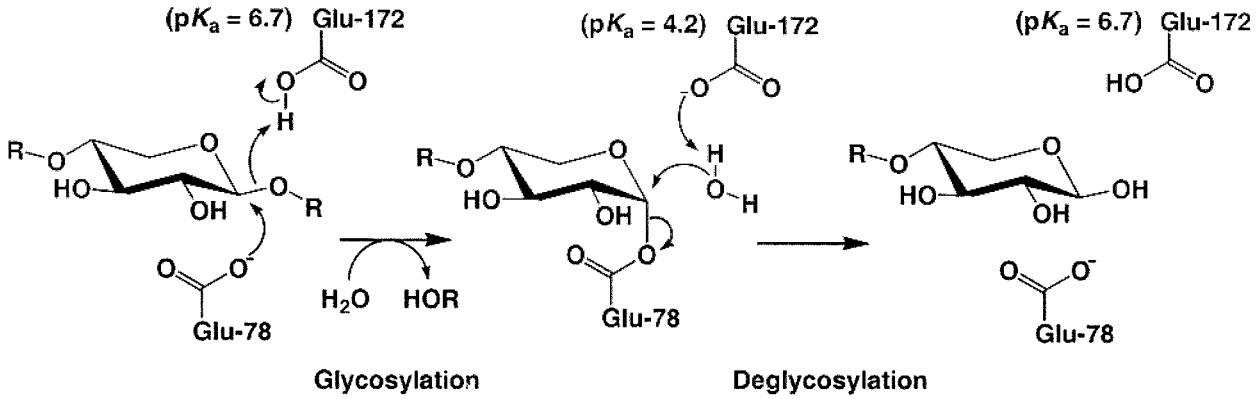


Figure 1. Double-displacement mechanism of retaining glycosidases.

high and that of Asp-70 (2.6) is unusually low (Table 2). In addition, these two carboxyls showed two-step titrations indicating a strong interaction and cooperative action between them, which is consistent with the X-ray structure showing them to be close to each other (*39*). For strong mutual charge-charge interactions, cooperative titrations are generally predicted and Oda et al. (*11*) further modeled the titrations of the Asp-10 and -70 and reproduced the two-step titrations, indicating the direct effect of the electrostatic interaction. This situation is similar to that of the two catalytic aspartic acid residues in the pepsin-like aspartic proteases and the catalytic carboxyls of Glu-35 and Asp-52 in hen egg-white lysozyme, which show smaller pK_a perturbations.

Examples of Proximal Positive Charges

One of the most "classic" examples of a catalytic group with a significantly perturbed pK_a is found in the enzyme acetoacetate decarboxylase (AAD, Table 2), which catalyzes decarboxylation of acetoacetate to yield acetone and carbon dioxide. Decarboxylation proceeds through formation of a Schiff base intermediate with the ε-amino group of Lys-116. On the basis of the measured >4-unit decreases in pK_a of both the active site Lys-116, as well as that of the phenol group of a reporter compound, Westheimer and coworkers (*40, 41*) proposed that the positive charge of the sequence-proximal lysine (Lys-115) could provide the driving force for deprotonation of Lys-116, necessary for Schiff base formation with the substrate. Westheimer's classic proposal withstood the test of experimental scrutiny when Highbarger et al. (*34*) performed pH-dependent chemical modifications of site-directed mutants of Lys-115 and Lys-116 and demonstrated that cationic Lys-115 generated the full decrease in pK_a of Lys-116.

Examples of Proximal Opposite Charges

Cysteine Proteinases. The cysteine proteinases, including papain, caricain, and ficin (Table 2), share a common active site composed of a Cys25-S^-/His159-Imidazole$(Im)^+$ ion pair (papain numbering) and catalyze proteolysis using the thiolate anion for nucleophilic attack at the peptide carbonyl group and the imidazolium cation as the general-acid catalyst for protonating the departing amine. The resulting thioacyl-enzyme intermediate is then hydrolyzed by a water molecule to regenerate the enzyme for the next catalytic cycle. Chemical modification studies indicated that the pK_a values of the nucleophilic cysteines are strongly perturbed downward to values of 3.3, 2.9, and 2.5 in papain, caricain, and ficin, respectively, and result primarily from close interaction with cationic histidine and secondarily from placement at the *N*-terminus of a α-helix (*22*). Because thioacylation is the rate-limiting step in this class of enzymes, the preexisting Cys-S^-/His-Im^+ ion pair does not provide for complete catalytic competence. In addition, there may be a requirement to rearrange the ion-pair geometry utilizing electrostatic contributions arising from substrate binding effects (*22*).

ArsC Enzyme. The ArsC enzyme (Table 2) uses the thiolate of Cys-12 to mediate reduction of arsenate [As(V)] to arsenite [As(III)] with reducing equivalents coming from either glutathione or glutaredoxin. Arsenite is the substrate of the ATP-coupled Ars pump that extrudes arsenite from cells of *E. coli*. Chemical modification and mutagenesis studies showed that the lowered pK_a value of Cys-12 of 6.3 is stabilized entirely by a positively charged residue at position 8, indicating that Cys-12 and His-8 form a thiolate-imidazolium ion pair (Table 2) (*19*). The pK_a of His-8 has not been determined.

Alanine Racemase. A two-base mechanism has been proposed for alanine racemase (Ala-Race, Table 2) in which both Lys-39 and Tyr-265 function as general acid-base catalysts to catalyze the interconversion of L- and D-alanine using pyridoxal phosphate as the cofactor. On the basis of mutagenesis and kinetic studies, Sun and Toney (*29*) proposed that the pK_a of Tyr-265 of 7.2 is decreased by 2.5 units, which would activate it as a general-base catalyst under the near neutral conditions found in the cell. From the X-ray structure of L-alanine phosphonate bound to alanine racemase (*42*), it can be seen that Tyr-265 is hydrogen bonded to His-166, which in turn is hydrogen bonded to Arg-219. Mutations of Arg-219 to the positively and like-charged lysine, the neutral-charged alanine, and the negatively and oppositely charged glutamate systematically

increased the pK_a of Tyr-265, indicating that a positive charge at this position is required to stabilize the decreased pK_a of Tyr-265 (*26*).

UDP-Galactose 4-Epimerase. UDP-galactose 4-epimerase (GalE, Table 2) catalyzes the interconversion of UDP-galactose and UDP-glucose in which tightly bound NAD^+ functions as the essential coenzyme for promoting epimerization at hexopyranosyl C4 of substrates. It has been proposed that epimerization is facilitated by nonstereospecific hydride abstraction from C4 by NAD^+, rotation of the ketohexopyranosyl intermediate about the P_β—O bond, and then nonstereospecific reduction of the 4-ketone by NADH. It is believed that hydride transfer from C4 to NAD^+ is activated by Tyr-149, which acts as a general base to abstract the C4—OH hydroxyl proton. In support of this mechanism, the measured pK_a of 6.1 for Tyr-149 is significantly decreased, and its side-chain hydroxyl group is positioned near the C4 hydroxyl group (*30*). The decreased pK_a value of 3.6 units for Tyr-149 is attributed mainly to the positive electrostatic field created by NAD^+ and Lys-153 (~3.0–3.2 units) and partly to hydrogen bonding with Ser-124 (~0.6–0.7 units).

CHARGE-DIPOLE INTERACTIONS

Hydrogen Bonding

The titratable group can also interact with partial charges or the *microscopic* dipoles found in polar residues and bound water molecules through the formation of a hydrogen bond. Hydrogen bonding involves the sharing of a hydrogen atom between two heteroatoms, usually oxygen, nitrogen, or sulfur. One heteroatom is a weakly acidic donor atom and the other is a weakly basic acceptor atom. A normal or weak hydrogen bond ranges from 2.7 to 3.0 Å in length, depending on the van der Waals contact distance between the two heteroatoms. The favorable energy gained on formation of a normal hydrogen bond (−1 to −3 kcal/mol) can be utilized to increase the pK_a of the hydrogen bond donor and decrease the pK_a of the hydrogen bond acceptor by the equivalent energetic amounts.

If the pK_a values of the hydrogen bond donor and acceptor are closely matched, then the hydrogen bond becomes unusually short (<2.6 Å), facilitating a stronger energy of formation (>7 kcal/mol) (*43*, *44*) and much larger pK_a perturbations can be observed. For example, proton "sponge" molecules contain two ionizable groups, which are sterically fixated in close proximity. To avoid electronic repulsion, the pK_a of one of the groups is significantly increased and a proton is tightly bound between the two groups (*45*). Hydrogen bonding can also indirectly affect the pK_a value of a heteroatom in a molecule far from the hydrogen bond by stabilizing a tautomeric form of the molecule (e.g., the catalytic triad of serine proteases is described next).

Serine Proteases. Serine proteases share a common active site composed of an oxyanion hole and a catalytic triad consisting of a serine, histidine, and aspartate (Fig. 2). The Nε nitrogen of His-57 in chymotrypsin functions first as a general-base catalyst, deprotonating Ser-195 for nucleophilic attack at the peptide backbone carbonyl group, which is polarized by the oxyanion hole, resulting in formation of the tetrahedral intermediate. His-57 NεH$^+$ then functions as a general-acid catalyst, protonating the departing amine of the leaving peptide. A water molecule then hydrolyzes the resulting acyl-enzyme intermediate to regenerate the free enzyme. NMR studies showed that the pK_a value for the His Nε nitrogen increases from near 7.5 in the free enzyme to pK_a values measured between 10.3 and 12.1 in complexes of chymotrypsin with various peptidyl trifluoroketones (Chy-TFKs, Table 2), analogs of the tetrahedral intermediate (*11*). Increased basicity for His Nε would enhance its reactivity in abstracting the proton from serine and thereby lower the activation barrier for formation of the tetrahedral intermediate.

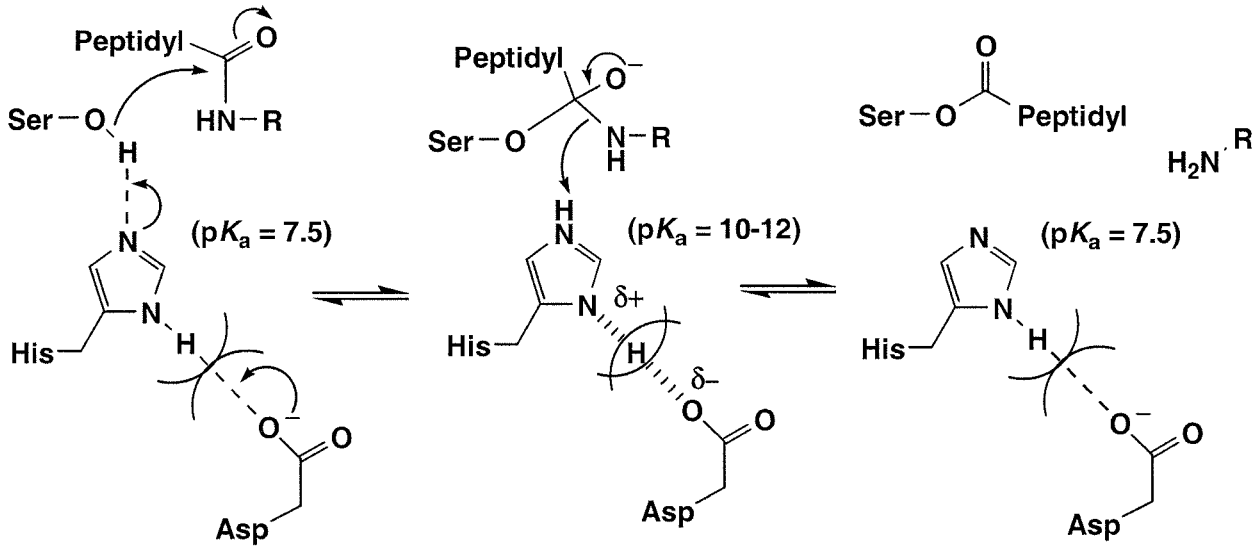


Figure 2. Mechanism of the acylation half-reaction of serine proteases.

The main mechanistic difference between the serine proteases and the cysteine proteases is the fact that the serine proteases require an additional catalytic carboxyl group to activate ionization of the nucleophilic serine. Fig. 2 shows that the carboxyl group of an aspartate is involved in a normal hydrogen bond to the His NδH and helps orient the His Nε nitrogen to activate serine for nucleophilic attack. The van der Waals contact distance between the carboxylate oxygen and the imidazole nitrogen is 2.7 Å and is indicated by the half spheres (Fig. 2). Upon binding the peptidyl group, the Asp···His hydrogen bond distance is apparently compressed to a distance shorter than the van der Waals contact, promoting formation of a shorter, stronger hydrogen bond (*46*) (Fig. 2). Although the distance between the oxygen and the nitrogen becomes shorter, the His Nδ—H bond is polarized or lengthened, increasing the negative charge on the imidazole ring, which serves to increase the basicity of the His Nε nitrogen.

Aspartate Aminotransferase. Aspartate aminotransferase (AspAT, Table 2) catalyzes reversible transamination reactions between dicarboxylic amino and keto acids. The amino group is transferred from aspartate to the coenzyme pyridoxal 5′-phosphate (PLP) to form oxalacetate and pyridoxamine 5′-phosphate (PMP). The amino group is then transferred from PMP to 2-oxoglutarate to form glutamate and regenerate PLP. The active-site lysine forms a Schiff base with the coenzyme PLP, and two arginines interact with the α- and ω-carboxylate groups of the substrates and products (Fig. 3). Kinetic and spectroscopic studies indicated that the pK_a of the imine nitrogen of the Schiff base has a decreased pK_a value of 6.9, compared to the value of around 10–12 of PLP Schiff bases in aqueous solutions. The lowered pK_a value for the Schiff base increases the fraction of unprotonated Schiff base, which facilitates the initial transaldimination reaction. On binding substrate, the pK_a increases by 2 units, and a proton is transferred from the substrate cationic amino group to the Schiff-base imine nitrogen. The substrate's neutral amino group is then activated as the nucleophile to attack the Schiff-base imine, which has increased its electrophilicity by its protonation.

Recent studies by Kagamiyama and coworkers (*31*, *47*) indicated that, contrary to the classical explanation, the electrostatic effect of the two active-site cationic arginines (Arg-292 and -386) accounts for only a partial amount (0.4 units) of the decrease in pK_a of the Schiff base in the PLP-lysine aldimine. Likewise, neutralization of these positive charges upon binding the dicarboxylate substrate accounts for only 0.7 of the 2-unit increased pK_a value. Rather, the mechanism that modulates the pK_a of the Schiff base in the PLP-lysine aldimine is attributed to changes of the imine-pyridine torsion angle (χ) (Fig. 3) (*31*). In the absence of substrate, the torsion angle around the C4—C4′ of the unprotonated form of the Schiff base is 35°. Because the imine bond is out of the plane of the pyridine ring, the pK_a of the Schiff base is lowered by 2.8 units because it cannot form a favorable hydrogen bond to the phenolic 3′-oxygen of the pyridine ring of PLP. Upon binding susbtrate, this torsion angle is reduced to 25°, favoring hydrogen bond formation to the 3-oxygen and stabilizing the protonated Schiff base by 1.3 units. The torsion angle between the Schiff base imine and the pyridine ring is maintained by a hydrogen-bond network including the guanidinium group of Arg-386, the amide group of Asn-194, and the 3′-oxygen of the pyridine ring. Thus, Arg-386 is considered

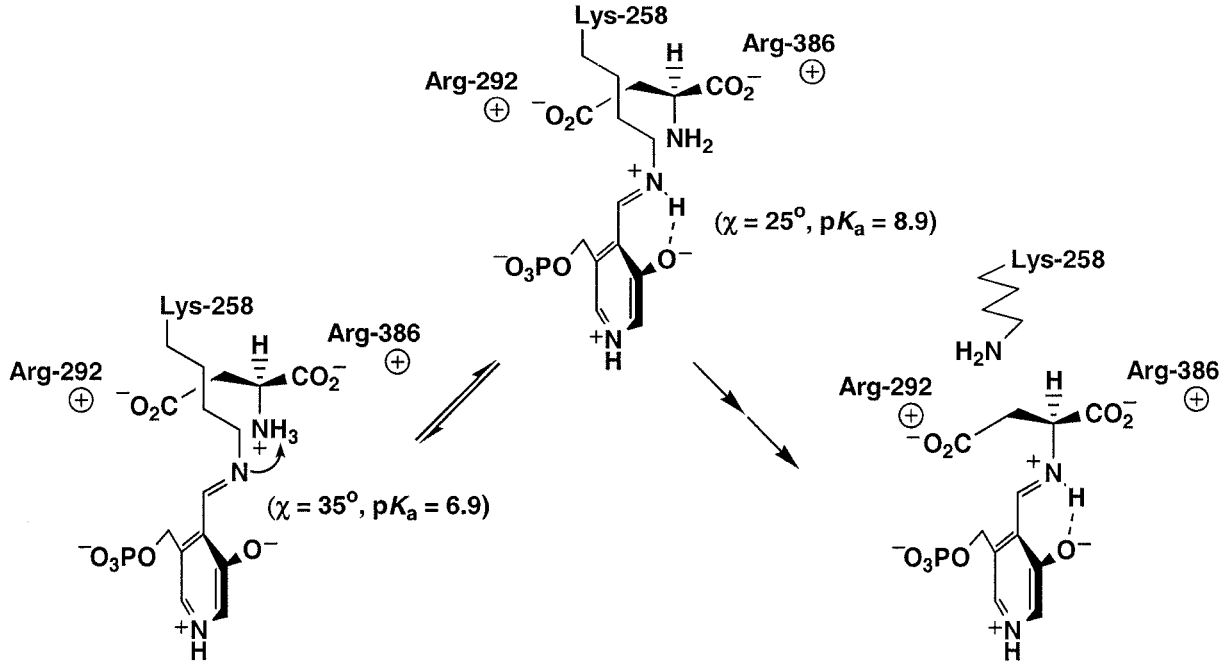


Figure 3. Reaction of pyridoxal phosphate with aspartate catalyzed by aspartate aminotransferase.

to have both electrostatic- and steric-mediated effects on the pK_a of the Schiff base. Interaction of the carboxylate group of the substrate with Arg-386 disrupts the Asn-194-mediated hydrogen-bond network, which relaxes the torsion angle.

Creatine Kinase. Human muscle creatine kinase (hmCK, Table 2) catalyzes the reversible phosphorylation of creatine, which plays a major role in energy homeostasis primarily in muscle and brain cells. To date, a conserved cysteine residue has been found to be essential for catalytic activity, but the function of this active-site residue remains contentious. Nevertheless, the pK_a of this cysteine in hmCK (Cys-282) has been measured and found to be decreased to 5.6 (*26*). The X-ray structure as well as mutagenesis studies suggest that the decreased pK_a results partially from hydrogen bonding to Ser-284, while the remaining contribution is undetermined (*26*).

N-Terminus of a α-Helix

The α-helices found in peptides and proteins have a *macroscopic* dipolar character arising from the parallel alignment of the dipolar peptide bonds of the helix (*48*). The alignment is such that the oxygens of the backbone carbonyl groups (C═O) are oriented parallel to the helix and point to the *C*-terminus of the helix, whereas the protons of the backbone amide groups (H─N) are hydrogen bonded to the backbone carbonyls and point back towards the *N*-terminus. The array of small dipoles along the helix backbone can be viewed as a purely electrostatic effect whereby a partial net charge exists at the helix termini, and nearby groups are perturbed by this field. In this sense, it has been estimated that the effect of the α-helix dipole to be that of a $+0.5$ elemental charge at the *N*-terminus plus a -0.5 elemental charge at the *C*-terminus (*49*). In addition, the two *ends* of a helix differ in their hydrogen-bonding properties. The *N*-terminal backbone amide and the *C*-terminal backbone carbonyl groups are available to function as hydrogen-bond donors and acceptors, respectively, to affect the pK_a of its hydrogen-bonded partner.

Triosephospahte Isomerase. Triosephosphate isomerase (TIM, Table 2) catalyzes the reversible tautomerization of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate, with Glu-165 removing the pro-*R* proton from C1 of dihydroxyacetone phosphate and His-95 polarizing the carbonyl group of the substrate. It was commonly believed that the pK_a of His-95 should be elevated so that it would retain positive charge at neutral pH, which would enhance its ability as an electrophilic catalyst to promote enolization of its substrate. Using NMR and ^{13}C- and ^{15}N-specific labeling, Lodi and Knowles (*50*) unambiguously demonstrated that the pK_a of His-95 is <4.5 so that it is surprisingly lowered by at least 2 units. The NMR studies further showed that the Nε nitrogen is protonated (*50*), consistent with the X-ray structure (*51*) showing the Nδ nitrogen to be hydrogen bonded to the backbone amide of Glu-97 so that the Nε─H bond vector is oriented to the substrate carbonyl.

The X-ray structure (*51*) further showed that the hydrogen bond to the amide of Glu-97 is facilitated by the placement of His-95 at the *N*-terminal end of a short α-helix (residues 95 to 102). Lodi and Knowles (*14*) investigated the energetic contributions to the lowered pK_a arising from the *N*-terminal "helix-dipole effect" of having either a partial positive charge or the possible hydrogen bond from the backbone amide by measuring the pK_a values of both His-95 and His-103 in the native and unfolded states of the enzyme. The pK_a of His-103, which is located at the *C*-terminus of the α-helix, is increased by 0.6 units upon protein folding, consistent with its placement near the partial negative charge at the *C*-terminus of the helix (*52*). Thus, the >2-unit decrease in the pK_a of His-95 must result from both the electrostatic and hydrogen bond contributions by the *N*-terminus of the α-helix. The use of *neutral* His-95 as an electrophilic catalyst can be rationalized in that the enzyme may have evolved to match the pK_a for ionization of neutral His-95 to its conjugate imidazolate anion base with the pK_a of the intermediate enediol(ate), facilitating the formation of a short, strong hydrogen bond (*43*, *44*).

Thioredoxin Family. Substantial pK_a perturbations are observed in the active sites of proteins in the thioredoxin family: thioredoxin (TRX), glutaredoxin (GRX), DsbA, DsbC, and two domains of the protein disulfide isomerase (PDI) (Table 2). All of these enzymes have similar folded conformations with active sites containing two cysteine residues, separated by two other residues in the sequence ─CXYC─, that catalyze reversible disulfide bond formation with target proteins. Although the various members of the thioredoxin family have been shown to possess very similar active-site structures, the stabilities of these disulfide bonds differ greatly. Stabilization of the reduced or thiol(ate) forms of these active site cysteines results from perturbation of the pK_a of Cys_N from its intrinsic value of 9.1 (Table 1A) to lower values ranging from a pK_a near 7.5 for thioredoxin to pK_a values of <5.5, 4.5, and 3.5 for glutaredoxin, protein disulfide isomerase, and DsbA, respectively (Table 2). Ionization of the solvent exposed Cys_N-SH to Cys_N-S^-, which can be monitored by ultraviolet spectroscopy and chemical modification, is necessary for all enzymes in this family to initiate nucleophilic attack for either subsequent reduction of its target disulfide (e.g., thioredoxin) or reoxidation of itself (e.g., glutaredoxin, protein disulfide isomerase, and DsbA). The range of decreased pK_a values of Cys_N residues in the thioredoxin family (Table 2) indicates that these enzymes provide different amounts of stabilizing interaction to the negative charge of the thiolate anion.

One mechanism of thiolate stabilization, common to all members of the thioredoxin family, involves the structural placement of Cys_N at the *N*-terminus of an α-helix (Table 2). In *E. coli* thioredoxin (bTRX, Table 2) (*15*), the electrostatic effect of having a net partial positive charge at the *N*-terminus of the helix can account for the slight stabilization of Cys_N-S^- ($\Delta pK_a = -1.6$). However, the increased amount of stabilization of Cys_N-S^- in other members of this family requires additional interactions such as hydrogen bonding or proximity to cationic residues. In human thioredoxin (hTRX, Table 2), the increased stabilization of Cys_N-S^- ($\Delta pK_a = -2.8$) appears to be provided by

the presence of a hydrogen bond with the backbone amide of Cys_C (*16*). In glutaredoxin, the increased stabilization of Cys_N-S^- ($\Delta pK_a = > -3.6$) appears to be provided by the presence of a hydrogen bond with Cys_C-SH (*17*). The assignments of these hydrogen bonds were made on the basis of NMR studies that showed strong interresidue NOEs from the respective proton donor to the CβH protons of Cys_N-S^-. In eukaryotic protein disulfide isomerase ($\Delta pK_a = -4.6$) (*18*) and *E. coli* DsbA ($\Delta pK_a = -5.6$) (*19*), the additional stabilization of Cys_N-S^- is provided by positioning it near a cationic histidine residue, which is conserved at the third position of the active site consensus sequence (—CX**H**C—).

Protein Tyrosine Phosphatases. Protein tyrosine phosphatases (PTP's, Table 2) catalyze hydrolysis of phosphotyrosines using a cysteine thiolate anion for nucleophilic attack at the phosphorus and an aspartic acid as the general-acid catalyst for protonating the departing tyrosine hydroxyl. This ionized aspartate then abstracts a proton from an attacking water molecule thereby activating hydrolysis of the thiophospho-enzyme intermediate, and the enzyme is regenerated for the next catalytic cycle. Members of the protein tyrosine phosphatase family contain a characteristic motif (H/V)CX_5R(S/T), which constitutes the phosphate binding loop (P-loop), utilizing backbone amide groups and the guanidinium group of the arginine residue to coordinate the nonbridging or equatorial oxygens of the phosphate group during substrate binding and catalysis. The pK_a values of the nucleophilic cysteine are perturbed downward to values of 4.7 in *Yersinia* PTP (yPTP, Table 2) (*23*), 5.6 in human PTP1 (hPTP1, Table 2) (*24*) and 6.7–7.5 in bovine liver low molecular weight PTP (blmPTP, Table 2) (*25*).

As noted before for the thioredoxin family, this broad range of downwardly perturbed pK_a values for cysteine results from various combinations of structural features. Slight stabilization of the thiolate anion is provided by placement of the active-site cysteine at the *N*-terminus of a α-helix in all PTPs. The hydroxyl group of the serine or threonine residue in the signature motif provides additional stabilization through hydrogen bonding. Mutagenesis studies indicate that the adjacent histidine residue in yPTP (*23*) and hPTP1 (*24*) provides the greatest and remaining thiolate stability, although it has not been firmly established whether the histidine exists in its cationic form as demonstated in the disulfide isomerases and cysteine proteinases. The valine residue found instead of the histidine in the low molecular weight PTPs likely results in the observed higher pK_a values of the nucleophile.

BORN OR DESOLVATION EFFECT

The Born or desolvation effect arises from the energetically unfavorable process of transferring a *single* charged group from an aqueous solvent of high polarity (water, $\varepsilon = 78$) to the hydrophobic interior of a protein in which the dielectric constant is estimated to be much lower. The effect of burying an ionizable group in the interior of a protein is best exemplified by the classical studies with staphylococcal nuclease, in which the buried Val-66 was replaced with either a lysine (*53*, *54*) or a glutamic acid (*55*). Equilibrium pH titrations of the V66K and V66E mutants revealed that the pK_a of Lys-66 decreased 4.9 units and that the pK_a of Glu-66 increased 4.3 units, favoring the uncharged state of protonation. According to the Born formalism (*56*, *57*), these ΔpK_as are energetically equivalent to the transfer of a charged group from water to a medium of $\varepsilon = 12$. In contrast, the static dielectric constants of dry protein powders range from $\varepsilon_{prot} = 2$–4 (*58–60*). Similar low values of ε_{prot} are predicted by a variety of theoretical calculations on the basis of normal mode analysis and on molecular dynamics simulations (*61–63*). The elevated dielectric constant of $\varepsilon_{prot} \cong 12$ can be rationalized in terms of exposure of buried ionizable groups to solvent either by global unfolding (*64*, *65*) or by the presence of buried water molecules in the native structure (*55*).

In cases where shifts in the pK_a of an ionizable group in a protein are determined primarily by the energetics of desolvation, with only minimal contributions by interactions with other charged or polar residues, the effective dielectric constant in the protein interior (ε_{prot}) can be estimated using a simple Born formalism (*56*, *57*), which relates the Gibbs free energy ($\Delta G^\circ = 2.3RT\Delta\ pK_a$) for transferring an ion of valence *Z* and cavity radius r_{cav} (Å) from water into the protein interior. Because the cavity radius differs among the various ionizable groups, a given ΔpK_a value will yield different values of ε_{prot} depending on the particular group.

Ketosteroid Isomerase. Ketosteroid isomerase (KSI, Table 2) catalyzes the conversion of Δ^5- to Δ^4-3-ketosteroids by sequential enolization and ketonization reactions. Asp-38 acts as a general base to remove a proton from C4 and subsequently return it to C6. Recent structural studies indicate that both Tyr-14 and Asp-99 act as hydrogen-bond donors to stabilize the negatively charged O3 oxygen of the dienolate intermediate (*66*). In accord with the proposed role of Asp-99 being such a hydrogen-bond donor, the pK_a of Asp-99 has been estimated by native gel electrophoresis and kinetic studies of mutants of ketosteroid isomerase to be >9 (Table 2) (*10*). Although not rigorously tested, the X-ray (*66*) and NMR structures (*67*, *68*) of ketosteroid isomerase suggest that the unusually high pK_a of Asp-99 is due to its location in a hydrophobic environment (HPE) as well as its proximity to the general base Asp-38. Consistent with a hydrophobic active site environment, the pK_a for Tyr-14 is also increased, and using the Born formalisms, yields an effective or local dielectric constant of $\varepsilon_{prot} = 18 \pm 2$, (*28*). The increased pK_a of Tyr-14 may serve to match the pK_a of the dieneolate oxygen of the reaction intermediate, facilitating formation of a short, strong hydrogen bond for stabilization of the intermediate (*69*).

4-Oxalocrotonate Tautomerase. 4-Oxalocrotonate tautomerase (4-OT, Table 2) catalyzes the isomerization of unconjugated α-keto acids such as 2-oxo-4-hexenedioate to its conjugated isomer 2-oxo-3-hexenedioate through a dienolate intermediate. The enzyme uses the *N*-terminal proline as the general base catalyst for transferring a proton between the two

different carbon atoms. Pro-1 can function as a general base under physiological conditions because of its unusually low pK_a of 6.4, which is 3 units lower than the pK_a of the model compound, proline amide (*70*). The X-ray structure of 4-OT (*71*) shows that within a sphere of 9 Å radius, Pro-1 is surrounded by hydrophobic residues, creating a site with a low dielectric constant. In addition, Arg-11 and Arg-39 are positioned 11 Å and 7 Å, respectively, from Pro-1 and are located at opposite sides of the active site, providing binding specificity to the terminal carboxyl groups of the substrate and product. Mutagenesis, NMR, and kinetic studies (*27*) indicate that the low pK_a of Pro-1 results solely from a low dielectric constant at the hydrophobic active site.

Antibody Aldolases. The catalytic lysine involved in Schiff base formation in a number of antibody aldolases (AbAld, Table 2) show downwardly perturbed pK_a values of 5.5 to 6.0 (*35*). In contrast to acetoacetate decarboxylase, X-ray structures show that the decreased pK_a values are due to placement in a hydrophobic microenvironment.

Bacteriorhodopsin. Bacteriorhodopsin (BR, Table 2) uses light energy to translocate protons from the intracellular to the extracellular side of the cell membrane. BR consists of seven α-helical transmembrane segments with short interhelical loops, and the chromaphore, retinal, is covalently bound in a Schiff base linkage to Lys-216 near the middle of the seventh helix (*72*). Absorption of a photon of light causes isomerization of retinal from the *all-trans* to the *13-cis* retinal isomer. After photoexcitation and retinal isomerization, the pK_a of the Schiff base drops from >13 to 8.3 (*32*, *33*), whereas the pK_a of Asp-85 increases from 2.2 to 6.9 (*5*). These coupled pK_a changes result in the transfer of a proton from the Schiff base to Asp-85, and a proton is released to the extracellular side of the membrane. Next, the Schiff base is reprotonated by Asp-96 when the pK_a of Asp-96, the primary proton donor to the Schiff base, drops from >12 to 7.1 (*7*). After reprotonation, retinal isomerizes back to *all-trans*, and the pK_a values of the Schiff base and Asp-96 increase back to their initially very high values and Asp-96 is reprotonated from the intracellular side of the membrane. The initial or ground state is ultimately regenerated when the pK_a of Asp-85 decreases back to its initial value, causing transfer of its proton to the proton-releasing group at the extracellular side of the membrane.

The pK_a changes that occur during the photocycle result from protein conformational changes that accompany the individual proton transfer reactions in the photorelaxation process. However, the specific structural details underlying the large changes in pK_a remain unclear. Both electron (*73*, *74*) and X-ray diffraction studies (*72*) indicate large movements of the F and G helices between the ground state and the intermediate conformational state obtained after release of the proton to the extracellular side of the membrane. It has been generally proposed that this movement serves to redistribute charges of a complex hydrogen-bonded network as well as alter solvent accessibility in order to alter the pK_a values of the key groups involved in proton transport.

Turner and collegues hypothesized that the driving force for the structural changes originate from concerted changes in a number of amino-acid side-chain interactions within the F and G helices (*75*). To identify these specific interactions, they have begun performing scanning mutagenesis of these regions in the D85N mutant of BR (*75*), which restricts the conformational equilibrium to only the reprotonation of the Schiff base by D96, because the neutral side chain of Asn-85 is effectively a "permanently" protonated Asp-85 (Fig. 4). Fig. 4A shows the results of these amino acid substitutions, which were obtained using a novel spectroscopic screening system. A light scattering integration sphere was attached to an analytical spectrophotometer enabling the ability to collect total reflectance spectra of whole cells expressing each of the second-site mutants of BR D85N. The reflectance spectra allow rapid semiquantitative comparison of BR spectra and pH-dependent spectral changes, without requiring protein purification. The amino-acid substitutions that caused significant shifts in the conformational equilibria coupled to the pK_a changes of the Schiff base and D96 are mapped onto the tertiary structure of BR (Fig. 4B). The observed patterns of perturbation indicate unique domains of peptide-peptide, peptide-lipid, and peptide-water interactions that are energetically coupled to the structural transitions that regulate the various pK_a changes. These experimental results, together with the results of various computational studies of BR, will provide the first and most detailed analysis of the structural basis of perturbed pK_a values of ionizing groups among the very large and poorly understood class of membrane-bound signaling and transport enzymes.

NUCLEIC ACIDS

The intrinsic pK_a values of the N1 nitrogens of adenylate (p$K_a = 3.5$) and guanylate (p$K_a = 9.4$) and the N3 nitrogens of cytidylate (p$K_a = 4.2$) and uridylate (p$K_a = 9.4$) (Table 1B) (*4*) make them strong candidates for use as catalytic groups in RNA and ribonucleoprotein molecules. Although the N3, N7, and exocyclic amino nitrogens of adenosine and guanosine are less likely candidates due to their very low intrinsic pK_a values (p$K_a < 2$) (*4*), it is plausible that the RNA or ribonucleoprotein can provide microenvironments that could strongly perturb these pK_a values near to physiological pH. The phosphodiester backbone, as well as the presence of bound metal ions and the aqueous-like environment found in nucleic acid structures, indicate that electrostatic effects may play a predominant role in perturbing pK_a values, as compared to perturbations of protein groups, which can further provide hydrophobic groups resulting in dielectric or Born effects.

Ribosomal Peptidyl Transferase Center (rPTC). One of the most intriguing examples is provided by X-ray structural (*76*) and chemical modification studies of the ribosomal peptidyl transferase center (rPTC, Table 2) (*36*) that suggested that the solvent-exposed N3 nitrogen of adenosine-2451 had an increased pK_a of 7.6, which could activate it as a general acid-base

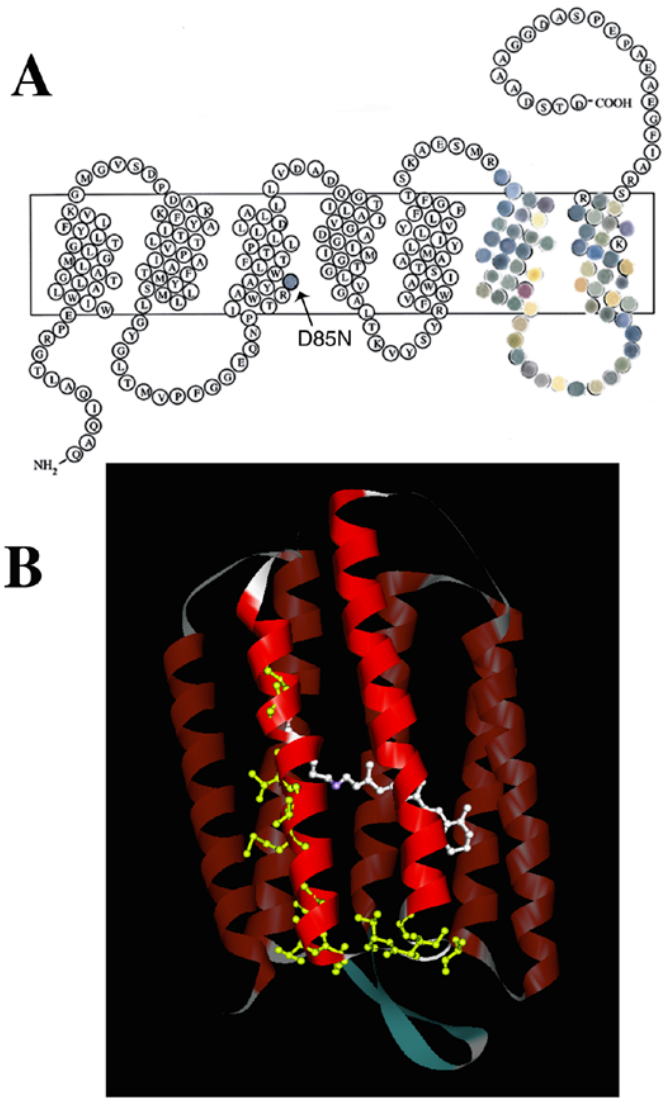


Figure 4. Scanning mutagenesis of the F and G helices of the D85N mutant of BR to reveal amino acids that regulate the coupled pK_a changes of the Schiff base and Asp-96 (*75*). The cytoplasmic and the extracellular membrane surfaces are oriented towards the top and bottom, respectively. (A) Phenotypes of whole cells expressing the various second-site mutants of the D85N mutant of BR at pH 7.0. Aliquots of whole-cell samples, expressing the different mutants, were transferred to nylon membranes and then pasted onto the corresponding amino-acid position in the two-dimensional representation of BR. The differences in color correspond to changes in the visible absorbance spectra of BR in whole cells, which reflect changes in the BR D85N structure as a result of the second-site mutation. The color and position of the D85N single mutant is indicated by the arrow in the C helix. (B) Tertiary map of BR showing the amino-acid residues (yellow) important in the regulation of the conformational equilibria, which are coupled to pK_a changes of the Schiff base and Asp-96. The chromophore, retinal, is shown in white in the interior of BR.

catalyst for peptide bond formation. However, subsequent studies by Strobel and coworkers (*77*) indicated that it was actually the buried N1 nitrogen of A2451 that was modified during the titration. In addition, the proposed role of A2451 functioning as the general acid-base catalyst has not withstood the test of mutation. Replacement of A2451 by other bases in the rPTC showed little change in catalytic activity (*78, 79*), indicating that ionization of N1 of A2451 has a less important role in catalysis than in tending to maintain conformational flexibility or an appropriate structure (*77–80*).

Hepatitis Delta Virus Antigenomic Ribozyme. A much smaller, but noteworthy increased pK_a value is observed for the N3 nitrogen of a catalytic cytidine (C76) in the *antigenomic* ribozyme (hdvAR, Table 2) (*37, 81*). The X-ray structure of the homologous *genomic* ribozyme indicates that a nonbridging phosphate oxygen of C22 could directly hydrogen bond to the exocyclic N4′-amino group of C75, shifting the equilibrium to the imino tautomer, which would serve to increase the pK_a of N3. By stabilizing the imino tautomeric form of C76, its N3 nitrogen could then be used as a general-base catalyst to accept the proton from the 2′-hydroxyl, activating it for attack of the bridging phosphate and formation of the pentacoordinated transition state. Protonated N3 could then be used as a general-acid catalyst to protonate the departing 5′-hydroxyl anion.

SUMMARY

Table 2 provides a comprehensive listing of ionizable, catalytic groups in enzyme active sites in which the pK_a value is significantly perturbed. The structural basis for each of the perturbed pK_a values has been analyzed, and the results indicate that three principle forms of molecular interactions account for the shifts in pK_a:

Charge-Charge Interactions. Cationic histidine and anionic cysteine ion pairs, as found in the cysteine proteases and protein tyrosine phosphatases, result in the lowering of the pK_a of the sulfhydryl, which activates it as a nucleophile. Pairs of proximal and catalytic lysine residues, as found in acetoacetate decarboxylase, result in the lowering of the pK_a of one of the lysines, activating it as a nucleophile. Pairs of proximal and catalytic carboxylate groups, as found in glycosidases, ribonucleases, aspartic proteases, and lysozyme result in the elevation of the pK_a of one of the carboxylates, which activates it as a hydrogen-bond donor or a general-acid catalyst.

Charge-Dipole Interactions. In addition to interactions between fully charged groups, the pK_a of a catalytic group can be strongly perturbed by interactions with partial charges such as those found in hydrogen bonds. Regulation of hydrogen bond distance and geometry provides a sensitive mechanism for regulating catalytic events that require systematic cycling of the pK_a value of a catalytic group that must perform more than one function such as the histidine in serine proteases, which serves first as a general-base catalyst and subsequently as a general-acid catalyst. Although the pK_a perturbations are considerably smaller, placement of a catalytic cysteine residue near

the *N*-terminus of a α-helix is found to be conserved throughout the thioredoxin family as well as the protein tyrosine phosphatase family.

Born or Desolvation Effect. The effect of transferring a charged group from an aqueous solvent of high polarity to the hydrophobic interior of a protein provides an alternative *macroscopic* force that strongly favors the neutral form of a titratable group. It seems to be a convenient mechanism for enzymes that bind hydrophobic substrates such as ketosteroid isomerase or for signaling and transport enzymes such as bacteriorhodopsin, which are embedded in the membrane.

Finally, it must be emphasized that the regulation of pK_a values of catalytic groups rarely results from the effect of one class of these types of interactions. Rather, a perturbed pK_a is modulated by a combination of these different interactions. Ligand binding often causes protein conformational changes, which modulate the balance between electrostatic contributions from different partially or fully charged groups as well as dielectric contributions arising from the exclusion or accessibility of water.

ACKNOWLEDGMENTS

This paper is dedicated to Professor Albert S. Mildvan, Department of Biological Chemistry, Johns Hopkins University School of Medicine. The authors thank the reviewers for insightful comments and suggestions.

REFERENCES

1. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology.* Dover Publications, Inc., New York.
2. Fersht, A. (1999) *Structure and Mechanism in Protein Science.* W. H. Freeman and Co., New York.
3. Jencks, W. P., and Regenstein, J. (1968) In *CRC Handbook of Biochemistry* (Sober, H. A., ed). pp. J150–J189. Chemical Rubber Co., Cleveland, OH.
4. Saenger, W. (1988) *Principles of Nucleic Acid Structure.* Springer-Verlag, New York.
5. Richter, H.-T., Brown, L. S., Needleman, R., and Lanyi, J. K. (1996) A linkage of the pK_a's of asp-85 and glu-204 forms part of the reprotonation switch of bacteriorhodopsin. *Biochemistry* **35**, 4054–4062.
6. Száraz, S., Oesterhelt, D., and Ormos, P. (1994) pH induced structural changes in bacteriorhodopsin studied by Fourier transform infrared spectroscopy. *Biophys. J.* **67**, 1706–1712.
7. Zscherp, C., Schlesinger, R., Tittor, J., Oesterhelt, D., and Heberle, J. (1999) In situ determination of transient pK_a changes of internal amino acids of bacteriorhodopsin by using time-resolved attenuated total reflection Fourier-transform infrared spectroscopy. *Proc. Natl. Acad. Sci. USA* **96**, 5498–5503.
8. Jeng, M.-F., and Dyson, H. J. (1996) Direct measurement of the aspartic acid 26 pK_a for reduced *Escherichia coli* thioredoxin by ^{13}C NMR. *Biochemistry* **35**, 1–6.
9. Qin, J., Clore, G. M., and Gronenborn, A. M. (1997) Ionization equilibria for side-chain carboxyl groups in oxidized and reduced human thioredoxin and in the complex with its target peptide from the transcription factor NFκB. *Biochemistry* **35**, 7–13.
10. Thornburg, L. D., Henot, F., Bash, D. P., Hawkinson, D. C., Bartel, S. D., and Pollack, R. M. (1998) Electrophilic assistance by Asp-99 of 3-oxo-Δ^5-steroid isomerase. *Biochemistry* **37**, 10499–10506.
11. Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., and Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease H1 from *Escherichia coli* determined by NMR. *Biochemistry* **33**, 5275–5284.
12. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a ^{13}C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* **35**, 9958–9966.
13. Lin, J., Cassidy, C. S., and Frey, P. A. (1998) Correlations of the basicity of His-57 with transition state analogue binding, substrate reactivity, and the strength of the low-barrier hydrogen bond in chymotrypsin. *Biochemistry* **37**, 11940–11948.
14. Lodi, P. J., and Knowles, J. R. (1993) Direct evidence for the exploitation of an α-helix in the catalytic mechanism of triosephosphate isomerase. *Biochemistry* **32**, 4338–4343.
15. Chivers, P. T., Prehoda, K. E., Volkman, B. F., Kim, B.-M., Markley, J. L., and Raines, R. T. (1997) Microscopic pK_a values of *Escherichia coli* thioredoxin. *Biochemistry* **36**, 14985–14991.
16. Forman-Kay, J. D., Clore, G. M., and Gronenborn, A. M. (1992) Relationship between electrostatics and redox function in human thioredoxin: characterization of pH titration shifts using two-dimensional homo- and heteronuclear NMR. *Biochemistry* **31**, 3442–3452.
17. Nordstrand, K., Åslund, F., Meunier, S., Holmgren, A., Otting, G., and Berndt, K. D. (1999) Direct NMR observation of the Cys-14 thiol proton of reduced *Echerichia coli* glutaredoxin-3 supports the presence of an active site thiol-thiolate hydrogen bond. *FEBS Lett.* **449**, 196–200.
18. Kortemme, T., Darby, N. J., and Creighton, T. E. (1996) Electrostatic interactions in the active site of the *N*-terminal thioredoxin-like domain of protein disulfide isomerase. *Biochemistry* **35**, 14503–14511.
19. Schirra, J. H., Renner, C., Czisch, M., Huber-Wunderlich, M., Holak, T. A., and Glockshuber, R. (1998) Structure of reduced DsbA from *Escherichia coli* in solution. *Biochemistry* **37**, 6263–6276.
20. McCarthy, A. A., Haebel, P. W., Törrönen, A., Rybin, V., Baker, E. N., and Metcalf, P. (2000) Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat. Struct. Biol.* **7**, 196–199.
21. Gladysheva, T., Liu, J., and Rosen, B. P. (1996) His-8 lowers the pK_a of the essential Cys-12 residue of the ArsC arsenate reductase of plasmid R773. *J. Biol. Chem.* **271**, 33256–33260.
22. Pinitglang, S., Watts, A. B., Patel, M., Reid, J. D., Noble, M. A., Gul, S., Bokth, A., Naeem, A., Patel, H., Thomas, E. W., Sreedharan, S. K., Verma, C., and Brockelhurst, K. (1997) A classical enzyme active center motif lacks catalytic competence until modulated electrostatically. *Biochemistry* **36**, 9968–9982.
23. Zhang, Z.-Y., and Dixon, J. E. (1993) Active site labeling of the *Yersinia* protein tyrosine phosphatase: the determination of the pK_a of the active site cysteine and the function of the conserved histidine 402. *Biochemistry* **32**, 9340–9345.
24. Lohse, D. L., Denu, J. M., Santoro, N., and Dixon, J. E. (1997) Roles of aspartic acid-181 and serine-222 in intermediate formation and hydrolysis of the mammalian protein-tyrosine-phosphatase PTP1. *Biochemistry* **36**, 4568–4575.
25. Zhang, Z.-Y., Davis, J. P., and Van Etten, R. L. (1992) Covalent modification and active site-directed inactivation of a low molecular weight phosphotyrosyl protein phosphatase. *Biochemistry* **31**, 1701–1711.
26. Wang, P.-F., McLeish, M. J., Kneen, M. M., Lee, G., and Kenyon, G. L. (2001) An unusually low pK_a for Cys-282 in the active site of human muscle creatine kinase. *Biochemistry* **40**, 11698–11705.
27. Czerwinski, R. M., Harris, T. K., Massiah, M. A., Mildvan, A. S., and Whitman, C. P. (2001) The structural basis for the perturbed pK_a of the catalytic base in 4-oxalocrotonate tautomerase: kinetic and structural effects of mutations of Phe-50. *Biochemistry* **40**, 1984–1995.
28. Li, Y.-K., Kuliopulos, A., Mildvan, A. S., and Talalay, P. (1993) Environments and mechanistic roles of the tyrosine residues of Δ^5-3-ketosteroid isomerase. *Biochemistry* **32**, 1816–1824.

29. Sun, S., and Toney, M. D. (1999) Evidence for a two-base mechanism involving tyrosine-265 from arginine-219 mutants of alanine racemase. *Biochemistry* **38**, 4058–4065.
30. Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., Holden, H. M., and Frey, P. A. (1997) Mechanistic roles of tyrosine 149 and serine 142 in UDP-galactose 4-epimerase from *E. coli*. *Biochemistry* **36**, 10675–10684.
31. Mizuguchi, H., Hayashi, H., Okada, K., Miyahara, I., Hirotsu, K., and Kagamiyama, H. (2001) Strain is more important than electrostatic interaction in controlling the pK_a of the catalytic group in aspartate aminotransferase. *Biochemistry* **40**, 353–360.
32. Druckmann, S., Ottolenghi, M., Pande, A., Pande, J., and Callender, R. H. (1982) Acid-base equilibrium of the Schiff base in bacteriorhodopsin. *Biochemistry* **21**, 4953–4959.
33. Brown, L. S., and Lanyi, J. K. (1996) Determination of the transiently lowered pK_a of the retinal Schiff base during the photocycle of bacteriorhodopsin. *Proc. Natl. Acad. Sci. USA* **93**, 1731–1734.
34. Highbarger, L. A., and Gerlt, J. A. (1996) Mechanism of the reaction catalyzed by acetoacetate decarboxylase. Importance of lysine 116 in determining the pK_a of active-site lysine 115. *Biochemistry* **35**, 41–46.
35. Barbas, III, C. F., Heine, A., Zhong, G., Hoffman, T., Gramatikova, S., Bjornestedt, R., List, B., Anderson, J., Stura, E. A., Wilson, I. A., Lerner, R. A. (1997) Immune versus natural selection: antibody aldolases with enzymic rates but broader scope. *Science* **278**, 2085–2092.
36. Muth, G. W., Ortoleva-Donnelly, L., and Strobel, S. A. (2000) A single adenosine with a neutral pK_a in the ribosomal peptidyl transferase center. *Science* **289**, 947–950.
37. Perrotta, A. T., Shih, I.-H., and Been, M. D. (1999) Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme. *Science* **286**, 123–126.
38. Tanford, C. (1961) *Physical Chemistry of Macromolecules.* John Wiley and Sons, Inc., New York.
39. Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Makamura, H., Ikehara, M., Matsuzaki, T., and Morikawa, K. (1992) Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution. *J. Mol. Biol.* **223**, 1029–1052.
40. Kokesh, F. C., and Westheimer, F. H. (1971) A reporter group at the active site of acetoacetate decarboxylase. II. Ionization constant of the amino group. *J. Am. Chem. Soc.* **93**, 7270–7274.
41. Frey, P. A., Kokesh, F. C., and Westheimer, F. H. (1971) A reporter group at the active site of acetoacetate decarboxylase. I. Ionization constant of the nitrophenol. *J. Am. Chem. Soc.* **93**, 7266–7269.
42. Stamper, C. G. F., Morollo, A. A., and Ringe, D. (1998) Reaction of alanine racemase with 1-aminoethylphosphonic acid forms a stable external aldimine. *Biochemistry* **37**, 10438–10445.
43. Cleland, W. W., and Kreevoy, M. M. (1994) Low-barrier hydrogen bonds and enzymatic catalysis. *Science* **264**, 1887–1890.
44. Harris, T. K., and Mildvan, A. S. (1999) High-precision measurement of hydrogen bond lengths in proteins by nuclear magnetic resonance methods. *Proteins* **35**, 275–282.
45. Frey, P. A., and Cleland, W. W. (1998) Are there strong hydrogen bonds in aqueous solutions? *Bioorg. Chem.* **26**, 175–192.
46. Cassidy, C. S., Lin, J., and Frey, P. A. (1997) A new concept for the mechanism of action of chymotrypsin: the role of the low-barrier hydrogen bond. *Biochemistry* **36**, 4576–4558.
47. Hayashi, H., Mizuguchi, H., and Kagamiyama, H. (1998) The imine-pyridine torsion of the pyridoxal 5′-phosphate Schiff base of aspartate aminotransferase lowers its pK_a in the unliganded enzyme and is crucial for the successive increase in the pK_a during catalysis. *Biochemistry* **37**, 15076–15085.
48. Wada, A. (1976) The α-helix as an electric macro-dipole. *Adv. Biophys.* **9**, 1–63.
49. Hol, W. G. J., van Duijnen, P. T., and Berendsen, H. J. C. (1978) The α-helix dipole and the properties of proteins. *Nature (London)* **273**, 443–446.
50. Lodi, P. J., and Knowles, J. R. (1991) Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. *Biochemistry* **30**, 6948–6956.
51. Davenport, R. C., Bash, P. A., Seaton, B. A., Karplus, M., Petsko, G. A., and Ringe, D. (1991) Structure of the triosephosphate isomerase-phosphoglycolohydroxamate complex: an analog of the intermediate on the reaction pathway. *Biochemistry* **30**, 5821–5826.
52. Sancho, J., Serrano, L., and Fersht, A. R. (1992) Histidine residues at the *N*- and *C*-termini of α-helices: perturbed pK_as and protein stability. *Biochemistry* **31**, 2253–2258.
53. Stites, W. E., Gittis, A. G., Lattman, E. E., and Shortle, D. (1991) In a staphylococcal nuclease mutant the side-chain of a lysine replacing valine 66 is fully buried in the hydrophobic core. *J. Mol. Biol.* **221**, 7–14.
54. Garciá-Moreno, E. B., Dwyer, J. J., Gittis, A. G., Lattman, E. E., Spencer, D. S., and Stites, W. E. (1997) Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophys. Chem.* **64**, 211–224.
55. Dwyer, J. J., Gittis, A. G., Karp, D. A., Lattman, E. E., Spencer, D. S., Stites, W. E., and Garciá-Moreno, E. B. (2000) High apparent dielectric constants in the interior of a protein reflect water penetration. *Biophys. J.* **79**, 1610–1620.
56. Rashin, V., and Honig, B. (1985) Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **89**, 5588–5593.
57. Friedman, H. L., and Krishan, C. V. (1973) In *Water. A Comprehensive Treatise* (Franks, F., ed). Vol. 3, pp. 1–118. Plenum Press, New York.
58. Bone, S., and Pethig, R. (1982) Dielectric studies of the binding of water to lysozyme. *J. Mol. Biol.* **57**, 571–575.
59. Bone, S., and Pethig, R. (1985) Dielectric studies of protein hydration and hydration-induced flexibility. *J. Mol. Biol.* **181**, 323–326.
60. Harvey, S. C., and Hoekstra, P. (1972) Dielectric relaxation spectra of water absorbed on lysozyme. *J. Chem. Phys.* **76**, 2987–2994.
61. Gilson, M. K., and Honig, B. H. (1986) The dielectric constant of a folded protein. *Biopolymers* **25**, 2097–2119.
62. Löffler, G., Schreiber, H., and Steinhauser, O. (1997) Calculation of the dielectric properties of a protein and its solvent: theory and a case study. *J. Mol. Biol.* **270**, 520–534.
63. Simonson, T., and Perahia, D. (1995) Internal and interfacial dielectric properties of cytochrome *c* from molecular dynamics in aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1082–1086.
64. Warshel, A. (1981) Calculations of enzyme reactions: calculations of pK_a, proton transfer reactions, and general acid catalysis reactions in enzymes. *Biochemistry* **20**, 3167–3177.
65. Warshel, A., Russell, S. T., and Churg, A. K. (1984) Macroscopic models for studies of electrostatic interactions in proteins: limitations and applicability. *Proc. Natl. Acad. Sci. USA* **81**, 4785–4789.
66. Kim, S. W., Cha, S.-S., Cho, H.-S., Kim, J.-S., Ha, N.-C., Cho, M.-J., Joo, S., Kim, K. K., Choi, K, Y., and Oh, B.-H. (1997) High-resolution crystal structures of Δ^5-3-ketosteroid isomerase with and without a reaction intermediate analogue. *Biochemistry* **36**, 14030–14036.
67. Wu, Z. R., Ebrahimian, S., Zawrotny, M. E., Thornburg, L. D., Perez-Alvarado, G. C., Brothers, P., Pollack, R. M., and Summers, M. F. (1997) Solution structure of 3-oxo-Δ^5-steroid isomerase. *Science* **276**, 415–418.
68. Massiah, M. A., Abeygunawardana, C., Gittis, A. G., and Mildvan, A. S. (1998) Solution structure of Δ^5-3-ketosteroid isomerase complexed with the steroid 19-nortestosterone hemisuccinate. *Biochemistry* **37**, 14701–14712.
69. Zhao, Q., Abeygunawardana, C., Talalay, P., and Mildvan, A. S. (1996) NMR evidence for the participation of a low-barrier hydrogen bond in the mechanism of Δ^5-3-ketosteroid isomerase. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8220–8224.
70. Stivers, J. T., Abeygunawardana, C., Mildvan, A. S., Hajipou, G., and Whitman, C. P. (1996) 4-Oxalocrotonate tautomerase: pH dependence of catalysis and pK_a values of active site residues. *Biochemistry* **35**, 814–823.
71. Taylor, A. B., Czerwinski, R. M., Johnson, W. H., Jr., Whitman, C. P., and Hackert, M. L. (1998) Crystal structure of 4-oxalocrotonate tautomerase inactivated by 2-oxo-3-pentynoate at 2.4 Å resolution: analysis and implications for the mechanism of inactivation and catalysis. *Biochemistry* **37**, 14692–14700.

72. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899–911.
73. Subramaniam, S., Gerstein, M., Oesterhelt, D., and Henderson, R. (1993) Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin. *EMBO J.* **12**, 1–8.
74. Ludlam, C. F., Sonar, S., Lee, C. P., Coleman, M., Herzfeld, J., Raj Bhandary, U. L., and Rothschild, K. J. (1995) Site-directed isotope labeling and ATR-FTIR difference spectroscopy of bacteriorhodopsin: the peptide carbonyl group of tyr-185 is structurally active during the bR→N transition. *Biochemistry* **34**, 2–6.
75. Martinez, L., Thurmond, R., Jones, P., and Turner, G. J. (2002) Subdomains in the F and G helices of bacteriorhodopsin regulate the conformational transitions of the reprotonation mechanism. *Proteins* (in press).
76. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930.
77. Muth, G. W., Chen, L., Kosek, A. B., and Strobel, S. A. (2001) pH-dependent conformational flexibility within the ribosomal peptidyl transferase center. *RNA* **7**, 1403–1415.
78. Thompson, J., Kim, D. F., O'Conner, M., Lieberman, K. R., Bayfield, M. A., Gregory, S. T., Green, R., Noller, H. F., and Dahlberg, A. E. (2001) Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50 S ribosomal subunit. *Proc. Natl. Acad. Sci. USA* **98**, 9002–9007.
79. Polacek, N., Gaynor, M., Yassin, A., and Mankin, A. S. (2001) Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature* **411**, 498–501.
80. Xiong, L., Polacek, N., Sander, P., Bottger, E. C., and Mankin, A. (2001) pK_a of adenine 2451 in the ribosomal peptidyl transferase center remains elusive. *RNA* **7**, 1365–1369.
81. Shih, I.-H., and Been, M. D. (2001) Involvement of a cytosine side chain in proton transfer in the rate-determining step of ribozyme self-cleavage. *Proc. Natl. Acad. Sci. USA* **98**, 1489–1494.