Hydrogen Bonds: Simple after All?

Daniel Herschlag*†‡§∥ and Margaux M. Pinney†¶

†Department of Biochemistry, ‡Department of Chemistry, §Department of Chemical Engineering, and ¶Stanford ChEM-H, Stanford University, Stanford, California 94305, United States

ABSTRACT: Hydrogen bonds play integral roles in biological structure, function, and conformational dynamics and are fundamental to life as it has evolved on Earth. However, our understanding of these fundamental and ubiquitous interactions has seemed fractured and incomplete, and it has been difficult to extract generalities and principles about hydrogen bonds despite thousands of papers published on this topic, perhaps in part because of the expanse of this subject and the density of studies. Fortunately, recent hydrogen bond proposals, discussions, and debates have stimulated new tests and models and have led to a remarkably simple picture of the structure of hydrogen bonds. This knowledge also provides clarity concerning hydrogen bond energetics, limiting and simplifying the factors that need be considered. Herein we recount the advances that have led to this simpler view of hydrogen bond structure, dynamics, and energetics. A quantitative predictive model for hydrogen bond length can now be broadly and deeply applied to evaluate current proposals and to uncover structural features of proteins, their conformational restraints, and their correlated motions. In contrast, a quantitative energetic description of molecular recognition and catalysis by proteins remains an important ongoing challenge, although our improved understanding of hydrogen bonds may aid in testing predictions from current and future models. We close by codifying our current state of understanding into five “Rules for Hydrogen Bonding” that may provide a foundation for understanding and teaching about these vital interactions and for building toward a deeper understanding of hydrogen bond energetics.

Hydrogen bonds lie at the heart of biology, as anticipated by Linus Pauling when, in 1931, he coined the term “hydrogen bond” and said the following:

*It has been recognized that hydrogen bonds restrain protein molecules to their native configurations, and I believe that as the methods of structural chemistry are further applied to physiological problems it will be found that the significance of the hydrogen bond for physiology is greater than that of any other single structural feature.*

Nevertheless, several decades later in 1973, in the wake of thousands of papers published on hydrogen bonds, Hopfinger concluded that, "The one definite fact about hydrogen bonds is that there does not appear to be any definite rules which govern their geometry.”

Two decades subsequent to this despondent claim, hydrogen bonds were additionally complicated by proposals that short, strong, or low-barrier hydrogen bonds could be responsible for much of catalysis by enzymes, while also drawing new interest to these old interactions.1–6

As is often the case, creative new ideas sparked discussion at multiple levels, comprising rounds of attacks, clarifications, and defenses.7–17 Although debates can distract from simpler well-established concepts, making it difficult for students and non-experts to navigate the literature, they also catalyze new thinking that can lead to new perspectives. Through this process, previously disconnected information is often integrated, initial proposals are shaped into more clearly defined models, and concrete tests are devised for these new models. Indeed, this has come to pass for hydrogen bonding, stimulated by the literature proposals and debates and building on data from decades of intensive research. Here we describe the understanding that has emerged: a remarkably simple, highly predictive model for hydrogen bond structure that will be of great value to a broad swath of biochemists, chemists, and biologists. Unsurprisingly, there remain important challenges in energetics that can now be more clearly delineated.

A SIMPLE VIEW OF HYDROGEN BOND STRUCTURE

Hydrogen bonds are interactions between the hydrogen atom covalently bound to an electronegative donor and the lone pair of electrons of an acceptor.18 While hydrogen bonds are often thought to be primarily or solely electrostatic in nature, surveys of hydrogen-bonded complexes show many hydrogen bonds being shorter than the sum of van der Waals radii for the donor and acceptor atoms and show angular constraints that are greater than expected for purely electrostatic interactions, indicating orbital overlap and partial covalent character.18

Special Issue: Current Topics in Mechanistic Enzymology

Received: February 21, 2018
Revised: April 9, 2018
Published: April 20, 2018
As differences in hydrogen bond length have been assumed to reflect differences in hydrogen bond “strength”, the shortest hydrogen bonds have been thought to be the strongest.\textsuperscript{3–6} We refer to “strength” in quotation because it is imprecise and requires further definition.\textsuperscript{b} In this section, we address the structure, or geometry, of hydrogen-bonded complexes, focusing primarily on length, as it is most readily accessible experimentally and as additional work is needed to codify our knowledge and understanding of hydrogen bond angles.\textsuperscript{18} In the following section we turn to energetics.

**Hydrogen Bond Donor/Acceptor Pair Properties as the Primary Determinants of Hydrogen Bond Length:** $\Delta pK_a$. Much of science entails searching for trends and trying to understand their origins. It was widely recognized that hydrogen bonds comprised of donor/acceptor pairs with matched, or nearly matched, $pK_a$ values gave the shortest hydrogen bonds, with O-O distances as short as $\sim 2.4$ Å, and that hydrogen bond lengths generally shorten as $\Delta pK_a$ increases in the donor and acceptor $pK_a$ values, decreases (Figure 1A).\textsuperscript{3,5,19–25} Building from these observations, and following the analyses of Gilli and Gilli,\textsuperscript{24} Sigala et al. plotted lengths of hydrogen bonds in organic, aprotic solvents (Figure 1B).\textsuperscript{26} To ensure a constant and well-established scale for comparison, 28 it was assumed that hydrogen bonds generally shorten as $\Delta pK_a$ increases, or nearly matched $pK_a$ fractions strongly suggestive of short, low-$\Delta pK_a$ hydrogen bonds in organic, aprotic solvents (Figure 2A).\textsuperscript{3,25,39,40} These hydrogen bond lengths were assumed to be long in water despite a lack of experimental evidence. Additionally, it was noted that HOH-OH\textsubscript{2} ($\Delta pK_a = 17.4$) and other high-$\Delta pK_a$ hydrogen bonds are long in water.\textsuperscript{7,39,40} However, these lengths were being compared to those of low-$\Delta pK_a$ hydrogen bonds. If, as shown from the linear relationship in Figure 1B, hydrogen bond lengths are sensitive to the nature of the donor and acceptor pair, then we cannot draw dependence of length on $\Delta pK_a$ over many orders of magnitude of proton affinity (slope = 0.020 Å/$pK_a$, $\Delta pK_a$ range of 0–20).

This relationship suggested a test for the generally held model that hydrogen bond lengths are intimately tied to the environment, i.e., the solvent in which the hydrogen bond is formed. This presumed relationship was at the heart of proposals for hydrogen bond energetics, which posited the following. (1) A non-aqueous active site or nonpolar solvent is required for short hydrogen bonds to form. (2) Shorter hydrogen bonds are stronger hydrogen bonds. Therefore, (3) hydrogen bonds become stronger upon removal from water (e.g., refs 5, 20, and 33–38 and Text S1). Nevertheless, there was no evidence for or against the first postulate, and below we describe a test that disproves this presumed relationship between hydrogen bond length and environment.

Short hydrogen bonds had been directly observed in crystals, as noted above, and there is extensive evidence from nuclear magnetic resonance (NMR), infrared (IR), and isotopic fractionation factors strongly suggestive of short, low-$\Delta pK_a$ hydrogen bonds in organic, aprotic solvents (Figure 2A).\textsuperscript{3,25,39,40} These hydrogen bond lengths were assumed to

![Figure 1. O-O hydrogen bonds follow a linear relationship with donor and acceptor $\Delta pK_a$ ($\Delta pK_a = pK_a^{\text{Donor}} - pK_a^{\text{Acceptor}}$). (A) As the donor/acceptor $\Delta pK_a$ increases from 0.0 (green) to 8.2 (red), the O-O hydrogen bond distance also increases (see footnote c). Lengths are reported from small molecule neutron diffraction structures (CSD IDs of NAHMAL01 and SUCACBO3 for green and red, respectively). (B) O-O hydrogen bonds in small molecule neutron diffraction structures. Reproduced from ref 26. Copyright 2015 American Chemical Society.](image1)

![Figure 2. $^1$H chemical shifts have a strong inverse correlation with hydrogen bond O-O distances. (A) As the hydrogen bond shortens, the $^1$H chemical shift of the hydrogen-bonded proton increases. This increase in chemical shift is due to deshielding of the hydrogen-bonded proton, which arises from the lengthening of the covalent O-H bond that accompanies shorter O-O ($r_{OH}$) and H-O ($r_{O-H}$) distances, as has been widely discussed.\textsuperscript{7,69,115,116} (B) Correlation of O-O hydrogen bond distances from X-ray crystal structures of diverse small molecules with $^1$H chemical shifts from solid-state NMR of the same compound. Panel B is reproduced from ref 117. Copyright 1999 John Wiley & Sons, Inc. Data are from ref 118. The empirical fit of the data in panel B (black line) has been used extensively in the literature to estimate hydrogen bond lengths from chemical shift data.\textsuperscript{117–125}](image2)
conclusions about environmental effects from comparisons involving hydrogen bond pairs with dissimilar ∆pK_a values. Indeed, the average HOH·OH_2 hydrogen bond length from >2000 hydrated small molecule crystal structures reasonably follows the length versus ∆pK_a correlation for hydrogen bonds in crystals (Figure S1).25 The common assumption that short hydrogen bonds are not formed in water and the difficulty in observing hydrogen bonds by ^1^H NMR in water due to exchange likely discouraged a search for short hydrogen bonds in water. Nevertheless, there were hints of short hydrogen bonds in mixed organic/aqueous solvents at low temperatures.41−44 For example, Mock observed highly deshielded proton chemical shifts indicative of short hydrogen bonds for a series of salicylates in a 10% water/90% acetone mixture at −50 °C (Figure S2).42 These observations suggested that hydrogen bond lengths might be able to be determined in water, without cosolvents and without extreme cooling, using high-field NMR and intramolecular hydrogen bonds to weaken exchange-induced line broadening. Observation of downfield ^1^H chemical shifts would provide evidence of such hydrogen bonds in water and allow their lengths to be inferred (Figure 2A,B), although the absence of these peaks would not be definitive as it could reflect longer hydrogen bonds or rapid solvent exchange.

Sigala et al. examined a series of salicylates, bearing different substituents to give a range of donor and acceptor ∆pK_a values (Figure 3A), and obtained NMR spectra in chloroform, acetone, and water (Figure 3B). Downfield chemical shifts were observed in water, indicative of short hydrogen bonds (Figure 3B), and the ^1^H chemical shifts in water were essentially identical to those in the organic solvents (Figure 3C), as were effects from ^2^H substitution on the chemical shifts, which provides information about the hydrogen bond potential.26 To assess whether the tight intramolecular nature of the salicylate hydrogen bonds might yield idiosyncratic results, several additional analyses were performed. (1) Hydrogen bond lengths from X-ray diffraction structures of substituted salicylates were shown to have the same linear

Figure 3. Substituted salicylates have the same hydrogen bond lengths in nonpolar solvents and water.26 (A) Substituted salicylates used by Sigala et al. and their donor and acceptor ∆pK_a values. (B) ^1^H NMR spectra of substituted salicylates in chloroform, acetone, and water at 4 °C. (C) ^1^H chemical shifts and estimated hydrogen bond lengths vs ∆pK_a (slopes = 0.8−1.0 ppm/pK_a unit; R^2 = 0.82−0.92 and 0.027−0.032 Å/pK_a unit; R^2 = 0.82−0.93). (D) ^1^H NMR spectra for the hydrogen-bonded proton of 2-hydroxyphenylacetate in chloroform and in a 10% water/90% acetone mixture. (E) One-dimensional potential energy curve for displacement of the hydrogen-bonded proton between the donor and acceptor oxygen atoms of the water–hydroxide dimer. Calculations were performed at the B3LYP level using the 6-311++G(df,p) basis set. Panels B−E are reproduced from ref 26. Copyright 2015 American Chemical Society.
dependence on $\Delta pK_a$ as observed for intermolecular hydrogen bonds from the neutron structures presented above [0.017 and 0.019 A/p$K_a$ for salicylates and intermolecular hydrogen bonds, respectively] (see also Figure S3 of ref 26). (2) Intermolecular hydrogen-bonded complexes typically exchange too fast in water to observe. However, 2-hydroxypyphenylacetate, which has an additional rotatable bond relative to the salicylates did (Figure 3D). (3) High-level quantum mechanical (QM) calculations gave short hydrogen bond lengths for both intra- and intermolecular complexes and very little sensitivity to dielectric over the range of 5–80 (Figure 3E). QM calculations and additional experimental observations also suggest that the shape of the hydrogen bond potential is nearly unperturbed across solvents and dielectrics. (4) The hydrogen bond lengths for substituted salicylates in chloroform, acetone, and water fall on the length versus $\Delta pK_a$ correlation line for hydrogen bonds in crystals determined by neutron diffraction (Figure S2).

To recap, there appears to be a universal relationship between hydrogen bond length and the difference in proton affinity of the hydrogen bond donor and acceptor ($\Delta pK_a$), and there appear to be two profound implications from this observation: (1) elimination of a foundational principle for hydrogen bond energetic proposals that rely on the postulate that hydrogen bonds shorten upon removal from water and (2) determination that a relationship can be used to predict hydrogen bond lengths. Given the potential importance of these statements, we address the following questions: can this relationship be extended more broadly to all known solution-phase O–O hydrogen bonds; can it be extended to N–O and N–N hydrogen bonds; and, critically for understanding protein and enzyme structure and function, does it hold for hydrogen bonds within proteins?

Figure 4A shows the hydrogen bond length versus $\Delta pK_a$ relationship from above extended to include available solution O–O hydrogen bond lengths estimated from $^1H$ NMR chemical shifts (Table S3). The overall strong correlation remains, with an $r^2$ of 0.83. Figure 4B shows the analogous plot for N–O hydrogen bonds (Tables S1 and S4). These hydrogen bonds follow a single correlation regardless of whether the proton is covalently attached to the nitrogen or oxygen, with $r^2$ values of 0.74 and 0.52 for neutron diffraction and NMR measurements, respectively. The O–O and N–O correlation lines for neutron diffraction measurements have the same dependence on $\Delta pK_a$, with slopes of 0.020 and 0.021 Å/Δ$pK_a$ respectively (Table 1). The N–O hydrogen bond lengths from neutron diffraction structures are uniformly displaced from the O–O hydrogen bonds by 0.146 Å, similar to the difference in van der Waals radius between oxygen and nitrogen (1.66 Å − 1.50 Å = 0.16 Å) and suggesting a steric origin for this difference. The correlation for N–N hydrogen bonds from neutron diffraction has a similar slope (0.015 Å/p$K_a$ (Table 1)) and is displaced

![Figure 4](image-url)

**Figure 4.** Hydrogen bond distances derived from $^1H$ NMR of small molecules follow a linear relationship with donor/acceptor $\Delta pK_a$ for (A) O–O, (B) N–O, and (C) N–N hydrogen bonds. Hydrogen bond lengths were estimated from $^1H$ chemical shifts of the hydrogen-bonded proton using the empirical correlation function from Figure 2A for O–O hydrogen bonds. N–O and N–N hydrogen bonds from small molecule crystal structures in the Cambridge Structural Database are, on average, 0.13 and 0.30 Å longer than O–O hydrogen bonds, respectively, reflecting the larger van der Waals radius of nitrogen (S. Alvarez, personal communication; see also footnote g). To account for this difference, N–O and N–N hydrogen bond distances predicted from $^1H$ NMR chemical shifts were uniformly corrected by the factors mentioned above, consistent with a previous analysis from ref 126 of N–O hydrogen bonds (see also Text S3 in ref 50). This correction does not influence the scale of the length versus $\Delta pK_a$ relationship and thus does not change the conclusions herein. The origin of the steeper relationship between O–O hydrogen bond lengths and $\Delta pK_a$ from $^1H$ NMR remains to be determined. One factor that may contribute to this difference is potential inaccuracies in converting $^1H$ NMR chemical shifts to hydrogen bond lengths due to the scatter in the empirical relationship in Figure 2B. Values are reported in Tables S1–S5 and Table S1 of ref 26. Slopes are reported in Table 1.

### Table 1. Dependencies of Hydrogen Bond Lengths on Donor and Acceptor $\Delta pK_a$ Values

<table>
<thead>
<tr>
<th>Hydrogen bond</th>
<th>Method</th>
<th>System</th>
<th>Count</th>
<th>Slope (Å/p$K_a$)</th>
<th>$r^2$</th>
<th>RMSD (Å)</th>
<th>Figure</th>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>O–O</td>
<td>Neutron diffraction</td>
<td>Small molecules</td>
<td>68</td>
<td>0.020</td>
<td>0.86</td>
<td>0.052</td>
<td>1B, 4A, 6A</td>
<td>S1 of ref 26</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR</td>
<td>Small molecules</td>
<td>291</td>
<td>0.033</td>
<td>0.82</td>
<td>0.038</td>
<td>4A</td>
<td>S3</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR and ≤1.0 Å X-ray diffraction</td>
<td>Proteins</td>
<td>59</td>
<td>0.020</td>
<td>0.77</td>
<td>0.052</td>
<td>5A</td>
<td>S6, S7</td>
</tr>
<tr>
<td>N–O</td>
<td>Neutron diffraction</td>
<td>Small molecules</td>
<td>86</td>
<td>0.021</td>
<td>0.74</td>
<td>0.070</td>
<td>4B, 6B</td>
<td>S1</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR</td>
<td>Small molecules</td>
<td>182</td>
<td>0.024</td>
<td>0.52</td>
<td>0.039</td>
<td>4B</td>
<td>S4</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR and ≤1.0 Å X-ray diffraction</td>
<td>Proteins</td>
<td>50</td>
<td>0.017</td>
<td>0.59</td>
<td>0.076</td>
<td>5B</td>
<td>S6, S7</td>
</tr>
<tr>
<td>N–N</td>
<td>Neutron diffraction</td>
<td>Small molecules</td>
<td>29</td>
<td>0.015</td>
<td>0.69</td>
<td>0.093</td>
<td>4C, 6C</td>
<td>S2</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR</td>
<td>Small molecules</td>
<td>18</td>
<td>0.012</td>
<td>0.92</td>
<td>0.020</td>
<td>4C</td>
<td>S5</td>
</tr>
<tr>
<td></td>
<td>$^1H$ NMR and ≤1.0 Å X-ray diffraction</td>
<td>Proteins</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>5C</td>
<td>S6, S7</td>
</tr>
</tbody>
</table>
from the N-O hydrogen by an additional 0.104 Å (Figure 4C and Tables S2 and S5).

We next turn to protein hydrogen bonds. In proteins, the relationship between hydrogen bond length and $\Delta pK_a$ gives slopes highly similar to those for small molecules in crystals and in solution. For example, in the active site of the enzyme ketosteroid isomerase (KSI), a series of bound substituted phenolate ligands of varying $pK_a$, which accept two hydrogen bonds from the KSI oxyanion hole hydrogen bond donors, gave slopes of 0.023 and 0.017 Å/$\Delta pK_a$.46 A similar behavior was observed in the photoactive yellow protein (PYP) binding site, with slopes of 0.030 Å/$\Delta pK_a$ for both hydrogen oxyanion hole hydrogen bond donors.47 There are also anecdotal reports of short hydrogen bond lengths in proteins from $^1$H NMR and ultra-high-resolution X-ray crystallography, and we have assembled protein hydrogen bond distances determined by these methods (Table 1 and Tables S6 and S7). Constructing plots of protein hydrogen bond length versus $\Delta pK_a$ as per analysis of the small molecule data sets, gives a correlation line with the same slope observed for small molecules: 0.020 and 0.017 Å/$\Delta pK_a$ for O-O and N-O hydrogen bonds, respectively (Figure 5A,B; we found far fewer examples of N-N hydrogen bonds in proteins and were therefore unable to determine the length vs $\Delta pK_a$ relationship). Nevertheless, outliers exist for both protein and small molecule hydrogen bond distances, and there remains a significant scatter in the data that is well beyond the experimental error for the small molecule neutron diffraction structures. These observations indicate features beyond $\Delta pK_a$ that influence hydrogen bond length. In the next section, we address the potential origins of, and what additional information may be learned from, this “scatter”.

To conclude this section, we reiterate that there is no indication of general properties of the protein environment that alter hydrogen bonds. The “good” news from these analyses is that hydrogen bond geometries are simpler than has been previously assumed or anticipated. The “bad” news is that one cannot just determine hydrogen bond lengths and directly read out an energetic contribution to binding or catalysis; as we discuss below, energetic properties are strongly environmentally dependent and are considerably more complex (see Hydrogen Bond Energetics: Simpler, but Still Not Simple).

**Coupling within Hydrogen Bond Networks and Other Factors That Affect Hydrogen Bond Length.** The correlations of hydrogen bond length with $\Delta pK_a$ are strong, but scatter remains. To assess whether this variation in length arises from factors beyond experimental uncertainty, we consider just the neutron diffraction data (Figure 6A,B). The neutron O-O and N-O hydrogen bond lengths vary from their

---

**Figure 5.** Hydrogen bond distances derived from $^1$H NMR and ultra-high-resolution (≤1.0 Å) X-ray crystal structures of proteins (cyan points) also follow a linear relationship with donor and acceptor $\Delta pK_a$ for (A) O-O, (B) N-O, and (C) N-N hydrogen bonds. Dark and light gray points are for small molecule hydrogen bonds from neutron diffraction and $^1$H NMR, respectively. Hydrogen bond lengths from $^1$H NMR were, again, estimated using the empirical correlation function from Figure 2B, and lengths of N-O and N-N hydrogen bonds derived from this relationship were corrected by 0.13 and 0.30 Å, respectively, as in Figure 4. Values are reported in Tables S6 and S7. Slopes are reported in Table 1.

**Figure 6.** Linear relationship between hydrogen bond distance and $\Delta pK_a$ that accounts for 86, 74, and 69% of the observed difference in hydrogen bond lengths for O-O, N-O, and N-N hydrogen bonds, respectively ($R^2 = 0.86$, 0.74, and 0.69 for panels A–C, respectively). However, there remains variation that is beyond experimental error in neutron diffraction structures. In neutron diffraction structures of 3,5-dinitrosalicylate, the intramolecular hydrogen bond length varies on the order of 0.1 Å (A, green) despite standard uncertainties of these structures of 0.001–0.010 Å.46

---

**DOI:** 10.1021/acs.biochem.8b00217

Biochemistry XXXX, XXX, XXX–XXX
respective correlation lines on the scale of $\sim 0.1$ Å, well beyond expected errors of $\sim 0.01$–$0.001$ Å. In addition, 36 X-ray structures of the same hydrogen-bonded compound, 3,5-dinitrosalicylate, in different crystal forms gave a range of hydrogen bond lengths on the scale of 0.17 Å, again well beyond the measurement error (Figure 6A, green). Thus, there are factors, beyond $\Delta pK_a$ and experimental uncertainty, that influence hydrogen bond length. The coefficients of determination of $R^2$ of 0.86, 0.74, and 0.69 for O-O, N-O, and N-N hydrogen bond length versus $\Delta pK_a$ respectively (Figure 6), indicate that the difference in electronic properties between the donor and acceptor account for 69–86% of the observed variation in hydrogen bond length. Inspection of the 3,5-dinitrosalicylate crystals revealed differences in hydrogen bonding patterns across the different crystals, suggesting that coupled hydrogen bonds could be responsible for much of the observed variation, and there is compelling evidence for such coupling from studies of hydrogen bond networks within proteins, as briefly described below. To explore coupling effects, active site hydrogen bond networks containing oxynion holes were perturbed for two variants of ketosteroid isomerase (KSI) and photoactive yellow protein (PYP) (Figure 7A), via a combination of site-directed mutagenesis and unnatural amino acid substitution at each position in the network. These studies revealed two basic properties. First, the effects from perturbation of the hydrogen bond network diminished for more remote residues in the network, consistent with successive polarization effects (see ref S0 and below). Second, there was an inverse relationship between hydrogen bond length for the two oxynion hydrogen bond donors such that lengthening of one of the hydrogen bonds (by a direct or proximal perturbation) shortened the other. Remarkably, this coupling follows a constant relationship for 19 different perturbations across the three proteins studied, with a slope of $-0.30 \pm 0.03$ (Figure 7B). This common relationship suggests a limited influence of the surroundings on hydrogen bond couplings. Thus, considering effects from both $\Delta pK_a$ and hydrogen bond networks should allow even more accurate prediction of hydrogen bond lengths, and deviations from the correlation may reveal coupling and other local factors, as discussed in the following section.
Other Factors Affecting Hydrogen Bond Length. We know that steric factors can alter hydrogen bonding. Grossly, a hydrogen bond formed with one protein ligand will not be made with an alternative ligand if the latter’s size prevents access to the site. We also know that other interactions can aid hydrogen bond formation, for example, if active site hydrophobic interactions with a ligand help position it to favor hydrogen bond formation, for example, if active site hydrogen bond formation to the site. We also know that other interactions can aid hydrogen bond formation, for example, if active site hydrogen bond formation to the site.

Here we address the subtler question of whether a protein–ligand hydrogen bond can be altered from its preferred length and geometry by protein structural features, by noting small molecule and protein examples with such effects. First, intramolecular hydrogen bonds are generally shorter than intermolecular hydrogen bonds of the same ΔpK_a, with O–H–O angles of <180°, rather than the favored near-180° angle observed in the absence of constraints (Figures 8 and 9A,B). These differences presumably reflect constraints from the covalent scaffold. Apparently, the energetic cost to reposition the hydrogen atom out of line is smaller than the penalty to rearrange the covalent scaffold to increase the O–O distance and allow the hydrogen bonding groups to relax to their otherwise favored linear configuration. A second example involves substituted phenolates within the KSI oxyanion hole, which donate short hydrogen bonds to para- and meta-substituted phenolates. Phenolates bearing two o-F substituents have hydrogen bonds that are ~0.1 Å longer than those of other phenolates with the same pK_a and bind ~1 kcal/mol more weakly (Figure 10); this disparity with di-ortho substituents is not seen for hydrogen bond complexes in solution. Thus, there appear to be active site constraints that prevent hydrogen bond shortening for the o-F_2 phenolates.

The cost to structurally rearrange the enzyme to allow the hydrogen bond to shorten by ~0.1 Å to its otherwise preferred length must be >1 kcal/mol, even over such a short distance. As we know that proteins are dynamic on this length scale and greater length scales,55 the inability to relax may arise from protein dynamics that are highly anisotropic, coupled, and idiosyncratic, a perspective that is also supported by recent room-temperature X-ray crystallographic results for this system (F. Yabukarski, J. Biel, M. Pinney, T. Doukov, J. Fraser, and D. Herschlag, manuscript in preparation).

What factors in addition to those addressed above might affect hydrogen bond structure? Metal ions in interaction networks would be expected to alter hydrogen bonds, presumably with larger perturbations than from hydrogen bond networks due to their higher charge densities and thus stronger polarizing effects, and there is evidence for such effects (see footnote d). Remote electrostatic fields in proteins could also, in principle, alter hydrogen bond lengths. This model might be tested via prediction of field effects on hydrogen bond lengths followed by experimental length determinations.

**Evaluating the Battle of Forces.** There are three general classes of models for hydrogen bond lengths that fall off the hydrogen bond length versus ΔpK_a correlation. (1) The assigned hydrogen bond length could be incorrect. (2) The

---

**Figure 9.** Intermolecular hydrogen bonds have angles of ~180°, whereas intramolecular hydrogen bonds of similar lengths generally have bent hydrogen bonds. (A) Intermolecular and intramolecular O–H–O hydrogen bond lengths and angles from selected small molecule neutron diffraction structures (CSD IDs are UROXAL01, SUCAB03, LHPAL01-02, and DHNAPH17, from top right). (B) O–O hydrogen bond lengths vs O–H–O angles for inter- and intramolecular hydrogen bonds from small molecule neutron diffraction structures collected in ref 27. Slopes in panel B are ~133°/Å and ~28°/Å for intra- and intermolecular hydrogen bonds, respectively.

**Figure 10.** Crystal structures of pKSI D40N bound to 2,6-difluorophenolate (pK_a = 7.1; orange; Protein Data Bank entry 2INX) and phenolate (pK_a = 10.0; blue; Protein Data Bank entry 2PZV). (A) Overlay of pKSI-2,6-difluorophenolate (orange) and pKSI-phenolate (blue) structures. (B) Side view of the overlay in panel A highlighting the ~15° rotation of 2,6-difluorophenolate relative to the phenolate ligand.55 This reorientation is consistent with repulsion that is partially relieved by the rotation. The 2,6-difluorophenolate hydrogen bond is longer than that for phenolates of the same pK_a (not shown).54
assigned ΔpK_a could be incorrect (because of an improper assignment of the hydrogen bond protonation states or partners). (3) Physical factors that alter the donor and acceptor pK_a values, such as hydrogen bond networks or internal forces, could lead to hydrogen bonds that deviate from their preferred lengths, as noted above.

We often, for the sake of convenience, refer to models as facts. This is particularly the case for structural models arising from X-ray diffraction of protein crystals, but even these models are evaluated for “allowable” bond rotations and interatomic distances during the process of refining and when deposing new structures. We suggest that analogous comparisons to predicted hydrogen bond length ranges might become standard, minimally for very high resolution structures. In addition, analyses of small molecule data sets from the Cambridge Structural Database could lead to predictions of preferred hydrogen bond angles and be implemented for proteins in cases in which hydrogen positions can be assigned (e.g., amide N—Hs). Our initial analysis of hydrogen bonds in high-resolution crystal structures from the Protein Data Bank (PDB) revealed a small set of protein hydrogen bonds with O—O distances that were considerably shorter than predicted. Upon closer inspection, the experimental electron density O distances that were considerably shorter than predicted. Upon closer inspection, the experimental electron density

A predictive model for hydrogen bond lengths could also be used to evaluate structures derived from molecular dynamics (MD), molecular mechanics (MM), and quantum mechanics/molecular mechanics (QM/MM) calculations. Given the often-limited sampling that is possible and the assumptions that underlie these models, assessing the structures that arise and are subsequently used to reach mechanistic and energetic conclusions would help in the evaluation of models and could identify molecular features that may cause distortions in hydrogen bonds, which could then be tested experimentally. As a first step, we evaluated a series of structural snapshots from a recent QM/MM study. We observed hydrogen bond lengths that differed substantially from prediction and had much broader spreads than observed among a large number of high-resolution X-ray structures for that enzyme, suggesting either that the QM/MM structures were not accurate or that this enzyme has large excursions that are not captured in standard X-ray crystallography experiments (unpublished results). To perform this evaluation, we had to request structure files from the study’s authors, and in this instance, our request was graciously accepted. Given that published conclusions are built from QM/MM and related structural models, it seems to be important for these models to be accessible through the scientific literature, consistent with requirements to deposit data that have been developed in other areas and have provided substantial value to those communities (e.g., refs 58–68).

Once hydrogen bond length outliers arising from experimental uncertainty and errors are eliminated and adjustments are made for hydrogen bond networks and coordination to metal ions, we may be left with an interesting set of experimental outliers that may help us better understand the protein interior. For example, how frequently do enzymes exhibit a “force” that lengthens or shortens a hydrogen bond? QM calculations suggest that lengthening a low-ΔpK_a hydrogen bond by 0.1 Å costs ~1 kcal/mol at dielectric constants from 5 to 80. The cost to stretch or bend hydrogen bonds is expected to decrease as ΔpK_a increases, so one can imagine a systematic study of length perturbation across ΔpK_a to assess “energetic vectors” of a protein landscape on this length scale. Given that hydrogen bond length changes going from ground to transition states are typically on a sub-angstrom scale, such knowledge may help us better understand enzyme catalysis and better evaluate catalytic models.

### HYDROGEN BOND ENERGETICS: SIMPLER, BUT STILL NOT SIMPLE

The simplest models for hydrogen bonds are purely electrostatic, with point charges assigned to the donor and acceptor heteroatoms and the central hydrogen atom. Unfortunately, such models do not reproduce the structural and energetic properties of hydrogen bonds and are thus of limited value (e.g., refs 71 and 72). For example, the vast majority of hydrogen bonds are shorter than the sum of van der Waals radii of the hydrogen bonding heavy atoms (Figure 5). Thus, there must be some orbital overlap, and corresponding contributions from partial covalency and charge transfer. Indeed, even quantum mechanics Hartree–Fock calculations, which do not take into account electron–electron correlations, overestimate hydrogen bond lengths. From these types of realizations, additional empirical terms and restraints were added to purely electrostatic models to better capture hydrogen bond distances and angles for use in protein structure prediction and design. Further emphasizing the need to go beyond point charge models, recent *ab initio* path integral calculations suggest that nuclear delocalization is necessary to correctly determine at least some hydrogen bond energies and lengths, with a particularly short hydrogen bond equilibrium incorrectly estimated by 4 kcal/mol upon omission of this term and presumably more modest contributions for longer hydrogen bonds.

The simplification of a transferable energy also does not hold for hydrogen bonds. Benson’s group additivity rules allow researchers to estimate the enthalpies of formation for organic compounds, with group values that can be transferred across the vast majority of compounds. If there were a hydrogen bond energy that was transferable across environments, then we could measure values for each class of hydrogen bond (at a given ΔpK_a) in a model system and use those values to predict hydrogen bond energies in proteins and active sites; thus, the effects of hydrogen bonds on protein stabilities and catalytic rates would be predictable. However, hydrogen bond formation energies are decidedly not transferable. For example, the F-HF hydrogen bond has a free energy of formation, ΔG_{11}^{HB}, of approximately −40 kcal/mol in the gas phase, but a value of only −0.8 kcal/mol in water. Similarly, ΔG_{11}^{HB} for the hydrogen bond between 4-nitrophenol and 4-nitrophenolate is −8 kcal/mol more favorable in acetonitrile and tetrahydrofuran than in water.

An additional complexity arises when considering ΔG_{11}^{HB}: we must analyze species in addition to the hydrogen-bonded complex itself. ΔG_{11}^{HB}, like all equilibria, is a function of the properties of the species on both sides of the equilibrium and all the interactions that are made (Figure 11). The situation becomes even more complex when we compare two equilibria or kinetic processes (Figure 11, top vs bottom), and many biochemical experiments involve such comparisons. For example, enzyme catalysis is assessed relative to a corresponding solution reaction, folding or catalysis of a wild-type enzyme is dissected by comparison to mutant proteins where specific residues of interest are altered (see also footnote h), enzyme
specificity is assessed by comparing reactions with different substrates, and catalyses by enzymes from different organisms are compared to better understand selective pressures and environmental niches. In each case, there are (at least) four states, and changes to any and all states must be considered to understand the physical origins of the energetic effect.

In the face of these pervasive complexities, one simplification arises from the experimental observations and QM calculations outlined in the prior section. As the geometry and potential energy surface of the hydrogen-bonded complex are nearly constant across different solvents and environments, the primary factors affecting free energies of hydrogen bond formation are expected to be differential interactions of the solvent or surroundings between the free and hydrogen-bound states. In contrast, positioning via hydrogen bonds and other structural networks have strong effects on binding and catalysis (e.g., refs 1, 57, 81, and 94–98); these important factors are mentioned below but are not addressed in depth.

A Thermodynamic Framework for Dissecting Hydrogen Bond Energetics: Does Catalysis Arise (Simply) from Exclusion of Water from Active Sites? Figure 12 presents a thermodynamic framework to aid in considering the environmental factors that affect hydrogen bond equilibria and hydrogen bond energetics. As the factors involved in overall catalysis and energetics are varied and complex, we simplified this scheme. Thermodynamic cycles, such as those in Figure 12, are powerful tools to build and test intuition and should stimulate further discussion but should not be mistaken as an attempt at an accurate representation.

It is often stated that hydrogen bonds are “stronger” in nonpolar solvents and, correspondingly, that enzymes can facilitate reactions by removing reactants from aqueous solution and creating a nonpolar-like environment (e.g., refs 98 and 99), but what does this mean with respect to the energetics of enzyme catalysis? Figure 12 allows us to analyze this common model and illustrates that there is not a de facto increase in the level of catalysis from the removal of hydrogen bonds from aqueous solution.

Figure 12 schematically shows equilibria for “hydrogen bond formation” (horizontal lines in each rectangle, Δ\(G^\text{HB}\)) and for changing the “environment” (front to back) from water (front, blue) to a nonpolar aprotic solvent (back, tan). The species in the top equilibria illustrate “hydrogen bond formation” to ground state complexes (Δ\(G^\text{HB}\)), here a carbonyl (C=O); those on the bottom illustrate “hydrogen bond formation” to a transition or intermediate state (Δ\(G^\text{HB}\)), here an oxanion (C−O−), as is common in enzymatic reactions. Thus, the downward dimension on the page is termed “Reaction”.

In the nonpolar environment, the hydrogen bonds are indeed “stronger”, as Δ\(G^\text{HB}_{\text{GS,NP}}\) (back) is more favorable than Δ\(G^\text{HB}_{\text{GS,W}}\) (front), as depicted by the length of the directional arrows for...

---

**Figure 11.** Hydrogen bond formation (Δ\(G^\text{HB}\)) between an acid [A (blue) or A’ (red)] and a base [B (blue) or B’ (red)], where Δ\(G^\text{HB}_A\) (red) is more favorable than Δ\(G^\text{HB}_B\) (blue).

**Figure 12.** Thermodynamic framework for dissecting hydrogen bond energetics in different environments. Equilibria for hydrogen bond formation (Δ\(G^\text{HB}\), horizontal lines) are depicted in water (front, blue) and a nonpolar solvent (back, tan) for a ground state (top line, Δ\(G^\text{HB}_{\text{GS}}\)) and a transition state (bottom line, Δ\(G^\text{HB}_{\text{TS}}\)). Equilibria for transfer between water and a nonpolar solvent (Δ\(G^\text{Env}\), de facto Gibbs free energy) are depicted for non-hydrogen-bonded (left) and hydrogen-bonded (right) ground state (Δ\(G_{\text{GS-W,HB}}\) and Δ\(G_{\text{GS-NP,HB}}\)) and transition state (Δ\(G_{\text{TS-W,HB}}\) and Δ\(G_{\text{TS-NP,HB}}\), top line) and transition state (Δ\(G_{\text{TS-W,HB}}\) and Δ\(G_{\text{TS-NP,HB}}\), bottom line) species. Finally, a schematic reaction from ground state to transition state (vertical “Reaction” coordinate) is depicted in water (front) or a nonpolar solvent (back) without (Δ\(G^\text{HB}_{\text{GS-W}}\) and Δ\(G^\text{HB}_{\text{GS-NP}}\), left) and with (Δ\(G^\text{HB}_{\text{TS-W}}\) and Δ\(G^\text{HB}_{\text{TS-NP}}\), right) hydrogen bond formation.
Δ\(G^{\text{HB}}\) in water and in the nonpolar solvent (Figure 12). Analogously, for the transition state mimic (C–O\(^{-}\)), the bottom \(Δ\!G_{\text{TS,NP}}^{\text{HB}}\) is more favorable than \(Δ\!G_{\text{TS,W}}^{\text{HB}}\).

But how does this relate to enzymatic catalysis? The free energies of hydrogen bond formation (e.g., \(Δ\!G_{\text{TS,NP}}^{\text{HB}}\) and \(Δ\!G_{\text{TS,W}}^{\text{HB}}\)) are not the equilibria that represent catalysis; rather, the vertical equilibria in Figure 12 represent reaction from the ground state to the transition state. We therefore must use the vertical dimension to assess the catalytic effects expected from simple removal of a hydrogen bond from water and placement in a nonpolar environment.

The hydrogen-bonded states on the right side of Figure 12 represent analogous ground states (top) and transition states (bottom) for reaction in aqueous (front) and nonpolar (back) environments.\(^7\) By first principles (Coulomb’s law), the localization of charge is disfavored, and disfavored more so in the lower dielectric of the nonpolar environment. Thus, charge localization in an oxyanionic transition state is disfavored more in the nonpolar environment than in water. Therefore, removing hydrogen bonding groups from water does not, in and of itself, provide catalysis, despite the formation of “stronger” hydrogen bonds [\(Δ\!G_{\text{GS,NP}}^{\text{HB}} < Δ\!G_{\text{GS,W}}^{\text{HB}}\) (Figure 12)]. Indeed, such a transition alone would hinder rather than promote reaction, so there must be additional compensating factors that enzymes use to achieve catalysis.

The conclusion described above is perfectly consistent with the observation of “stronger” hydrogen bonds in the nonpolar solvent, as can be understood through the linked thermodynamic cycle of Figure 12 by considering the other dimension, that of changing the “environment”. “Naked” charges and polar groups are highly unstable in nonpolar solvents. In other words, removing solvating water is energetically costly, and the species without hydrogen bonds are highly favored to partition to water (Figure 12, \(Δ\!G_{\text{GS,HB}}^{\text{TS}}\) and \(Δ\!G_{\text{TS,HB}}^{\text{GS}}\)). Thus, the strong driving force for hydrogen bond formation in the nonpolar solvent, relative to water (\(Δ\!G_{\text{GS,NP}}^{\text{HB}}\) vs \(Δ\!G_{\text{GS,W}}^{\text{HB}}\)), arises because of the absence of competing hydrogen bonding (from water molecules) and the much poorer charge screening or solvation from the low-dielectric, nonpolar solvent than from water. There is no need to invoke “stronger” bonds within the hydrogen-bonded complex; indeed, as described above, there is compelling evidence against such changes in the nature of hydrogen bonds as the environment changes (e.g., Figure 3B).

Finally, we again consider “transition state” species with increased charge density. The higher the charge density, the more the donor and acceptor want to escape from being “naked” in a nonpolar solvent. There is then a strong driving force to form a hydrogen-bonded complex, and \(Δ\!G_{\text{TS,HB}}^{\text{GS}}\) is highly favorable; however, there is an even stronger driving force to escape into water, \(Δ\!G_{\text{TS}}^{\text{GS}}\). Thus, creating a nonpolar environment does not stabilize the charge-dense, hydrogen-bonded transition state, relative to that same transition state in water. As noted above, other factors are needed for catalysis.

**Beyond Hydrogen Bonds.** Of course, a protein “environment” is very different from the nonpolar solvent described above. The properties of this environment can alter hydrogen bond energetics and also provide multiple additional catalytic features. These additional features include side chains that act as nucleophiles and general acids and bases, cofactors that license particular types of chemical transformations, binding interactions that position substrates with respect to one another and with respect to catalytic groups, and the overarching protein scaffold that is responsible for positioning of all of the groups noted above. With respect to hydrogen bonds, the protein environment can better position hydrogen bond donors and acceptors, so that less rearrangement is needed relative to hydrogen bonds in solution, with positioning “paid for” by protein folding energy and ligand binding energy.\(^6\) It is even possible that the hydrogen bond arrangement is precise enough to be suboptimal in the ground state and optimal in the transition state, with folding and binding energy used to enforce this positioning. Additionally, there may be more hydrogen bonds formed within an active site than on average in water, and the active site hydrogen bond donors and acceptors can have charge densities higher than those of water and, as a result, undergo a greater stabilization going from the ground to transition state (assuming an equivalent “solvent” background or enzyme environment for comparison).\(^100,101\)

Another fundamental difference between the protein environment and a nonpolar solvent is that proteins have many dipoles: each peptide bond, many of the side chains, and nearby water molecules.\(^1\) Thus, charge screening on enzymes might be similar overall to charge screening in water, although the detailed energetic accounting would differ. On the other hand, if enough nearby enzyme charges and dipoles were sufficiently immobile and sufficiently oriented to stabilize the transition state and if solvating water molecules were sufficiently remote, then there may be substantial electrostatic advantages to catalysis from these groups. Indeed, some models for enzyme catalysis seem to imply that long-range electrostatic fields play important roles in enzymatic catalysis,\(^102–105\) although experiments to date provide no indication of predicted “enhanced” electrostatic factors within enzyme active sites (e.g., refs 46 and 93), and additional experimental tests are needed and in progress.

Most generally, these and other models accounting for enzymatic catalysis require explicit and specific predictions, preferably quantitative, that allow the models to be tested and distinguished. While we can now readily make a discrete list of catalytic features, they are inextricably linked, and as a consequence, enzyme energetics are highly nonadditive.\(^57,106,107\) The interconnectivity of catalytic mechanisms precludes a simple divide-and-conquer approach to understand enzyme catalysis, necessitating instead more sophisticated models and tests. It is now common to computationally reproduce rate effects from previous experiments, but given the large number of interactions present—and thus potential variables in a computational model—simply matching one or a few overall rate effects is unlikely to adequately distinguish between models that assign different weights to interactions and catalytic features. Indeed, it is not uncommon for different computational models to reproduce prior kinetic results, despite having different underlying mechanisms. In the face of these difficulties and complexities, we need to develop clearly defined models that deliver nontrivial predictions that are different for different models and that are made prior to experiment, so that experiments can be used to test, distinguish, and refine these models.\(^56,108\)

### Closing Remarks: Five “Rules” for Hydrogen Bonding

We close by codifying our current state of understanding into five rules for hydrogen bonds that may aid in future hydrogen bond discussions and may provide a foundation for teaching students of biochemistry about these vital interactions. These rules may help researchers consider structure—function
relationships across the enormous variety of biological systems in which hydrogen bonds play important roles. The development and testing of models for hydrogen bonds described herein and summarized by these rules may also inform the development of predictive models for the more complicated problem of overall enzymatic catalysis and ensuring that these models are testable and distinguishable.

Rule 1: Hydrogen Bond Lengths Do Not Change with the Environment (Figure 3B). This is the powerful simplification outlined herein that allows us to consider a given hydrogen bond species, like a covalent species, as impervious to environment, to a first approximation.

Rule 2: Hydrogen Bonds Shorten as ΔpK_a Decreases (Figure 1B). This rule holds nearly universally (Figures 4 and 5). Deviations from this behavior allow the identification of model error or interesting system features, such as hydrogen bond coupling, limiting steric interactions, and competing structural preferences.

Rule 3: The Free Energy of Formation of a Hydrogen-Bonded Complex (ΔG^HB) Is More Favorable in Nonpolar Environments (Figure 10A). These environments lack competing hydrogen bond donors and acceptors and are far worse than water at solvating separated hydrogen bonding partners, leading to a large driving force for hydrogen bond formation.

Rule 4: Hydrogen Bonds (or Hydrogen-Bonded Complexes) Are Not More Stable in Nonpolar Environments (Figure 10A). The free energy of transfer of a charged or highly dipolar hydrogen-bonded complex from water to a nonpolar solvent is unfavorable. By this simple measure, hydrogen bonds or hydrogen-bonded complexes are, in general, more stable in water.

Rule 5: ΔG^HB Becomes More Favorable as ΔpK_a Decreases in All Solvents, but More So in Nonpolar Solvents. This observation is the basis for one model for how enzymes may take advantage of hydrogen bond behaviors to facilitate catalysis (e.g., refs 5, 20, 33–38, and 101). The dependence of ΔG^HB on ΔpK_a may also provide an assessment of the electrostatic properties of the enzyme environment.13,4,100,101 In the interest of space, we did not present prior results that suggest that the enzyme interior gives energetic behavior similar to that in aqueous solution,66,91 and this and related models are currently undergoing additional tests that will be presented and summarized in the near future.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b00217.

Supplementary tables of compiled neutron diffraction and 1H NMR-derived hydrogen bond lengths and ΔpK_a values, as well as additional clarifying text and figures (PDF)

AUTHOR INFORMATION

Corresponding Author
*E-mail: herschlag@stanford.edu.

ORCID

Daniel Herschlag: 0000-0002-4685-1973
Margaux M. Pinney: 0000-0002-7735-5032

Funding

National Science Foundation (NSF) Grant MCB-1714723. M.M.P. was supported by an NSF Graduate Research Fellowship.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank S. Alvarez for analysis of O-O, N-O, and N-N hydrogen bond lengths in the Cambridge Structural Database, J. Fraser for investigation of hydrogen bond lengths in high-resolution crystal structures from the PDB, F. Yabukarski, D. Mokhtari, and C. Stark for comments on the manuscript, and members of the Herschlag lab for stimulating discussions.

ADDITIONAL NOTES

“Unless otherwise noted, we use hydrogen bond length to refer to the distance between heteroatoms, as shown in Figure 1A. The term “hydrogen bond strength” implies a focus on the hydrogen-bonded complex itself, or perhaps its enthalpy of formation, but is often used in discussions of the free energy of hydrogen bond formation. We refer to the free energy of hydrogen bond formation as ΔG^f^HB herein. Most generally, explicit definitions of processes and equilibria under discussion are critical for clear analyses and communication. For hydrogen bonds, we must consider both the hydrogen-bonded and non-hydrogen-bonded states, as we describe (see Hydrogen Bond Energetics: Simpler, but Still Not Simple).

“A proxy for the electronic properties of hydrogen bond donors and acceptors is their pK_a values. These values are highly sensitive to changes in the electron distributions that result from differences in induction, electronegativity, resonance, and/or polarization and thus provide an experimental readout of these effects. For the sake of simplicity, we use a single common scale of aqueous pK_a values as these values are readily available and correlate with electronic properties derived from quantum chemical descriptions (see refs 28–32).

“In a 0.9 Å X-ray structure of a thermostabilized mutant of carbonic anhydrase (PDB entry 6B00), the interaction between a bound zinc ion is positioned such that it could polarize His94, resulting in a hydrogen bond between His94 and the carbonyl of Gln92 shorter than predicted from its ΔpK_a (ΔpK_a = 16.2, N-O\_observed = 2.79 Å vs N-O\_predicted = 2.93 Å). This length corresponds to that predicted for a hydrogen bonding pair with a ΔpK_a of 5.4 (Figure 4B).

“The relationship between hydrogen bond distance and donor and acceptor ΔpK_a developed herein is for the distance between heavy atoms (O=O, N-O, and N-N) and can therefore be used to evaluate structural models in which proton positions are not known. A linear dependence on ΔpK_a is also observed for H-O distances (see Figure S2 in ref 26; see also ref 27).

“We omit for the sake of this discussion consideration of C=O hydrogen bonds (see refs 69 and 70).

“The sums of heavy atom van der Waals radii are 3.00, 3.16, and 3.32 Å for O = O, N = O, and N = N hydrogen bonds, respectively (van der Waals radii from ref 45). Similar distances are obtained using other van der Waals radius scales (e.g., Bondi109 or Batsanov110).

“Early site-directed mutagenesis work by Fersht80 led to empirical rules for the energetic consequences of mutating neutral and charged hydrogen bonds (0.5–1.5 kcal/mol for...
neutral hydrogen bonds and 3.5–4.5 kcal/mol for charged hydrogen bonds). These values represent complex comparisons of $\Delta G$ values for wild-type and mutant enzymes ($\Delta \Delta G$), encompassing four species in total, any or all of which can be altered.\(^{57,108}\) Nevertheless, the effect of removing a hydrogen bond can depend profoundly on the nature of the mutation, in particular whether water or other groups can substitute for the removed hydrogen bond or whether a hydrogen bond partner is "stranded."\(^{81,82}\) These substitutions may be accompanied by structural rearrangements, which may also contribute to observed energetics (i.e., $\Delta \Delta G$). Rearrangements tend to be observed when nearby interactions can (partially) compensate for the elimination of favorable interactions or offset the introduction of unfavorable interactions (e.g., refs 83–85). Overall, the range of observed values and underlying idiosyncratic effects limit the ability to use such empirical rules in prediction.

There are differences in interactions in gas-phase and even low-dielectric environments, as undamped electrostatic forces can extend over long distances in the absence of charge screening. Indeed, in the gas phase, compounds that exhibit virtually no attraction in condensed-phase environments can exhibit enormous attractive forces (e.g., the enthalpy of formation of the $\text{CF}_3\text{CH}_2\text{F}^-$ complex in the gas phase is $\sim 30$ kcal/mol).\(^{\text{(11)}}\)

The vertical (“Reaction”) comparison is akin to $k_{\text{cat}}$ and is used herein as this is simpler to visualize in Figure 12 than $k_{\text{cat}}/K_M$. Analogous conclusions hold for $k_{\text{cat}}/K_M$.

We use the hydrogen-bonded complex with the hypothetical acid, $\text{H}–\text{A}$, in water to provide the most straightforward comparison, although the hydrogen bond donated by $\text{H}–\text{A}$ will compete with hydrogen bonds donated from water and may not be favored.

To understand hydrogen bond energetics, we focus on the change in charge at the position of the oxanion in Figure 12. We do not address energetics associated with the accompanying positive charge development elsewhere on the transition state. For example, it is possible that upon substrate binding, an anionic general base is isolated from solvent and destabilized, via the use of binding energy, with this destabilization eliminated upon the proton transfer to the general base that accompanies oxanion formation (e.g., see ref 112).

Although active sites are often removed from solution, they are not deeply buried. As a result, electrostatic screening is not expected to differ dramatically within the active site. In addition, some enzymes have active sites that are exposed to water yet do not appear to exhibit anomalous catalytic properties. Consistent with this expectation, energetic contributions of functional units have been shown to be constant despite nearby mutations that expose active sites to solution (ref 113; see also ref 114).

**REFERENCES**


