JOURNAL OF PHYSICAL ORGANIC CHEMISTRY
J. Phys. Org. Chem. 2004; **17:** 511–520
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/poc.769

Strong hydrogen bonding in chymotrypsin and other serine proteases†,‡

Perry A. Frey*

Department of Biochemistry, University of Wisconsin–Madison, 1710 University Avenue, Madison, Wisconsin 53726, USA

Received 4 September 2003; revised 31 October 2003; accepted 31 October 2003

ABSTRACT: The hydrogen bond linking His57-Nδ1 and Asp102-Oδ1 in chymotrypsin (Cht) at low pH and in transition state analogue complexes of Cht with peptide trifluoromethylketones (peptide-TFKs) at pHs up to 12 has been assigned as a low barrier hydrogen bond (LBHB). The hydrogen bonds in these species of Cht display the physicochemical properties of LBHBs, as follows: 1) The proton NMR signals are far downfield, 18.1 ppm for Cht at low pH and 18.6–18.9 ppm for peptide-TFK complexes. 2) The D/H fractionation factors are low, 0.3–0.4 for the peptide-TFK complexes. 3) The deuterium and tritium isotope shifts (δ_D–δ_H, δ_T–δ_H) are negative. 4) The enthalpies of activation for solvent exchange (ΔH_{ex}) are high, 10–19 kcal mol^{-1}. The LBHB is postulated to increase the base strength of His57 in the transition state for the nucleophilic addition of Ser195 to the peptide acyl group of a substrate. This property of His57 is displayed by His57 in the complexes of Cht with peptide-TFKs, in which its p*K*a lies between 10.6 and 12 depending on the structure of the peptide. These values are optimal for an acid/base catalyst that both abstracts a proton from Ser195 in the formation of the tetrahedral intermediate and donates a proton to the leaving *N*-terminal amino group in the decomposition of the tetrahedral intermediate. Strong hydrogen bonds in simple molecules can be studied in both aqueous and nonaqueous solutions, and the conditions for their existence are discussed. Copyright © 2004 John Wiley & Sons, Ltd.

KEYWORDS: Low barrier hydrogen bond; chymotrypsin; transition state analogue; nuclear magnetic resonance; fractionation factor; isotope shift; deshielded proton; peptide trifluoromethyl ketone; peptide boronic acid

THE DOWNFIELD PROTON IN CHYMOTRYPSIN

Until recently, all hydrogen bonds in biological systems were thought to be of a single type, weak electrostatic attractions between non-bonding electron pairs of heteroatoms and weakly acidic protons covalently bonded to other heteroatoms. However, it became clear that a single hydrogen bond in chymotrypsin (Cht) displayed physicochemical properties very different from those of hundreds of other hydrogen bonds in the molecule.[1,2] Similar protons have been found in other serine proteases in the class of Cht, and the unique protons bridge His57-Nδ1 and Asp102-Oδ1 in the catalytic triads of these enzymes.[3,4]

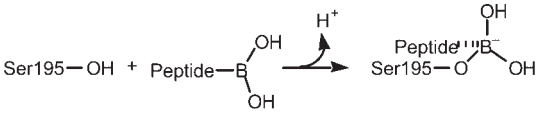
The first sign of uniqueness was the observation that this proton in Cht could be observed in the ^{1}H NMR spectrum, and the other unique feature was that the signal assigned to this proton appeared far downfield from those of all other protons in the molecule, at 18 ppm.[2] The observation of such a proton by NMR meant that its chemical exchange with protons of the solvent must be slow relative to rates for typical acidic and basic groups, which undergo chemical exchange too rapidly to be observed in an NMR experiment. In general, the slowly exchanging protons in proteins are peptide-amide protons engaged in α-helices and β-sheets, which are held in place by highly cooperative hydrogen bonding networks. In contrast, the catalytic triad is not an element of secondary structure.

The downfield position of the NMR signal further defines the uniqueness of the bridging proton in the catalytic triad, in that it must reside in an unusual magnetic environment. Downfield, protons can be found in proteins that incorporate paramagnetic metal ions, which induce contact shifts in neighboring protons. Also, ring currents generated by neighboring aromatic rings can shift a signal downfield, albeit not as far as 18 ppm. However, the structure of Cht does not include a paramagnetic metal ion; nor are there aromatic residues near the catalytic triad that might perturb its magnetic environment. Something within the active site must create the magnetic environment of the proton bridging His57-Nδ1 and Asp102-Oδ1 in Cht. The simplest and most obvious interaction that would shift the signal downfield is the hydrogen bond itself.

**Correspondence to:* P. A. Frey, Department of Biochemistry, University of Wisconsin–Madison, 1710 University Avenue, Madison, Wisconsin 53726, USA.
E-mail: frey@biochem.wisc.edu
Contract/grant sponsor: National Institutes of Health; *Contract/grant numbers:* GM30480; DK28607.
†This paper is dedicated with appreciation and gratitude to Professor William P. Jencks, a great scientist, a great educator and a great friend.
‡Selected paper part of a special issue entitled 'Biological Applications of Physical Organic Chemistry.'

The unique 18 ppm proton in Cht appears in acidic solutions, conditions under which His57-Nε2 is protonated.[2] At pH > 7.5 the His-Nδ1 proton is found at 15 ppm in the NMR spectrum, still a low field for protons. For many years the downfield proton in acidic solutions was not considered mechanistically significant because of the inactivity of Cht at low pH, and the fact that His57 must be in its neutral, basic form to function in catalysis. However, in the reaction mechanism His57 becomes protonated in the tetrahedral intermediate as the direct result of its action as a base. The tetrahedral intermediate is not normally observed because of its short lifetime. However, transition-state analogues of this intermediate can be prepared that have indefinite lifetimes, and they incorporate both protonated His57 and the downfield proton, even further downfield in the best mimics of the tetrahedral intermediate than is found in Cht itself.

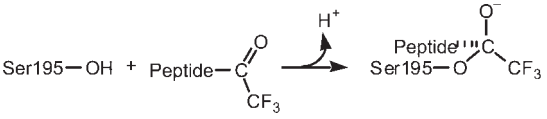
The first tetrahedral complexes of Cht to be studied were those arising in the reactions of boronate, benzene boronic acid, phenylethylboronic acid, and peptide boronic acids with Cht according to Eqn (1):[5,6]



(1)

The proton generated in Eqn (1) is not released from the enzyme but is transferred to His57-Nε2. In these complexes, His57 is in its protonated state at neutral pH, but the NMR signal for the proton bridging His57-Nδ1 and Asp102-Oδ1 in boronate complexes is significantly upfield from that for free Cht at low pH. This may be because the boronate adducts are not strict analogues of the tetrahedral intermediate; the negative charge of the adduct resides on boron, as shown in Eqn (1). A hydroxy group on boron occupies the oxyanion site that binds the oxyanionic group of the tetrahedral intermediate in the catalytic mechanism. The electrostatic attraction between the boronide ion and the adjacent imidazolium ring of His57 may weaken the hydrogen bond bridging His57-Nδ1 and Asp102-Oδ1, and this would move the NMR signal upfield.

Analogues more similar to the tetrahedral intermediate are those resulting from the reaction of a peptide trifluoromethyl-ketone (peptide-TFK) with Cht according to the equation.[7–9]



(2)

A hemiketal adduct of a peptide-TFK with Ser195 is very similar to a tetrahedral intermediate, with the CF_3 group in place of the leaving group NHR. The tetrahedral carbon bears an oxyanion residing in the oxyanion binding site, similar to the oxyanionic tetrahedral intermediate in catalysis. His57-Nε2 accepts the proton generated in Eqn (2), and the proton bridging His57-Nδ1 and Asp102 appears in the NMR spectrum at 18.6–19.0 ppm, depending on the structure of the peptide moiety. This proton appears well downfield from its position in free Cht, and it persists in the NMR spectrum at pHs well above neutrality.[1,8–11]

One way to conceptualize the relationship of the downfield proton in Cht with well-studied hydrogen bonds in chemistry is to consider the classes of hydrogen bonds as defined by physical chemists and physical organic chemists.

WEAK, LOW BARRIER AND DEEP WELL HYDROGEN BONDS

Hydrogen bonds have been controversial through much of the past century.[12,13] Among the reasons may have been their apparently ephemeral nature. One striking aspect is the great difference in hydrogen bond strengths, from 2 to 40 kcal mol^{-1} (1 kcal = 4.184 kJ), a 20-fold range. In contrast, the energies of covalent bonds extend from 25 to 120 kcal mol^{-1}, only a 5-fold range. Another initially confusing aspect was the question of whether a weak attraction should be regarded as bonding. However, by the mid-twentieth century the rules for weak hydrogen bonding had been revealed, and the consequences for molecular structure became known, and weak hydrogen bonding was generally accepted.

The ultra-strong, 40 kcal mol^{-1} hydrogen bond in hydrogen difluoride $[F\cdots H\cdots F]^-$ was also discovered before mid-twentieth century, although there was a tendency to regard it as an anomaly. In the second half of the twentieth century, the spectroscopic and chemical properties of the proton in hydrogen difluoride became better known, and much new information about special hydrogen bonds in many other molecules also appeared. Consequently, other strong and very strong hydrogen bonds were discovered. The physical properties of a large number of molecules led naturally to the definition of the three classes of hydrogen bonds, weak, strong and very strong, in a comprehensive review.[14]

The key physical properties distinguishing weak, strong and very strong hydrogen bonds are those in Table 1.[14] The NMR downfield chemical shifts are the most universal single property of strong and very strong hydrogen bonds. The downfield positions vary with the size of the heteroatom and should be compared within types, e.g. O···H···O, N···H···O, N···H···N, etc. The entry > 16 ppm in Table 1 is somewhat arbitrary and refers to the shortest of all hydrogen bonds, that in hydrogen difluoride, in which the F···F distance is 2.26 Å. A very strong hydrogen bond between atoms like O separated by 2.4 Å would appear at 21 ppm in the NMR spectrum. Strong and very strong hydrogen bonds are formed when the heteroatoms display comparable proton affinities and are separated by less than twice the sum of their van der Waals radii. For O···H···O systems

Table 1. Physical properties that distinguish weak, strong and very strong hydrogen bonds

Property	Weak	Strong (LBHB)[a]	Very strong (DWHB)[a]
δ_H (ppm)	5–12	>16[b]	>16[b]
ϕ[c]	1.0–1.2	0.3–0.7	<1.0
δ_D–δ_H (ppm)	0	−0.2 to −0.8	+0.2 to +0.5
ν_{AH}/ν_{AD}[d]	1.4	1.0–1.2	1.2–1.3

[a] LBHB refers to low barrier hydrogen bond and DWHB to deep well hydrogen bond (see Fig. 1).
[b] The positions of downfield protons in LBHBs and DWHBs depend on the heteroatoms. For O···H···O and N···H···O the range is 17–19 for LBHBs and 19–21 for DWHBs, in which the hydrogen is equally shared.
[c] D/H fractionation factor, generally smaller for LBHBs than for DWHBs.
[d] The normal ratio is 1.4 and governed by the ratio of reduced masses of H and D.

the O···O distance must be <2.55 Å, and for N···H···O systems the N···O distance must be <2.70 Å.

All of the parameters in Table 1 display analogous behavior when plotted against the distance R_{AB} separating the heteroatoms; they pass through either a minimum $\{\phi, \nu_{AH}/\nu_{AD}\}$ or a maximum $\{\delta_H, \Delta(\delta_D - \delta_H)\}$ as the distance between heteroatoms decreases.[14] In the cases of O and N the distance can never be as close as in hydrogen difluoride, so that a 16 ppm signal in a hydrogen bond involving O or N does not indicate extraordinarily strong hydrogen bonding. In general, strong hydrogen bonds involving only O and/or N display NMR chemical shifts of 17–19 ppm, and very strong hydrogen bonds appear at 19–21 ppm.

The physical basis for the downfield signal is elongation of the covalent bond linking the proton to a heteroatom. Elongation decreases shielding of the proton from the external magnetic field by the nonbonding electrons of the heteroatom. Strong hydrogen bonding lengthens the covalent bond much more than weak hydrogen bonding and exerts a greater deshielding effect. Structural support for this concept is provided by an analysis of the covalent bonds O—H and hydrogen bonds H···O in a large number of O—H···O interactions in small molecules. The structures unmask two important relationships. The longer the covalent bond O—H the shorter the hydrogen bond H···O, until the two became equal at 1.2 Å in symmetrical, very strong hydrogen bonds as O···H···O.[13,15] Implicit in this relationship is the apparently counter-intuitive fact that the shorter the distance R_{OO} for O···O the longer the covalent bond O—H. This makes sense because the more the proton is shared between heteroatoms the less closely it is held by either of them. Hence the stronger the hydrogen bond the longer the covalent bond and the further downfield the proton signal until symmetry is attained. The relationship is not linear, as shown in a plot of R_{OO} against chemical shifts, but the values approach 22 ppm in the limit.[16]

It is worth noting that an unshielded, free proton displays a chemical shift of 30 ppm. Thus, a proton in the category of an LBHB or deep well hydrogen bond is substantially deshielded relative to a typical, strictly covalently bonded proton (δ_H 1–10 ppm).

Isotope effects on physical properties support the correlations between crystal structures and NMR chemical shifts. The D/H fractionation factors ϕ, the heavy hydrogen isotope effects $\Delta(\delta_D$–$\delta_H)$ on the chemical shift, and the ratios of hydrogen and deuterium stretching frequencies ν_{AH}/ν_{AD} define three classes of hydrogen bonds, as shown in Table 1. Each of these parameters

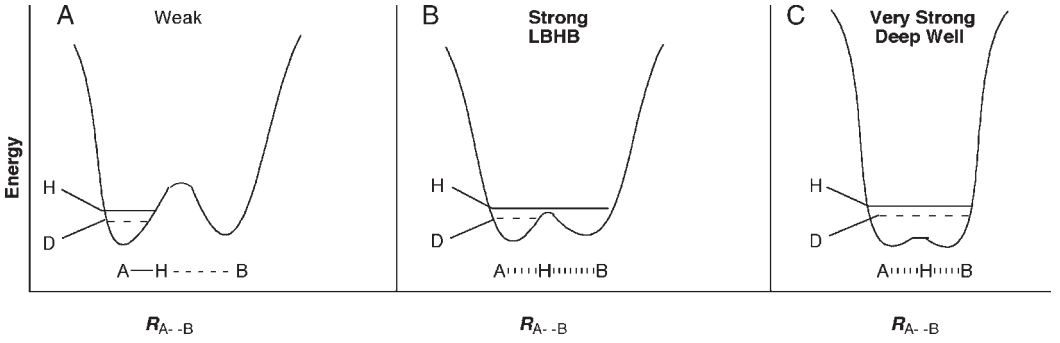


Figure 1. Zero point energy effects in three classes of hydrogen bonds. (A) In a weak hydrogen bond, the proton is bonded covalently to heteroatom A in one side of a double well potential, with its zero point energy deeply below the barrier. The zero point energy of deuterium is even lower. (B) In a strong hydrogen bond, the heteroatoms are closer together and the barrier in the double minimum is lowered significantly to slighly below the zero point energy of hydrogen. Then the zero point energy of deuterium (and tritium) are below the barrier, giving rise to substantial deuterium isotope effects on several physical parameters (Table 1). Because the zero point energy of hydrogen is above the barrier, the proton is drawn toward heteroatom B and away from heteroatom A but is not equally shared between them. Because the barrier is low but still influential, the strong hydrogen bond has been called a low barrier hydrogen bond or LBHB.[1] (C) In a very strong hydrogen bond, the heteroatoms are very close together and the barrier is either not present or far below the zero point energies of hydrogen, deuterium and tritium. The proton is similarly attracted to both heteroatoms and essentially between them. Zero point energy effects on physical constants are absent because of the absence of a barrier effect. Because the proton is drawn away from the heteroatoms in both a strong (LBHB) and a very strong hydrogen bond, it is deshielded from an external magnetic field, and its resonance appears far downfield in an NMR spectrum (Table 1)

distinguishes the strong class of hydrogen bonds from the weak and very strong classes. The isotope effects signal the intervention of zero point energies in the relationships among the three classes and lead to the diagrams in Fig. 1. According to this model, in a strong hydrogen bond the zero point energy for the proton lies slightly above the barrier in a double well potential, leaving the zero point energy of deuterium (or tritium) below the barrier and bringing about the isotope effects. For this reason, the strong hydrogen bonds can be regarded as low barrier hydrogen bonds, or LBHBs.[1] In the very strong hydrogen bond, the barrier is well below the zero point energies of both hydrogen and deuterium or absent entirely, abolishing the zero point energy effects. Because of the absence of a barrier effect, the very strong hydrogen bonds can be regarded as deep well hydrogen bonds.

The very strong hydrogen bonds have sometimes been called single well hydrogen bonds.[1] However, this designation is less useful than deep well hydrogen bonds because of the implication of the absence of a double minimum potential, a rare event. Most very strong hydrogen bonds display double minima that lie well below the zero point energies of hydrogen and deuterium, and isotope effects attributable to zero point effects are absent. In the case of weak hydrogen bonds, the zero point energies of both hydrogen and deuterium lie well below the barrier in the double well potential and no zero point effects are observed.

The strengths of the three classes of hydrogen bonds depend on the physical state. Reported strengths often refer to the hydrogen bond in a vacuum and to the energy difference with and without the hydrogen bond between two heteroatoms at a given R_{AB}. Quoted values under these conditions are 2–12 kcal mol^{-1} for weak bonds, 12–24 kcal mol^{-1} for strong hydrogen bonds (LBHBs) and >24 kcal mol^{-1} for very strong hydrogen bonds (deep well hydrogen bonds).[14] Transfer of a very strong hydrogen bond from a medium of dielectric constant of 1.0—a vacuum—to a medium of dielectric constant 80 corresponding to water weakens it by 50%.[17] The strengths of the three classes in media of effective dielectric constants between 10 and 80 would be relevant to the action of enzymes.

THE LBHB IN THE CATALYTIC TRIAD

The assignment of the proton bridging His57-Nδ1 and Asp102-Oδ1 as an LBHB in acidic solutions of Cht and in adducts with peptide-TFKs[1] has been further strengthened in other studies.[10,11,18–22] The LBHB in chymotrypsinogen at low pH displays a low fractionation factor and a high value of the activation enthalpy for exchange with solvent protons, $\Delta H^{\ddagger}_{ex}$.[18] The NMR signal for the downfield proton in serine proteases is broadened by the influence of chemical exchange at a rate comparable with spectrometer frequencies, so that temperature effects on signal width can be employed to evaluate $\Delta H^{\ddagger}_{ex}$.[18,23] High values of $\Delta H^{\ddagger}_{ex}$ are regarded as indicative of strongly bonded protons.

The LBHBs in peptide–TFK adducts of Cht are stronger than that in Cht itself at low pH, as indicated by their physical properties in Table 2. The inhibitors vary in structure, from *N*-acetyl-L-phenylalanine (AcF-CF_3), to *N*-acetylglycyl-L-phenylalanine (AcGF-CF_3), to *N*-acetyl-L-valyl-L-phenylalanine (AcVF-CF_3) and to *N*-acetyl-L-leucyl-L-phenylalanine (AcLF-CF_3). The downfield chemical shifts, low H/D fractionation factors and negative tritium isotope shifts all confirm the assignment of LBHBs, as distinguished from either weak or very strong hydrogen bonds. Further, the values of $\Delta H^{\ddagger}_{ex}$ are very large and point to strongly held protons. The rate constants for chemical exchange with solvent protons decrease with increasing strength of the LBHBs, as do the inhibition constants for the four compounds. The inhibition constants refer to the hydrated compounds in water,

Table 2. Properties of the downfield proton in tetrahedral complexes of peptide trifluoromethyl ketones with chymotrypsin

	Inhibitor			
Physical property	AcF-CF_3	AcGF-CF_3	AcVF-CF_3	AcLF-CF_3
δ_H (ppm)[a]	18.6	18.7	18.9	19.0
ϕ[b]	0.32	0.34	0.38	0.43
δ_T–δ_H ppm[c]	−0.63	—	−0.65	−0.68
k_{ex} (s^{-1})[d]	282	123	—	12.4
$\Delta H^{\ddagger}_{ex}$(kcal mol^{-1})[e]	15	16.0	—	19
K_I (mM^{-1})[f]	17, 30 20, 40	18, 12	2.8, 4.5	1.2, 2.4
pK_a[g]	10.7	11.1	11.8	12.1

[a] His57-Hδ1.[11]
[b] D/H fractionation factor of His57-Hδ1.[19]
[c] Tritium isotope shift for His57-Hδ1.[22]
[d] Rate constant for exchange of His57-Hδ1 with solvent protons at 25 °C.[19]
[e] Activation enthalpy for exchange of His57-Hδ1.[19]
[f] Inhibition constants reported.[7–9]
[g] pK_a of His57-H$^{\varepsilon}$2 reported.[10,11]

not to the dissociation constants for the ketone forms. The hydration constant for AcLF-CF_3 is 4500, so that the value of K_d for the dehydrated, ketone form of this inhibitior is 0.2 nM, indicative of a transition-state analogue.[9] The hydration constants for the other inhibitors are likely to be similar.

The crystal structures of Cht with AcF-CF_3 or AcLF-CF_3 bound to Ser195, at a resolution of 1.8 Å, supported the assignment of an LBHB between His57-Nδ1 and Asp102-Nδ1.[24] The Nδ1···Oδ1 distances were 2.5–2.6 Å, satisfying the criterion of $<$2.7 Å for an LBHB of the type N···H···O. The spacing has been confirmed at a resolution of 1.4 Å.[21]

The assignment of LBHBs in the catalytic triads of serine proteases in general are further supported by studies of mechanistically related enzymes. The LBHB proton in the transition-state analogue complex of *N*-acetyl-L-leucyl-L-leucine trifluoromethyl ketone with subtilisin labeled with ^{15}Nδ1 displays proton-^{15}N coupling, and the magnitude of this interaction indicates that the proton is 20–30% transferred.[25] This proton also displays a value of 0.55 for the D/H fractionation factor. Further, an atomic resolution (0.78 Å) crystal structure of subtilisin allowed the electron density of hydrogen atoms to be examined.[26] The proton bridging His-Nδ1 and Asp-Od1 in the catalytic triad was modeled as an LBHB, with an Nδ1···H distance of 1.2 Å and a H···Oδ1 distance of 1.5 Å. The N—H bond is clearly elongated, and to the degree indicated by the NMR spectroscopy. Ultrahigh resolution structures of elastase (0.95 Å) and proteinase K (0.98 Å) gave similar results.[27,28]

The chemical shift values and fractionation factors indicate that the LBHB in Cht at low pH is not as strong as that in the peptide–TFK adducts. The 18 ppm signal of free Cht is significantly upfield from the 18.6–19 ppm signal of the Cht–peptide-CF_3 complexes. The D/H fractionation factor for chymotrypsinogen is 0.5,[18] higher than the 0.32–0.43 for Cht–peptide complexes.[19] Moreover, the value of $\Delta H^{\ddagger}_{ex}$ for chymotrypsinogen is 10 kcal mol^{-1}, significantly lower than the 14–19 kcal mol^{-1} for the Cht–peptide-CF_3 complexes. The differences may be due to tightening of the contact between His57 and Asp102 induced by the binding of the transition-state analogue. Alternatively, the presence of the transition-state analogue may decrease the effective dielectric constant in the active site, thereby strengthening the LBHB. In either case, the LBHB is a common feature of Cht at low pH and Cht in the transition-state analogue.

THE LBHB IN CATALYSIS

The mechanism for acylation of chymotrypsin, including the LBHB in the structure of the tetrahedral intermediate, can be formulated as in Fig. 2. Upon reaction of a peptide with Cht, His57-Nε2 abstracts the proton from the 3-hydroxy group of Ser195 as the oxygen undergoes

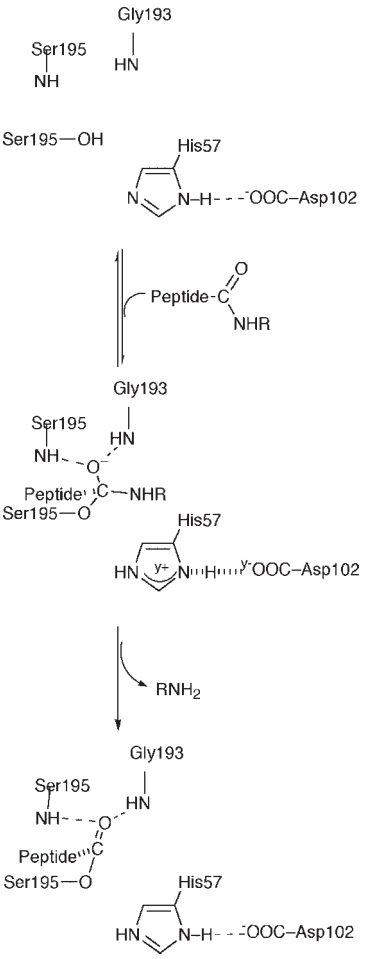


Figure 2. Role of the LBHB in the catalytic triad of chymotrypsin. The acylation of Cht takes place in two chemical steps following substrate binding to form the Michaelis complex, addition of Ser185 to the acyl carbonyl group of the substrate to form a tetrahedral addition intermediate followed by elimination of the leaving group to cleave the peptide and form the N-terminus of the product. In peptide hydrolysis, the first step, formation of the tetrahedral adduct, limits the rate. Proton transfer is essential in both steps, from Ser195 to His57 in the first step and from His57 to the leaving N-terminal amino group in the second. In the tetrahedral intermediate, His57-Nε2 is protonated. Structural analogue of the tetrahedral intermediate that lack a leaving group incorporate an LBHB between His57-Nδ1 and Asp102-Oδ1 (Table 2). It is postulated that the strength of the LBHB facilitates the formation of the tetrahedral intermediate by increasing the basicity of His57 in the transition state[1]

nucleophilic addition to the peptide carbonyl group. The resulting tetrahedral intermediate incorporates the LBHB between His57-Nδ1 and Asp102-Oδ1, an oxyanionic group in the oxyanion binding site and the leaving group NHR adjacent to His57-Nε2, which bears a proton. In Fig. 2, the LBHB is flanked by the imidazole ring of His57 and the β-carboxyl group of Asp102, which are assigned charges of $y+$ and $y-$, respectively, where $1 > y > 0.5$.[1] This notation signifies the partial depolarization of the ionic bond by the LBHB. In the second step, the intermediate eliminates the leaving group under the

driving force of the oxyanion and with catalysis by proton transfer from His57-Nε2 to the leaving group. The scheme is Fig. 2 describes the process but does not give information about the significance of the LBHB.

The structures of adducts formed between Ser195 of Cht and peptide-TFKs are very similar to those of tetrahedral intermediates in the catalytic mechanism; they differ essentially by the replacement of the leaving group (NHR) in an intermediate with the CF_3 group in the inhibition complexes. The groups NHR and CF_3 are both electronegative and electrostatically neutral, but the CF_3 group cannot be protonated and is not a leaving group. Because CF_3 cannot leave, the peptide-TFK adducts can be studied. The inhibition constants for the peptide-TFKs in Table 1 are well correlated with the values of k_{cat}/K_m for the Cht-catalyzed hydrolysis of the corresponding methyl ester substrates; the higher the value of k_{cat} for the ester the lower in the value of K_I for the trifluormethyl ketone, and the two parameters obey a linear free energy relationship.[9] The same relationship extends to the values of chemical shift and pK_a of His57.[11] There is every reason to regard the peptide-TFK adducts as analogue of the corresponding tetrahedral intermediates. It follows that the LBHB is almost certainly an integral part of the structure of the tetrahedral intermediate.

It has been pointed out that when the LBHB is disrupted, either by methylation of His57 or mutation of Asp102, the catalytic rate decreases by 10^4–10^5-fold.[1] The kinetic consequences of methylating His57 is a 10^5-fold decrease in rate,[29] and mutation of His57 in trypsin to Asn decreases the rate by 10^4-fold.[30] These effects offer a guide to the kinetic importance of the LBHB and indicate that it can lower the activation energy by up to 5.5–7 kcal mol^{-1}, provided that the kinetic mechanism remains the same after methylation of His57 or mutation of Asp102. Should the alterations change the rate-limiting step, the effect of the LBHB could be larger. The effect seems substantial in either case.

It has been suggested that the LBHB might decrease the barrier to the formation of the tetrahedral intermediate by increasing the basicity of His57-Ne2 in the transition state.[1] The test of this hypothesis would be a detailed structure–function analysis varying the basicity of His57-Nε2 and the strength of the LBHB and correlating them with the kinetic consequences. As a start on this, one can consider that the Cht–peptide-TFK complexes are transition-state analogues. They are analogues of the tetrahedral intermediate, which based on the Hammond postulate is similar to the transition state. Therefore, the properties of tetrahdral inhibitor complexes can report on the properties of the transition states for reactions of analogous substrates.

The plot of δ_{LBHB} for the Cht–peptide-TFKs in Table 2 against $\log(k_{cat}/K_m)$ for the hydrolysis of the corresponding peptide methyl esters is linear and displays a positive slope.[11] This indicates that the rate depends on the strength of the LBHB. One measure of the basicity of His57-Nε2 in the transition state is the pK_a of His57 in the peptide-TFK complexes with Cht. These pK_as can be measured by observing the transition of the downfield NMR signal for the LBHB from 19 ppm for the adduct to 15 ppm with increasing pH. The pK_as are indeed high, in accord with the postulated role of the LBHB, and they vary with the structure of the peptide and plot as a line against $\log(k_{cat}/K_m)$, again with a positive slope, showing that the rate increases with increasing basicity of His57-Nε2.[11] Because of the narrow ranges of δ_{LBHB} and pK_a in the available studies, the results should be accepted with caution; however, it is significant that the correlations are in the mechanistically meaningful direction, that is, the slopes are neither zero nor negative.

The pK_a values for His57 in the peptide-TFK adducts fit nicely into the mechanistic requirements for base catalysis in the acylation of chymotrypsin, and also in deacylation. The ideal base to catalyze acylation would be strong enough to abstract the proton from Ser195 (p$K_a \approx 13$) but not so strong as to be unable to donate a proton to the leaving group, the N-terminal amino group of a peptide (p$K_a \approx 9$). The pK_a values of 10.6–12 for the peptide-TFK adducts lie well within the optimal range for base catalysis in the action of chymotrypsin.

REACTION OF Cht WITH AcLF-CHO

An important question about the LBHB has to do with the pK_a of 7 for His57-Nε2 in free Cht.[5,6] If the LBHB increases the basicity of His57-Nε2 in catalysis, and protonation of His57-Nε2 leads to LBHB formation, how does it happen that His57 in free Cht displays a normal pK_a? The properties of *N*-acetyl-L-leucyl-L-phenylalanal (AcLF-CHO) as an inhibitor of Cht sheds light on this issue.[21] AcLF-CHO differs from the best TFK inhibitor, AcLF-CF_3, by the substitution of H for CF_3, which transforms the trifluoromethyl ketone into an aldehyde. AcLF-CHO is also an inhibitor of Cht, and it also forms a covalent adduct with Ser195; however, the dissociation constant for the aldehyde is 8000 times that for the trifluoromethyl ketone. Moreover, although the structure of the Cht–AcLF-CHO complex is very similar to that of the complex with AcLF-CF_3, there are important differences. There are also important chemical differences, as revealed by ^{1}H and ^{13}C NMR spectroscopy and biochemical analysis.

The high-resolution crystal structure of the Cht–AcLF-CHO complex shows many similarities with and one important difference from that of Cht–AcLF-CF_3.[21] The overall structures are almost identical, including the contacts between Cht and the peptide moiety. The spacing between His57-Nδ1 and Asp102-Oδ1 is very similar to that in Cht–AcLF-CF_3, 2.6 Å, characteristic of an LBHB. However, the hemiacetal oxygen originating with the aldehyde carbonyl group does not reside exclusively

in the oxyanion binding site of Cht. Instead, the adduct consists of two epimers in a ratio of about 40:60, in which the epimeric center is the hemiacetal carbon. The epimer mixture arises through the addition of Ser195 to either face of the aldehyde group in AcLF-CHO. In one epimer the hemiacetal oxygen is in the oxyanion site, and in the other epimer the oxygen is within hydrogen bonding distance of His57-Nε2.[21]

NMR spectroscopy and the biochemical properties of the Cht–AcLF-CHO complex revealed the reason for the epimer mixture.[21] The ^{1}H NMR spectrum showed that in neutral solution His57 was not protonated; the downfield proton appeared at 15.1 ppm, and there was no LBHB. His57 could be titrated with acid and became protonated with a pK_a of 6.5 to generate an LBHB between His-Nδ1 and Asp102-Oδ1 that appeared at 17.8 ppm in the ^{1}H NMR spectrum. The pH dependence of the ^{13}C NMR signal of the Cht complex with AcLF-^{13}CHO showed no evidence that the hemiacetal oxygen underwent an ionization between pH 6 and 13, suggesting that it was a hydroxyl group with a p$K_a > 13$. The NMR studies suggested that both the hemiacetal and the imidazole ring of His57 were neutral at pH ≥ 7 and above. That is, the addition of Ser195 to the aldehyde carbonyl proceeded with proton transfer from serine to the hemiacetal oxygen and not to His57. If so, then the binding of AcLF-CHO could not involve proton uptake or release, and this was proved to be the case. The epimeric structure of the complex could then be assigned as A and B in Fig. 3. The structures show that the hemiacetal adducts of AcLF-CHO with Cht are not transition-state analogues. The key difference is the destination of the proton derived from Ser195. Its placement on the hemiacetal oxygen neutralizes both the tetrahedral adduct and His57.

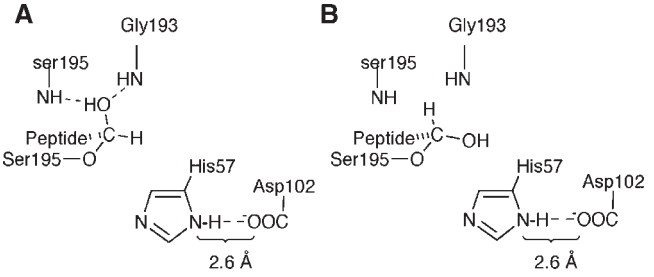


Figure 3. Epimeric structures of the tetrahedral adducts of AcLF-CHO with Cht. The reaction of the peptide aldehyde AcLF-CHO with Ser195 in Cht leads to an epimeric mixture of hemiacetals. In both hemiactals, the 3-hydroxyl group of Ser195 is added to the carbonyl group of the inhibitor and the proton liberated from Ser195 is bonded to the hemiacetal oxygen. In epimer A, the hemiacetal OH group lies in the oxyanion binding site of Cht and in epimer B it lies within hydrogen bonding distance of His57-Nε2. The spacing between His57-Nδ1 and Asp102-Oδ1 is similar to that in the hemiketal complexes of peptide-TFKs with Cht and with that in Cht at acidic pHs. However, there is no LBHB in the AcLF-CHO complexes because His57-Nε2 is not protonated. The proton liberated upon hemiacetal formation with Ser195 resides in the hemiacetal hydroxyl group

The absence of electron withdrawal on the tetrahedral carbon in the Cht–AcLF-CHO-complex makes the adduct more basic than His57-Ne2, so that in the competition for the proton from Ser195 the hemiacetal wins. In the tetrahedral intermediate, with NHR as the fourth substituent of the central carbon, the intrinsic pK_a of the oxygen can be expected to be between 11 and 12, with reference to ionization in water, similar to that of His57-Nε2 in the Cht–AcLF-CF_3 complex. Then His57-Nε2 can be protonated and the tetrahedral intermediate negatively charged, as in the mechanism of Fig. 2. The Cht–peptide-TFK complexes are similar to the intermediate because the pK_a of a peptide-TFK hemiketal is $\sim$9.5, well below that of His57-Nε2, and the proton derived from Ser195 resides on His57-Nε2. In contrast, the pK_a of a hemiacetal of AcLF-CHO is at least 13.8, well above that of His57-Ne2, so that His57 cannot compete with the Cht–AcLF-CHO adduct for the proton.[21]

The ionization properties of His57 in the Cht–AcLF-CHO complex are similar to those of free Cht, with modest differences in the values of pK_a and δ_H for His56-Nδ1. Therefore, the presence of the peptide and the tetrahedral carbon do not affect the electrostatic properties of the active site in an important way. The crucial difference from the peptide-TFK complexes is brought about by the presence of a hydrogen atom in place of the CF_3 group, which eliminates the electron-withdrawing inductive effect on the tetrahedral carbon. The consequences of this difference reveal much about the properties of the transition-state analogue complexes with peptide-TFKs.

The pK_a of His57-Nε2 is governed by the overall electrostatic charge in the active site, and it is remarkably little affected by other contacts between Cht and peptide tetrahedral adducts. The active site is here defined by Ser195 and its substituents, His57 and Asp102. The site is stable when it bears an overall charge of -1, and any deviation from this charge presents a barrier that must be overcome, either by perturbations in the ionization constants at the active site or by some sort of stabilizing interaction. The ionization behavior of the Cht–peptide-TFKs are instructive. Consider the ionizations of the hemiketal oxygen and His57, as illustrated in Fig. 4(A). Both ionizations alter the preferred charge of -1, and the price for this is a pK_a altered by $\sim$5 units from the standard aqueous value, corresponding to 7 kcal mol^{-1} in terms of free energy at 25 °C.

The ionizing properties of His57 in free Cht and in the Cht–AcLF-CHO complex are similar and dramatically different from the transition-state analogue complexes. The ionization of His57 proceeds with a change in net charge of the active site, but with little apparent perturbation in the pK_a of His57 relative to values for histidine peptides in aqueous solution, as shown in Fig. 4(B).[21] How is the barrier to ionization in the active site overcome? The simplest and most obvious rationale is the interaction between His57 and Asp102, the LBHB. In a thought experiment, the absence of an interaction can be

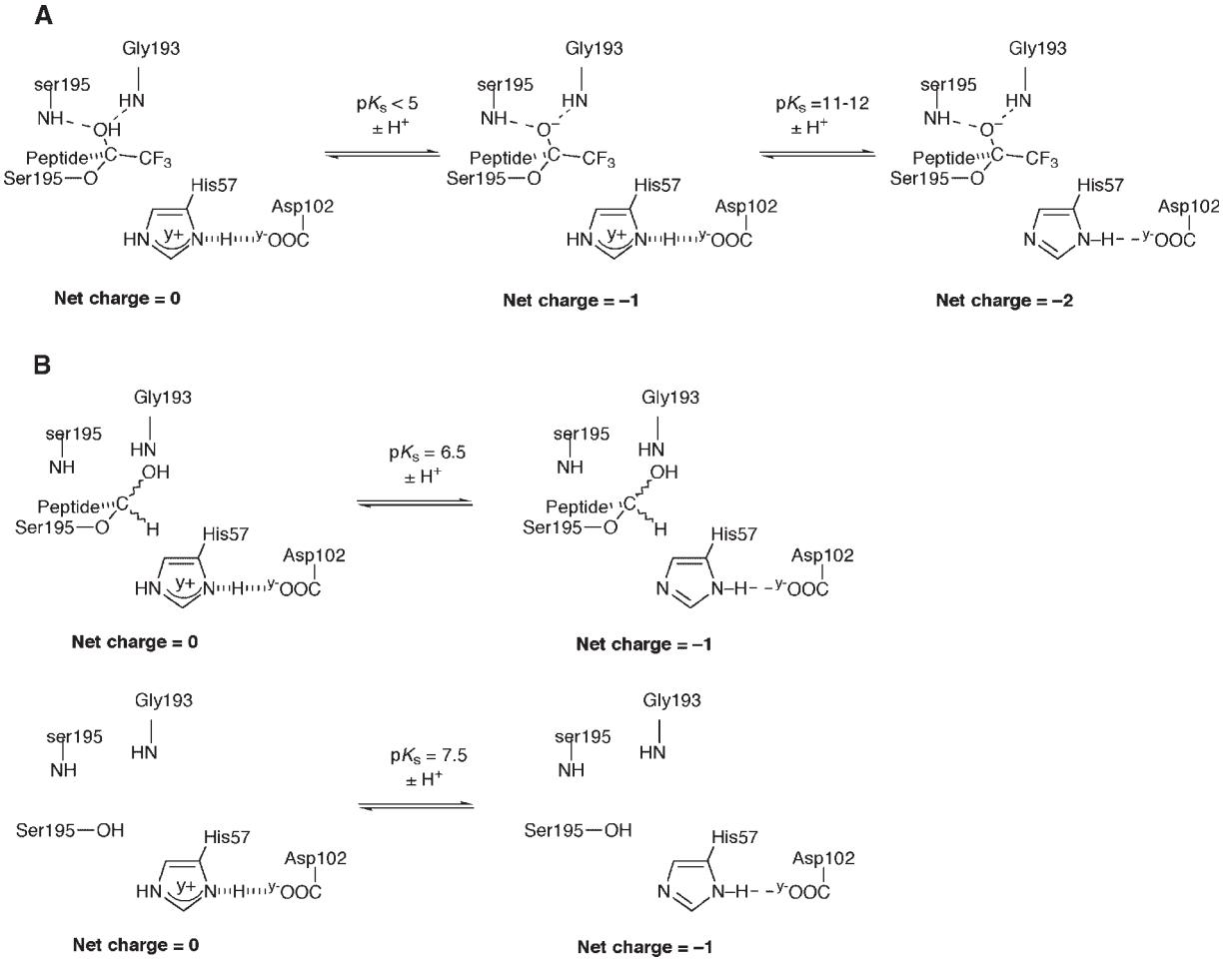


Figure 4. Ionizing properties of Cht and tetrahedral complexes with peptide derivatives. (A) The structures and ionizing properties of the hemiketal adducts of Cht with peptide-TFKs are illustrated. The preferred overall charge in the active site is −1 and the ionization constants of His57 and of the hemiketal are correspondingly perturbed. (B) The structures and ionizing properties of free Cht and its hemiacetal complex with a peptide aldehyde are illustrated. The ionizing properties of His57 and the hemiacetal appear normal. The discussion in the text puts forward a hypothesis to resolve the apparently paradoxical differences in ionizing properties of the structures in (A) and (B)

imagined. In such an experiment, protonation of His57 would lead to an unstable state where the net charge would be zero, it would be difficult to protonate His57 and the pK_a would be very low. However, in reality the His–Asp interaction is stabilizing and eases the acquisition of a proton. Consequently, the pK_a is brought into the neutral range. In this framework, the His–Asp interaction does elevate the pK_a, but it elevates it into the neutral range, where it is mechanistically significant.

HYDROGEN BONDING IN COMPLEXES OF *N*-METHYLIMIDAZOLE WITH CARBOXYLIC ACIDS

The foregoing rationalizes the ionizations at the active site of Cht and transition-state analogue complexes, and it explains the catalytic role of His57 as a strong base in the transition state for tetrahedral intermediate formation. However, it does not explain the role of the LBHB. Could not a simple ionic bond between His57-Nδ1 and Asp-Oδ1 suffice? The LBHB is ionic, albeit depolarized by the LBHB. Experiments prove the presence of the LBHB, but they are essentially descriptive in nature and do not explain its role. Studies of complexes between *N*-methylimidazole and a series of carboxylic acids in chloroform shed further light on the conditions for LBHB formation.[31,32]

^{1}H NMR and FTIR spectroscopic data show that the formation of strongly hydrogen-bonded complexes between 1-methylimidazole and carboxylic acids in chloroform critically depends on the pK_a of the carboxylic acid. Among carboxylic acids of varying strength, pK_a-0.23–4.8 in water, only 2,2-dichloropropionic acid forms a strongly hydrogen-bonded complex with 1-methylimidazole. Weaker acids form exclusively neutral complexes and stronger acids form exclusively ionic complexes with 1-methylimidazole. The results show that an acid with the pK_a of the β-carboxyl group of aspartate would form a neutral complex with 1-methylimidazole. Therefore, the formation of an LBHB between a histidine and an aspartate residue in a protein would require a

special structural environment. The very close and stereoelectronically favorable contact between His57 and Asp102 might suffice. Alternatively, interactions that would increase the acid strength of Asp57 would also promote LBHB formation.

Enthalpy measurements complement the spectroscopic characterization of strong hydrogen bonding in complexes of 1-alkylimidazoles and 2,2-dichloropropionic acid.[33] The enthalpies of complexation for 1-methyl-, 1-*n*-butyl- and 1-*tert*-butylimidazole with 2,2-dichloropropionic acid are 12.3, 11.5 and 14.9 kcal mol^{-1}, respectively. These heats of complexation are in the expected range for LBHBs and are likely to represent the strength of the LBHB bridging His57-Nδ1 and Asp102-Oδ1 in the catalytic triad. If the enthalpy of the His–Asp hydrogen bond prior to formation of the tetrahedral intermediate in Fig. 2 is 5–7 kcal mol^{-1}, then the LBHB in the tetrahedral intermediate would be stronger by 7–10 kcal mol^{-1}. This would account for the decrease in enzymatic activity upon disruption of the LBHB by methylation of His57 or mutation of Asp102.[1,29,30] It would also account for the high basicity of His57 in the tetrahedral adduct and the transition state for its formation.

The results with *N*-methylimidazole and acids of varying strength complemented the analogous results of experiments with trifluoracetic acid and pyridines of varying basicity, which gave comparable results.[34,35] An elegant cryogenic NMR study of the structures of a large number of hydrogen-bonded complexes formed between amines and acids of varying basicity showed the full range of hydrogen bond types, including symmetrical hydrogen bonds.[36]

STRONG HYDROGEN BONDING IN AQUEOUS MEDIA

Until recently, strong hydrogen bonding was not regarded as compatible with an aqueous environment, which had been thought to exert a leveling effect on hydrogen bonding. However, LBHBs have been characterized in Cht and other serine proteases, acetylcholinesterase[37] and Δ^5-3-ketosteroid isomerase.[38–40] In each case, the LBHB was observed in a structural analogue of a metastable reaction intermediate, which was thought to be similar to the transition state. All of these LBHBs appeared in aqueous solutions, suggesting that an aqueous medium *per se* would not prevent LBHB formation. Actually, the existence of strong hydrogen bonds in aqueous media was anticipated by observations of the acid–base properties of proton sponges in aqueous-organic solvents.[14] It was suggested that certain dicarboxylic acids might engage in strong hydrogen bonding in aqueous solutions, and that this could in part explain the great differences in the first and second ionization constants in acids such as maleic and *cis*-caronic acids.[41]

To probe the question of possible strong hydrogen bonding in aqueous media, the ^{1}H NMR spectra of hydrogen dicarboxylates were examined in 10% aqueous acetone-d_6 (0.31 mole fraction H_2O). The experiments were carried out at low temperature (−55 °C) to depress the rate of chemical exchange with the solvent. Downfield protons (19–21 ppm) were observed in hydrogen 2,2-dimethylmalonate, hydrogen *cis*-cyclohexane-1,2-dicarboxylate, hydrogen maleate and hydrogen cyclopropane-1,1-dicarboxylate.[42,43]

The D/H fractionation factor for the internal hydrogen bond in hydrogen maleate had been shown to be low, 0.77, in water.[44] The fractionation factor for hydrogen *cis*-cyclohexane-1,2-dicarboxylate proved to be 0.69 in water and 0.52 in 10% aqueous acetone.[45]

Further characterization of the strong hydrogen bond in hydrogen maleate in aqueous acetone media under cryogenic conditions showed the activation enthalpy for its exchange with solvent protons to be $\Delta H^{\ddagger}_{ex}$ = 8 kcal mol^{-1}, and the value for hydrogen *cis*-cyclohexane-1,2-dicarboxylate was 7.3 kcal mol^{-1}.[45] The solvent exchange proved to be specifically catalyzed by maleic acid and stronger acids but not by weaker acids and not by bases, which implied the mechanism in Fig. 5. In this

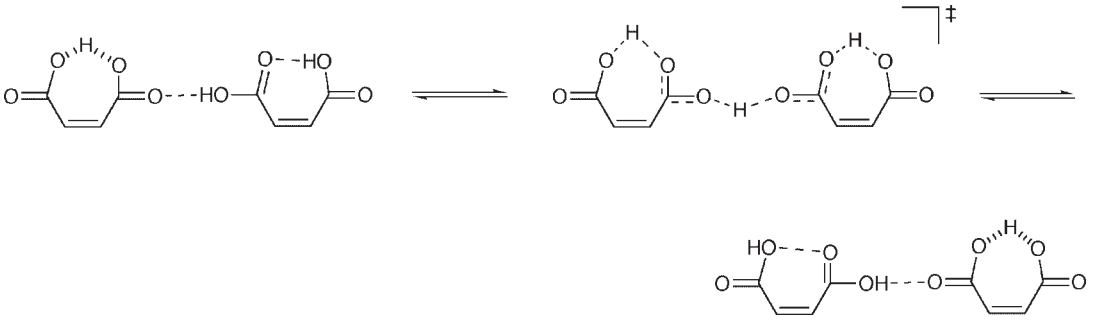


Figure 5. A mechanism for the exchange of the very strong hydrogen bond in hydrogen maleate. The exchange of the very strongly hydrogen-bonded proton in hydrogen maleate is strictly acid catalyzed, and this allows a mechanism in which the very strong hydrogen bond is partially but not completely broken in the transition state. This mechanism accounts for the high activation enthalpy (8 kcal mol^{-1}) for the exchange reaction. This value is much higher than the typical values of 1–2 kcal mol^{-1} for the exchange of weakly hydrogen bonded protons. It is not as high as the energy of the hydrogen bond itself because the hydrogen bond is never completely broken in the mechanism. In the transition state, the very strong hydrogen bond in one molecule is partially broken and a new, very strong hydrogen bond in the second molecule is partially formed

mechanism the maleic acid generated by added hydronium ion undergoes proton exchange with hydrogen maleate by a mechanism in which hydrogen bonding is retained in the transition state. While the internal, strong hydrogen bond is weakened in the transition state, it is not broken. The value of 8 kcal mol^{-1} for $\Delta H^{\ddagger}_{ex}$ represents a substantial weakening of the strong hydrogen bond in the transition state, but not a complete loss of hydrogen bonding. In view of the low activation energies typical of exchange between carboxylic acids, the values for hydrogen maleate and hydrogen *cis*-cyclohexane-1,2 carboxylate indicate strong internal hydrogen bonding in aqueous acetone solutions.

CONCLUSIONS

Predictions that strong hydrogen bonds would participate in enzymatic catalysis[46–48] have been realized in the reactions of Cht, acetylcholinesterase, and Δ^5-3-ketosteroid isomerase. Of the strong hydrogen bonds originally postulated,[46–48] a few have been excluded by physicochemical or biochemical data. Others remain for future experimental examination. The LBHB in Cht is well characterized by standard physicochemical methods in peptide-TFK complexes, which are structurally and electrostatically similar to the tetrahedral intermediates in catalysis. Similar characterizations of LBHBs in the actions of acetylcholinesterase and Δ^5-3-ketosteroid isomerase attest to an expanded role of LBHBs in enzymatic catalysis. The catalytic effects of LBHBs seem to lower the activation energy by 5–7 kcal mol^{-1}, based on available data for Cht and trypsin. The effect may be buffered by the likelihood that disruption of the LBHB results in a change in the rate-limiting step.

Acknowledgments

The author is grateful to the National Institutes of Health for Grants Nos GM30480 and DK28607. Research in his laboratory on the subject of strong hydrogen bonding was completed in collaboration with the graduate students and postdoctoral fellows John Tobin, Constance Cassidy, Jing Lin, Yaoming Wei and David Neidhart. The collaboration of his colleagues W. W. Cleland, J. L. Markley, W. M. Westler and L. A. Reinhardt is gratefully acknowledged.

REFERENCES

1. Frey PA, Whitt SA, Tobin JB. *Science* 1994; **264**: 1927–1930.
2. Robillard G, Shulman RG. *J. Mol. Biol.* 1972; **71**: 507–509.
3. Markley JL. *Biochemistry* 1978; **17**: 4648–4656.
4. Bachovchin WW. *Proc. Natl. Acad. Sci. USA* 1985; **82**: 7948–7951.
5. Robillard G, Shulman RG. *J. Mol. Biol.* 1974; **86**: 541–558.
6. Zhong S, Haghjoo K, Kettner C, Jordan F. *J. Am. Chem. Soc.* 1995; **117**: 7048–7055.
7. Imperiali B, Abeles RH. *Biochemistry* 1986; **25**: 3760–3767.
8. Liang T-C, Abeles RH. *Biochemistry* 1987; **26**: 7603–7608.
9. Brady K, Abeles RH. *Biochemistry* 1990; **29**: 7608–7617.
10. Cassidy CS, Lin J, Frey PA. *Biochemistry* 1997; **36**: 4576–4584.
11. Lin J, Cassidy CS, Frey PA. *Biochemistry* 1998; **37**: 11940–11948.
12. Pauling L. *The Nature of the Chemical Bond*. Cornell University Press: Ithaca NY, 1960.
13. Jeffrey GA. *An Introduction to Hydrogen Bonding*. Oxford University Press: New York, 1997.
14. Hibbert F, Emsley J. *Adv. Phys. Org. Chem.* 1990; **26**: 255–379.
15. Steiner Th, Saenger W. *Acta Crystallogr., Sect. B* 1994; **50**: 348.
16. McDermott A, Ridenour CF. In *Encyclopedia of NMR*, Richard JP (ed). Wiley: Chichester, 1996; 3820.
17. McAllister MA. *Can. J. Chem.* 1997; **75**: 1195–1202.
18. Markley JL, Westler WM. *Biochemistry* 1996; **35**: 11092–11097.
19. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA. *Proc. Natl. Acad. Sci. USA* 1998; **95**: 14664–14668.
20. Cassidy CS, Lin J, Frey PA. *Biochem. Biophys. Res. Commun.* 2000; **273**: 789–792.
21. Neidhart D, Wei Y, Cassidy CS, Lin J, Cleland WW, Frey PA. *Biochemistry* 2001; **40**: 2439–2447.
22. Westler WM, Frey PA, Lin J, Wemmer DE, Morimoto H, Williams PG, Markley JL. *J. Am. Chem. Soc.* 2002; **124**: 4196–4197.
23. Mildvan AS, Harris TK, Abeygunawardana C. *Methods Enzymol.* 1999; **308**: 219–245.
24. Brady K, Wei A, Ringe D, Abeles RA. *Biochemistry* 1990; **29**: 7600–7607.
25. Halkides CJ, Wu YQ, Murray CJ. *Biochemistry* 1996; **35**: 15941–15948.
26. Kuhn P, Knapp M, Soltis SM, Ganshaw G, Thoene M, Bott R. *Biochemistry* 1998; **37**: 13446–13452.
27. Katona G, Wilmouth RC, Wright PA, Berglund GI, Hajdu J, Neutz R, Schofield CJ. *J. Biol. Chem.* 2002; **277**: 21962.
28. Betzel C, Gourinath S, Kumar P, Kaur P, Perbandt M, Eschenburg S, Singh TP. *Biochemistry* 2001; **40**: 3080–3088.
29. Henderson R. *Biochem. J.* 1971; **124**: 13–18.
30. Craik CS, Roczniak S, Largeman C, Rutter WJ. *Science* 1987; **237**: 909–913.
31. Tobin JB, Whitt SA, Cassidy CS, Frey PA. *Biochemistry* 1995; **34**: 6919–6924.
32. Cassidy CS, Reinhardt LA, Cleland WW, Frey PA. *J. Chem. Soc., Perkin Trans. 2* 1999; 635.
33. Reinhardt LA, Sacksteder KA, Cleland WW. *J. Am. Chem. Soc.* 1998; **120**: 13366.
34. Dega-Szafran Z, Dulewicz E. *J. Chem. Soc., Perkin Trans. 2* 1983; 345.
35. Barczynski P, Dega-Szafran Z, Szafran M. *J. Chem. Soc., Perkin Trans. 2* 1985; 765.
36. Smirnov SN, Golubev NS, Denisov GS, Benedict H, Schah-Mohammedi P, Limbach HH. *J. Am. Chem. Soc.* 1996; **118**: 4094.
37. Massiah MA, Viragh C, Reddy PM, Kovach IM, Johnson J, Rosenberry TL, Mildvan AS. *Biochemistry* 2001; **40**: 5682–5690.
38. Zhao Q, Abeygunawardana C, Talalay P, Mildvan AS. *Proc. Natl. Acad. Sci. USA* 1996; **93**: 8220–8224.
39. Cho HS, Choi G, Choi KY, Oh BH. *Biochemistry* 1998; **37**: 8325–8330.
40. Choi C, Ha NC, Kim SW, Kim DH, Park S, Oh BH, Choi KY. *Biochemistry* 2000; **39**: 903–909.
41. Frey PA, Cleland WW. *Bioorg. Chem.* 1998; **26**: 175.
42. Cassidy CS, Lin J, Tobin JB, Frey PA. *Bioorg. Chem.* 1998; **26**: 213.
43. Hess RA, Reinhardt LA. *J. Am. Chem. Soc.* 1999; **121**: 9867.
44. Jarrett RM, Saunders M. *J. Am. Chem. Soc.* 1985; **107**: 2648.
45. Lin J, Frey PA. *J. Am. Chem. Soc.* 2000; **122**: 11258.
46. Cleland WW. *Biochemistry* 1992; **31**: 317.
47. Gerlt JA, Gassman PG. *Biochemistry* 1993; **32**: 11943.
48. Cleland WW, Kreevoy MM. *Science* 1994; **264**: 1887.