PCCP

Newly Identified C–H···O Hydrogen Bond in Histidine

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-05-2022-002048.R1
Article Type:	Paper
Date Submitted by the Author:	07-Jul-2022
Complete List of Authors:	Steinert, Ryan; Wichita State University Fairmount College of Liberal Arts and Sciences, Department of Chemistry and Biochemistry Kasireddy, Chandana; Wichita State University Fairmount College of Liberal Arts and Sciences, Department of Chemistry and Biochemistry Heikes, Micah; Wichita State University Fairmount College of Liberal Arts and Sciences, Department of Chemistry and Biochemistry Mitchell-Koch, Katie; Wichita State University Fairmount College of Liberal Arts and Sciences, Department of Chemistry and Biochemistry

Newly Identified C–H···O Hydrogen Bond in Histidine

Ryan M. Steinert[1], Chandana Kasireddy[2], Micah E. Heikes[3], Katie R. Mitchell-Koch[4*]

[1]Department of Chemistry and Biochemistry, Wichita State University, 1845 Fairmount Street, Wichita, KS 67260-0051

[2]Department of Chemistry and Biochemistry, Wichita State University, 1845 Fairmount Street, Wichita, KS 67260-0051. Current Address: 1973 Kings Knot Ct, Apex, NC, 27519.

[3]Department of Chemistry and Biochemistry, Wichita State University, 1845 Fairmount Street, Wichita, KS 67260-0051

[4]Department of Chemistry and Biochemistry, Wichita State University, 1845 Fairmount Street, Wichita, KS 67260-0051; *Corresponding author email address: katie.mitchell-koch@wichita.edu

1. ABSTRACT

New Cδ–H···O histidine hydrogen bonding interactions in various proteins are identified by neutron diffraction and computationally characterized. Neutron diffraction data shows several H-bond motifs with the Cδ–H moiety in histidine side chains, including interactions in β-sheets and with coordinated waters, mostly with histidinium and τ-tautomers. In yellow protein, an active site histidine H-bonds via Cδ–H to a main chain carbonyl while the Cε–H bond coordinates a water molecule. Although the H-bonding ability of Cε–H bonds in histidine have been previously identified, analysis of neutron diffraction structures reveals Cε–H H-bonds in notable active site interactions: for the proximal histidine in myoglobin; a zinc-bound histidine in human carbonic anhydrase II; within the Ser-Asp-His catalytic triad of the trypsin active site; and a histidine in the proton shuttle mechanism of RNase A, in addition to more general roles of coordinating water and forming H-bonds with carbonyl groups in β-sheets within a number of proteins. Properties of these H-bonds were computationally investigated using 5-methylimidazole and 5-methylimidazolium as models for histidine and histidinium. The π- and τ-tautomeric states of 5-methylimidazole were investigated, as both histidine tautomers are observed in the crystal structures. The newly characterized Cε–H···O and Cδ–H···O model complexes with water and acetone meet the overwhelming majority of IUPAC H-bonding criteria. 5-methylimidazolium forms complexes that are nearly twice as strong as the respective neutral τ-5-methylimidazole and π-5-methylimidazole complexes. While the τ- and π-tautomers form Cε–H···O complexes of similar strength, the τ-Cδ–H···O interaction is approximately twice as strong as the π-Cδ–H···O interaction. Calculated charges on C–H (and N–H) hydrogens not participating in the H-bond are only slightly perturbed upon complex formation, implying that formation of one H-bond does not diminish the molecule's capacity for further H-bond formation at other sites in the imidazole ring. Overall, findings indicate that the Cδ–H···O interaction may be important for β-sheet stability, conformation, interactions with solvent, and mechanisms in the active site. Recognition of C–H bond polarity and hydrogen bonding ability in histidine may improve molecular modeling and provide further insight into the diverse roles of histidine in protein structure-function-dynamics.

2. INTRODUCTION

The C–H···X hydrogen bonding interaction, where X is most commonly oxygen, often plays an important role in the structure and function of biomolecules. In nucleic acids, the C6–H of pyrimidines and C8–H of purines can interact with oxygen of the phosphate group in single nucleotides, forming an intramolecular H-bond.[1,2] The backbone-backbone axial C2'(H)···O4' and C5'(H)···O2 interactions between two nucleotides that can lead to 7-membered ring formation in RNA is a consequence of hydrogen bonding.[3] In the β-sheet structures of proteins, carbonyl groups along the main chain can form hydrogen bonding interactions with Cα–H donors of an adjacent chain.[4, 5] For a more in-depth survey of C–H···X interactions in biological systems, including roles in structure, molecular recognition, and enzyme catalysis, the excellent review by Horowitz and Trievel is recommended.[6] In general, hydrogen bonds in proteins are directional interactions that play a key part in governing protein folding and structure. The strength and directionality of H-bonding interactions are also important in molecular recognition.[7]

Histidine is a chemically versatile amino acid, with a variety of hydrogen bonding motifs, that is present in ~50% of enzyme active sites.[8] Within protein environments, it can exist as the τ-tautomer, π-tautomer, or be protonated as histidinium (the free amino acid side chain has an aqueous pKa ~ 6.0). Among the side chains, it offers a wealth of different hydrogen bond motifs. In particular, the protonated histidine residue, an imidazolium species, offers a rich arrangement of potential hydrogen bond donor sites, e.g., the Nε1–H (protonated in the τ–tautomer) and Nδ2–H (protonated in the π–tautomer) H sites, and the Cε–H and Cδ–H sites (see Figure 1). Furthermore, our previous work indicated from calculated charges on the histidine ring that the polarized ring of imidazolium even contains polarized C–H bonds in the τ-tautomer and protonated form, leading to this inquiry on the ability of the C–H groups to hydrogen bond.[9] Buncel *et al*. have presented data showing that the hydrogen/tritium exchange rate constant at the Cε–H site at 85°C is over two orders of magnitude larger for histidinium than for neutral τ-histidine.[10] It follows that hydrogen bonding at Cε–H is enhanced in histidinium over the neutral τ-histidine species, as we will see below.

Data from Nanda and Schmiedekamp suggests, based on structural database studies, that C–H···X interactions involving a neutral histidine residue are not statistically biased toward achieving a linear configuration.[11] It might in fact be that Cε–H···X and Cδ–H···X interactions in neutral histidine energetically

prefer linearity, but by and large are unable to overcome protein packing forces to direct a linear H-bonding interaction. In fact, this study will show that unconstrained geometric optimizations of Cε–H···O and Cδ–H···O complexes of 5-methylimidazole do tend toward such a linear configuration. It should be noted that data from Nanda and Schmiedekamp revealed that all N–H···O interactions involving nitrogen of histidine (and histidinium) showed a propensity for the H-bond to dictate the geometry of the atoms involved in the hydrogen bond.[11] In our previous studies of fluorohistidine,[9 12 13 14] evaluation of charges calculated by electrostatic potential indicated that the τ-tautomer and protonated (histidinium) forms of histidine have a quite polar Cδ–H bond, which prompted this investigation.

The focus of the current study is twofold. First, protein structures from neutron diffraction data were scrutinized to search for histidine residues involved in potential C–H···X hydrogen bonding interactions. Through neutron diffraction, it is easy to differentiate the tautomeric and protonation states of histidine side chains, as this technique allows acquisition of well-defined positions of all the atoms, including hydrogen. Secondly, a computational study involving implicitly solvated 5-methylimidazole (both tautomers and 5-methylimidazolium) and its various complexes with water and acetone (see Figure 4) help to elucidate the hydrogen bonding capacity of the different tautomeric states of histidine; specifically, this study seeks to examine whether or not the tautomeric form of the imidazole species impacts its hydrogen bonding potential at the Cε–H and Cδ–H sites. Experimental ^{13}C NMR evidence found that the tautomeric ratio (τ : π) for the imidazole ring in basic solution at 32 ± 3°C is approximately 4 : 1,[15] and so it would seem quite relevant to seek to investigate the hydrogen bonding potential of each tautomeric state.

5-methylimidazole and 5-methylimidazolium were used as model compounds for histidine and histidinium, respectively. H-bonding complexes of these model compounds with water and acetone are investigated below, with acetone being used to model interaction of the C–H hydrogen bond donor group with a C=O hydrogen bond acceptor group. The numbering scheme for 5-methylimidazole ring atoms is shown in Figure 1, with ring numbering beginning at the nitrogen atom adjacent to the carbon bearing the methyl substituent. Here, the number following the atomic symbol simply represents the ring position of each atom. The complexes of interest to this work involve the C2–H (Cε–H) and C4–H (Cδ–H) groups acting as H-bond donors.

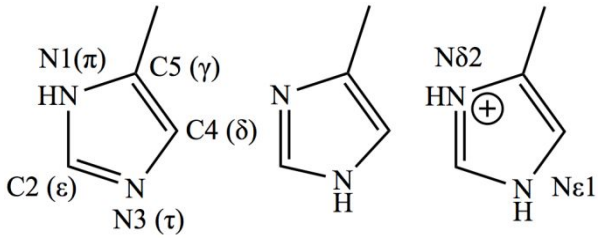

Figure 1. Left to Right: π-5-methylimidazole, τ-5-methylimidazole, 5-methylimidazolium. The ring numbering scheme shown for the π-tautomer is the same for the τ and the imidazolium species. Alternate nomenclature for the nitrogen atoms is indicated in the imidazolium structure.

3. METHODS

Structural Data. The protein database was searched for all available neutron diffraction structures, so that proton positions and histidine tautomeric form/protonation state could be determined. Given the relative paucity of neutron diffraction structures, the histidine environments were evaluated manually. Thus, all solved neutron diffraction structures containing histidine, and having a resolution cutoff of ≤ 2.0 Å, were examined as follows: 1.) In PyMOL (The PyMOL Molecular Graphics System, Versions 2.2 and 2.5, Schrödinger, LLC.), the backbone structure was visualized. 2.) All residues and water molecules within 5 Å of any histidine were selected and highlighted with all atoms visualized. 3.) Each histidine was visually inspected. The tautomeric/protonation state was determined. 4.) Atoms in proximity to C–H bonds were inspected, and if they qualified as hydrogen-bond acceptors (O, N on side chains, O in crystallographic water), the C–H···X distance and angle were determined using PyMOL tools (precision of 0.1 Å measurements). For this work, three researchers independently evaluated the protein structure in Table 1, for validity and reproducibility of the manual data analysis.

Electronic Structure Calculations. Structural optimizations were performed by Gaussian 09 (G09)[16] after preparing initial geometries with GaussView 5.0.[17] The G09 default for closed shell systems of restricted (as opposed to unrestricted or restricted open-shell) was used, along with the default "fine" integration grid

for DFT calculations. The G09 default SCF=Tight was used as convergence criterion for all calculations. The textual format for a given method will be written as "level of theory/basis set"; for example, MP2/aug-cc-pVTZ. The two DFT functionals used for optimizations were the global-hybrid meta-GGA functional M06-2X[18] and the hybrid GGA functional BHandHLYP.[19] Unless otherwise noted, all calculations were performed using the conductor-like polarizable continuum (C-PCM) implicit solvation model,[20,21] with water as the solvent (ε = 78.36). Gas phase calculations were used for $H_3N{:}\rightarrow F_{(1)}{-}F_{(2)}$ and $H_2O{:}\rightarrow Cl_{(1)}{-}Cl_{(2)}$ adduct analysis, and also NMR calculations at the PBE0[22]/cc-pVTZ level. All geometry optimizations were unconstrained. Frequency calculations were performed in tandem with optimizations to confirm that a given stationary point on the potential energy surface of a molecule or complex was an equilibrium geometry; only optimized geometries with all real frequencies are presented herein. Natural population analysis (NPA) calculations[23] and Hirshfeld population analysis[24] calculations were carried out within the G09 set of codes. For calculation of isotropic shielding values for NMR analysis, Gauge-Including Atomic Orbital (GIAO)[25] and Continuous Set of Gauge Transformations (CSGT)[26] calculations were performed with G09 on structures optimized at the BHandHLYP/6-31+G(d,p) level. For the bond critical point (bcp) analysis, MP2/aug-cc-pVTZ single point energy calculations were carried out on BHandHLYP/6-31+G(d,p) optimized geometries. Checkpoint files from the single point calculations were used to generate density cube files using the G09 cubegen utility. The density cube files were then analyzed for bond critical points using AIM-UC v1.6.4 software.[27] For analysis of electrostatic potential surfaces, MP2/aug-cc-pVTZ single point energy calculations were performed on BHandHLYP/6-31+G(d,p) optimized geometries, density and potential cube files were generated in the G09 cubegen utility, and the electrostatic potential was mapped onto the 0.002 au isodensity surface in GaussView 5.0. For more details and figures pertaining to the electrostatic potential analysis, see S12-S17 in supporting information.

The concern of basis set superposition error (BSSE) and basis set incompleteness error (BSIE) was considered. Many of the calculations in this paper are based on double-zeta 6-31+G(d,p) geometries without using a counterpoise correction. To investigate the potential for significant geometric errors, the geometries of the water and acetone complexes were also optimized at the BHandHLYP/6-311+G(d,p) level. A basis set of at least triple-zeta quality is recommended to avoid BSSE and reduce other BSIE errors.[28] It was found that the optimized geometries from BHandHLYP/6-31+G(d,p) and BHandHLYP/6-311+G(d,p) methods were remarkably similar; full results are shown in Table S.5. As such, 6-31+G(d,p) geometries, and other calculations based on these geometries, should be satisfactory.

4. RESULTS

X-RAY AND NEUTRON DIFFRACTION DATA

Twelve neutron diffraction structures were found to contain hydrogen bonds containing C–H groups of histidine as hydrogen bond donor, with even the possibility of multiple C–H bonds within a single histidine. The results are presented in Table 1, which provides the protonation state/tautomeric form of each histidine and describes the H-bond geometry and donor/acceptor pair. Hydrogen bonds between both Cε (C2) and Cδ (C4) were observed between histidine side chains and oxygen-containing functionalities. As can be seen, the hydrogen bonds satisfy geometric criteria of linearity/near-linearity and most heavy atom distances at or below 3.3 Å.

Hydrogen bonding between histidine C–H groups and main chain (backbone) carbonyl groups were observed in structures of D-xylose isomerase (Cδ–H in π-tautomer His71; PDB ID 4QDW); in a signaling protein prototype (Cε–H in τ-tautomer of His166; PDB ID 4RSG); human heart fatty-acid binding protein (Cε–H in τ-tautomer of His 119; PDB ID 5CE4, shown in Figure 2a); photoactive yellow protein (PYP) (Cδ–H in protonated His108; PDB ID 2ZOI); HIV-1 protease (Cδ–H in protonated His69; PDB ID 2ZYE); PYP mutant E46Q (Cδ–H in protonated His 108; PDB ID 5GX9); and class A β-lactamase (Cδ–H in τ-tautomer His 112; PDB ID 5A90). Also observed is an intra-helix C–H to backbone hydrogen bond that appears to position the side chain of His71 in D-xylose isomerase so that its N–H group can coordinate water within a solvent hydrogen bond network, as seen in Figure 2b. An intra-helix H-bond with a Cδ–H group is also seen in β-lactamase (Figure 2c). This hydrogen bond may act as an anchor for the side chain, inhibiting rotation, for this residue is shown coordinating three water molecules that are in turn part of a broader hydration network. Thus, both Cδ–H and Cε–H bonds are seen to interact with the peptide backbone in β-sheets and α-helices (viewed in Figure 2a-c), ostensibly contributing to the stabilization of the secondary structures of proteins.

In the active site of trypsin, within the Ser-Asp-His catalytic triad, Cε–H in protonated/cationic His57 is observed to hydrogen bond to the main chain carbonyl group of Ser214. Although it is Ser195 that is a member of the catalytic triad in trypsin, the interaction between His57 and Ser214 is demonstrated in an *ab initio* study to be important in the transition state, contributing to the interaction energy.[29] The same *ab initio* study indicates through mutant studies that together Asp102 and Ser195 position the His57 ring into a catalytically competent state; when the His57 ring is in a flipped conformation, there is an additional energetic penalty for catalysis. Further, the Ser214 residue is implicated in stabilizing the ionic intermediate state in which His57 is cationic. Our structural analysis shows a heavy atom distance of 3.1 Å between Cε of His57 and the carbonyl oxygen of Ser195. It is remarkable that the *ab initio* simulations find this distance to be 3.065 Å during the tetrahedral intermediate transition via single proton transfer mechanism, lending further credence to the importance of the Cε–H hydrogen bond within the critical His57-Ser195 interaction.

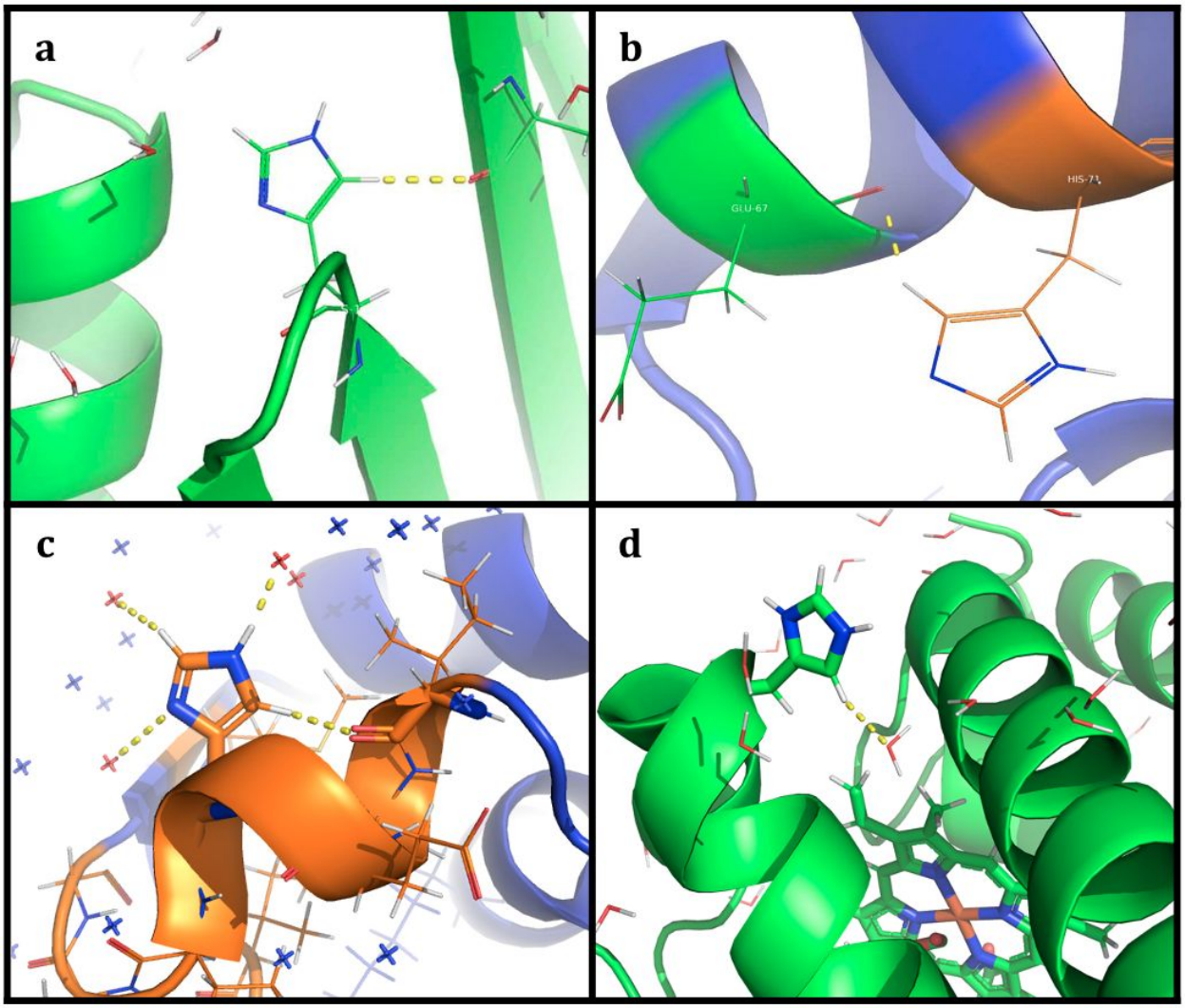

Figure 2. Cδ–H hydrogen bonds in histidine observed in neutron diffraction structures**. a.** His119 in human heart fatty-acid binding protein, PDB ID 5CE4: τ-tautomer Cδ–H···O=C **b.** His71 in D-xylose isomerase, PDB ID 4QDW: π-tautomer Cδ–H···O=C **c.** His112 in β-lactamase, PDB ID 5A90: τ-tautomer Cδ–H···O=C **d.** $His^{+}36$ in myoglobin, PDB ID 2MB5: $histidinium^{+}$ Cδ–H···O_{water}. H-bonds designated by yellow dashed lines. Atom coloring scheme: carbon (orange, green), hydrogen (white), nitrogen (blue), oxygen (red).

In addition to interactions with carbonyl groups, another common motif in the hydrogen bond data in Table 1 is hydrogen bonds with water molecules. Hydrogen bonds between histidine C–H groups and water oxygen atoms are seen twice in met- (oxidized) myoglobin (Cε–H in π–tautomer His 24 and τ-tautomer His119; PDB ID 1L2K); carbonmonoxide-bound myoglobin (Cδ–H in protonated His36; PDB ID 2MB5; shown

in Figure 2d); and photoactive yellow protein (PYP) mutant E46Q (Cε–H in protonated His108; PDB ID 5GX9). Note that in this last example, the structure of PYP also shows that the same histidine, His108, interacts with a backbone carbonyl through the Cδ–H bond! In fact, Figure 3d shows that His$^+$108 of PYP is capable of engaging in four hydrogen bonds. The Cδ–H hydrogen bonding interaction is conserved in the wild-type protein, too. As shown below in the electronic structure calculations, protonated/cationic histidine is the most highly polarized form of histidine, and the most strongly interacting in hydrogen bonds, primarily due to its charge. In bovine pancreatic ribonuclease A (PDB ID 3A1R), water is also observed to be coordinated through multiple hydrogen bonds of protonated His12 as shown in Figure 3a. In this structure, the histidine is fully protonated and coordinating water molecules at the N3, Cδ–H, and Cε–H positions. This histidine residue is part of the proton transfer process, in the highly-studied enzymatic mechanism.[30 31] Although solution state tautomeric form has not been definitively established, experimental and theoretical studies suggest His12 of RNase A is in the π-tautomer prior to accepting a proton (and protonated afterward, as seen in this structure).[13 32]Hydrogen bonds help stabilize changes in charge.

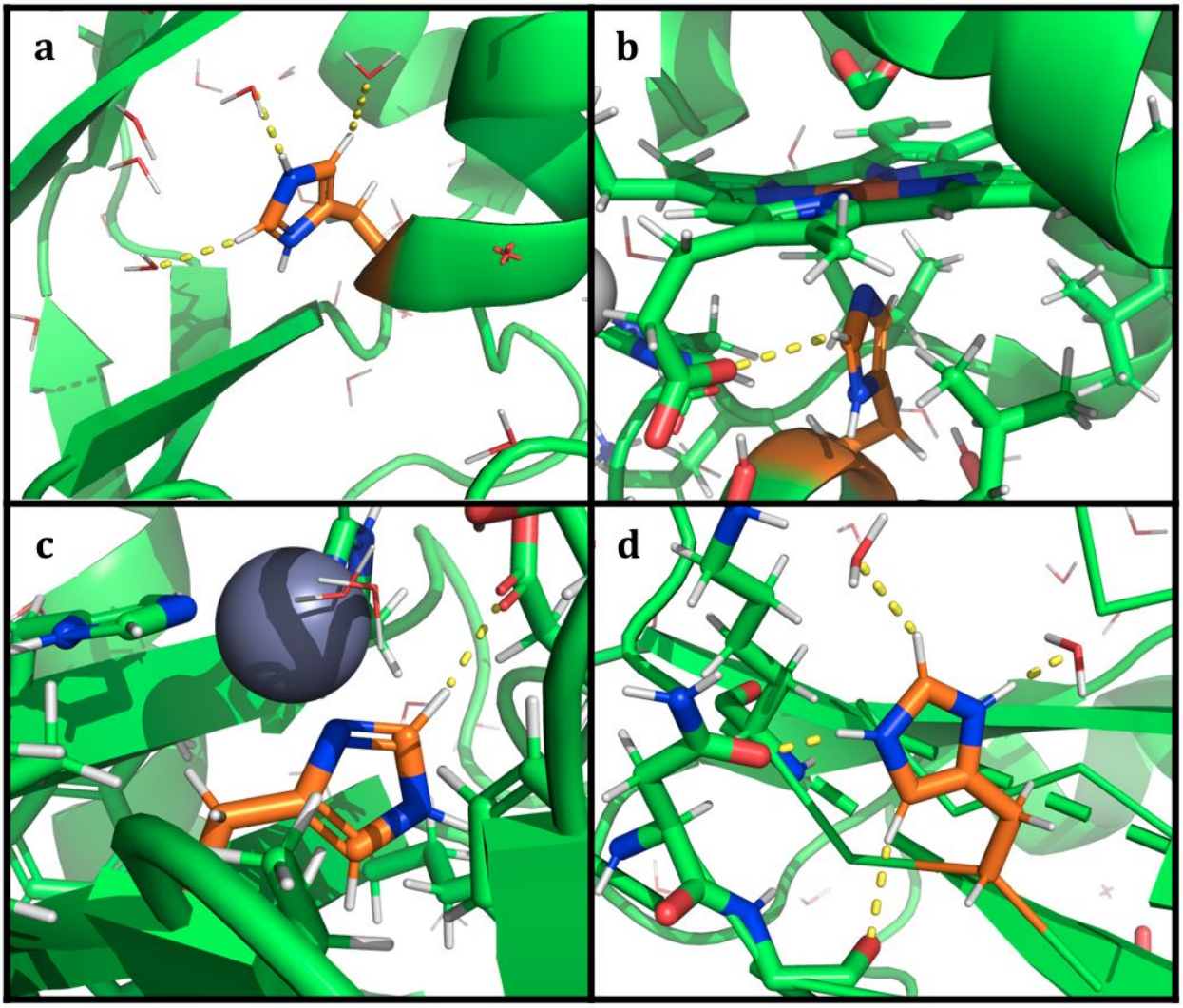

Figure 3. Histidine C–H hydrogen bonds in neutron diffractions structures. **a.** His$^+$12 in RNase A, PDB ID 3A1R has both Cδ–H and Cε–H hydrogen bonded to water molecules. **b.** His93 in the π-tautomer, which is the proximal histidine in myoglobin, PDB ID 2MB5, has a hydrogen bond between Cε–H and a carboxylate group on the heme. **c.** His119 in the τ-tautomer in human carbonic anhydrase II, PDB ID 4Q49, has a hydrogen bond between Cε–H and the carboxylate of Glu106 in the active site. **d.** His$^+$108 in photoactive yellow protein, PDB ID 5GX9 has a Cδ–H hydrogen bond to a backbone carbonyl and a Cε–H hydrogen bond to water; note that all four X–H sites of the imidazole ring are involved in hydrogen bonds. H-bonds designated by yellow dashed lines. Atom coloring scheme: carbon (orange, green), hydrogen (white), nitrogen (blue), oxygen (red).

For two active site histidines coordinated to metals, C–H hydrogen bonds are observed. In the *apo* form of human carbonic anhydrase II (HCAII) (PDB ID 4Q49), there are 3 histidines coordinated to zinc. As seen in Figure 3c, one of them, His119, has a Cε–H bond to the carboxylate of Glu106. This glutamate residue has been shown to control interactions around the zinc ion, positioning zinc-bound hydroxide in an optimal geometry for attack on the substrate CO_2. The present results suggest that Cε–H of His119 may also contribute to the intricate hydrogen bond network in human carbonic anhydrase that mediates the water network and active site architecture.[33]

Meanwhile, in carbonmonoxy-myoglobin structure 2MB5, the active site residue His93 is the proximal histidine directly bound to the heme iron.[34] It has been noted previously that metal-bound histidine side chains are potent hydrogen bond donors.[35 36 37] The His93 is hydrogen bonded to a carboxylate group on the heme via the Cε–H bond, as shown in Figure 3b. The heme is known to undergo a conformational change in which the Fe–N bond distance between His93 and heme adjusts after release of CO/O_2.[38]

Table 1. *Data showing Cε and Cδ carbon in hydrogen bonding interactions with electronegative atom captured in protein crystal structures. the C to O distance, represented by O - - - C, is the Cδ/Cε to O_{water} or Cδ/Cε to $O_{carbonyl}$ distance. Values not in parenthesis are from crystallographic data, and values in parenthesis are data from the corresponding model complexes optimized at the BHandHLYP/6-31+G(d,p) level (see Table 2a). Percent difference values between crystal structure and model complex data are listed. The average and standard deviation of the percent differences are given at the bottom of the table.*

PDB ID	Protein description	Residue	His type	H-bond donor	H-bond acceptor	H-bond distance (Å)	O - - - C distance (Å)	H-bond angle (degrees)
4QDW[39]	D-xylose isomerase	His71	π	Cδ–H	C=O of backbone in α-helix	2.1 (2.46) 16%	3.1 (3.54) 13%	147 (179) 20%
1L2K[40]	Metmyoglobin	His24	π	Cε–H	oxygen of water	2.2 (2.27) 3.1%	3.3 (3.34) 1.2%	164 (179) 8.7%
2MB5[41]	Carbonmonoxy myoglobin	His93	π	Cε–H	COO^- on heme	2.6 (2.30) 12%	3.4 (3.38) 0.59%	130 (176) 30%
4Q49[42]	Human carbonic anhydrase II	His119	τ	Cε–H	COO^- of Glu106 in active site	2.2 (2.30) 4.4%	3.2 (3.38) 5.5%	165 (177) 7.0%
1L2K[40]	Metmyoglobin	His119	τ	Cε–H	oxygen of water	2.3 (2.26) 1.8%	3.3 (3.34) 1.2%	160 (179) 11%
4RSG[43]	RAS GTPase	His166	τ	Cε–H	C=O of backbone	2.3 (2.30) 0%	3.2 (3.38) 5.5%	146 (177) 19%
5A90[44]	Class A β-lactamase	His112	τ	Cδ–H	C=O of backbone in α-helix	2.4 (2.34) 2.5%	3.3 (3.41) 3.3%	146 (178) 20%
5A90[41]	Class A β-lactamase	His112	τ	Cε–H	oxygen of water	2.8 (2.26) 21%	3.6 (3.34) 7.5%	~180 (179) 0.56%
5CE4[45]	Human heart fatty-acid binding protein	His119	τ	Cδ–H	C=O of backbone on β-loop/turn	2.4 (2.34) 2.5%	3.3 (3.41) 3.3%	162 (178) 9.4%
2ZOI[46]	Wild-type PYP	His108	His^+	Cδ–H	C=O of backbone in β-sheet	2.3 (2.17) 5.8%	3.4 (3.24) 4.8%	160 (170) 6.1%
2ZYE[47]	HIV-1 Protease	His69	His^+	Cδ–H	C=O of backbone on β-loop/turn	2.3 (2.17) 5.8%	3.2 (3.24) 1.2%	160 (170) 6.1%
3A1R[48]	Bovine pancreatic ribonuclease A	His12	His^+	Cε–H	oxygen of water	2.3 (2.03) 12%	3.2 (3.11) 2.9%	133 (179) 29%
3A1R[48]	Bovine pancreatic ribonuclease A	His12	His^+	Cδ–H	oxygen of water	2.6 (2.14) 19%	3.4 (3.22) 5.4%	116 (176) 41%

5GX9[49]	Mutant (E46Q) PYP	His108	His^+	Cε–H	oxygen of water	2.3 (2.03) 12%	3.2 (3.11) 2.9%	145 (179) 21%
5GX9[49]	Mutant (E46Q) PYP	His108	His^+	Cδ–H	C=O of backbone on β-sheet	2.3 (2.17) 5.8%	3.3 (3.24) 1.8%	163 (170) 4.2%
5MNX[50]	Trypsin	His57	His^+	Cε–H	C=O of backbone	2.4 (2.04) 16%	3.1 (3.12) 0.64%	122 (175) 36%
2MB5[38]	Carbonmonoxy myoglobin	His36	His^+	Cδ–H	oxygen of water	2.5 (2.14) 16%	3.4 (3.22) 5.4%	145 (176) 19%
					Average ± standard deviation	(9 ± 7)%	(4 ± 3)%	(17 ± 12)%

Examples of Cδ–H···O interactions in protein crystal structures are shown in Figure 2. While Cε–H···O have been more thoroughly investigated in previous work, these present results show that Cδ–H···O_w and Cδ–H···O=C interactions with backbone amides appear to be true hydrogen bonding interactions, as seen by the geometries reported in Table 1. In addition to the structural characterizations of C–H hydrogen bonds in the neutron diffraction data, the hydrogen bond lengths, geometries, and heavy atom distances for model complexes are provided in parentheses in the last three columns of Table 1, along with the percent difference. Although protein structures impose steric constraints on hydrogen bond geometries, it can be seen that Cδ–H interactions have the possibility to undertake energetically-preferred near-linear configurations in protein structures, with possibilities of both shorter and longer heavy atom distances than are found in the model complexes that are fully characterized in the subsequent section.

Hydrogen bonds are observed for all protonation states of histidine, but are found in this data set to be most common for histidinium and least common for the π-tautomer. For Cε–H···O interactions, the τ-tautomeric state was most common, with histidinium the next most prevalent. The Cδ–H bond is sterically hindered relative to the previously identified histidine Cε–H hydrogen bond donor. However, this work suggests that the Cδ–H···O interaction, when sterically feasible, should be considered an H-bonding interaction, and researchers should be aware of it when analyzing protein structures to understand stability and structure-function relationships. Interaction energies and molecular properties characterized by electronic structure calculations in the sections below suggest that the C–H···O interactions can be considered to play an important role in protein folding and stabilization.

Computational Analysis

H-bonding interactions involving X–H···Y–Z may be identified through a combination of geometric parameters, energetic data, electron density data, and IR and NMR data. For characterization of H-bonding interactions in the 5-methylimidazole···water and 5-methylimidazole···acetone complexes displayed in Figure 4, IUPAC criteria[51] for hydrogen bond identification will be utilized. Koch and Popelier[52] have previously proposed eight conditions to serve as necessary criteria for the existence of H-bonding, a few of which (bond critical point analysis and loss of charge of the hydrogen atom) will also be discussed below. Three forms of 5-methylimidazole are examined: the τ (tau) tautomer, the π (pi) tautomer, and the imidazolium ion in which both ring nitrogen atoms are protonated.

Geometric data provides evidence for H-bonding interactions. IUPAC notes (for a Z–Y:→H–X hydrogen bonding interaction) that "historically, the X to Y distance was found to be less than the sum of the van der Waals radii of X and Y and this shortening of the distance was taken as an infallible indicator of hydrogen bonding. This empirical observation is true only for strong hydrogen bonds. This criterion is not recommended."[51] However, for the sake of curiosity, this van der Waals concept applied to H-bonding is explored. In the present work, Z–Y:→H–X may be written as Z–O:→H–$C_{imidazole}$, and the C to O distance is the $C_{imidazole}$ to O_{water} or $C_{imidazole}$ to $O_{acetone}$ distance, represented by O - - - C in Table 2a.

tau-5-methylimidazole /C4(δ)–H···water **pi-5-methylimidazole /C4(δ)–H···water** **5-methylimidazolium /C4(δ)–H···water**

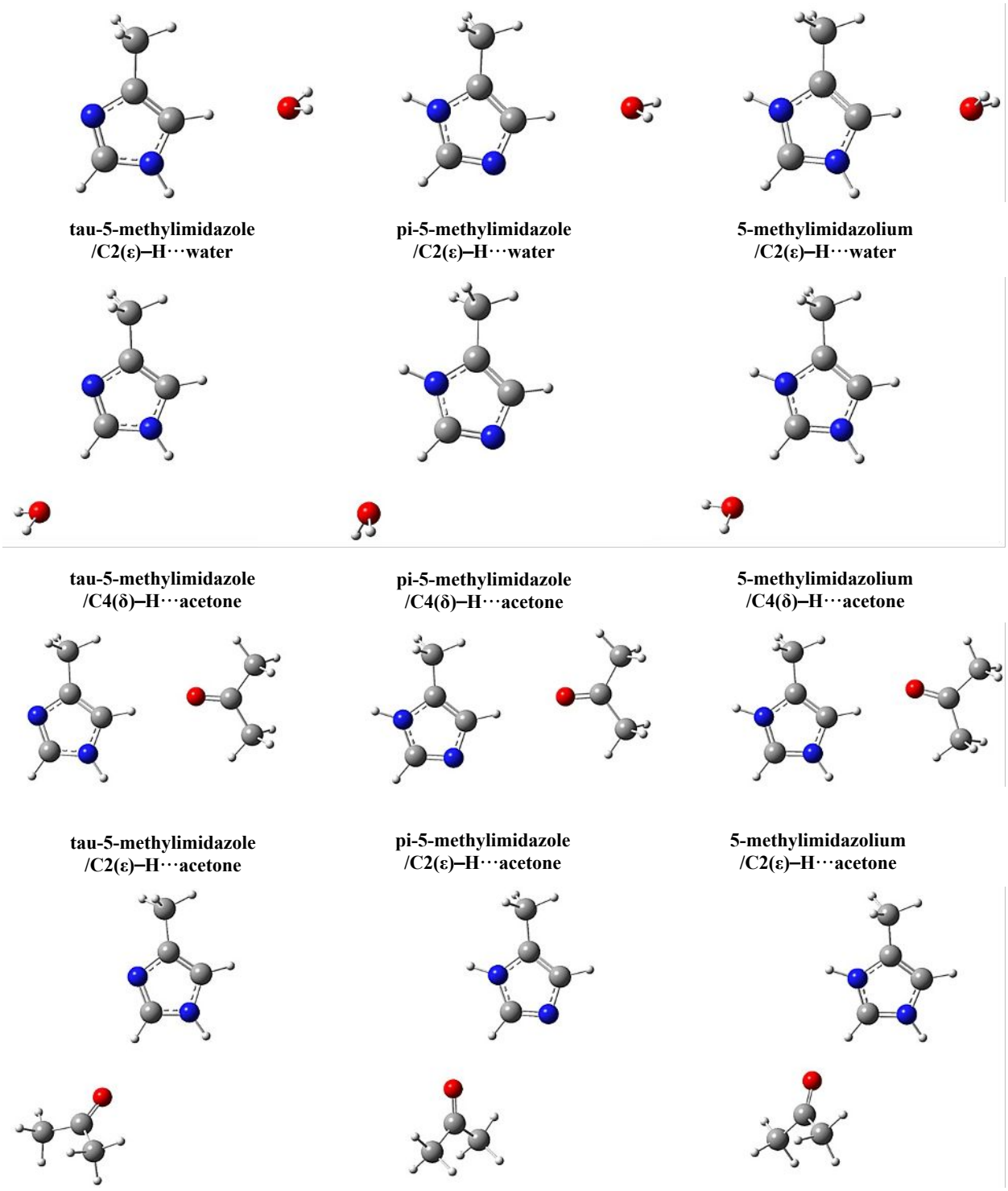

Figure 4. BHandHLYP/6-31+G(d,p) optimized geometries of the water and acetone complexes. Atoms are displayed with covalent radii scaled to 50%. Atom coloring scheme: carbon (grey), hydrogen (white), nitrogen (blue), oxygen (red). Visual images captured from GaussView 5.0 software. Note that the imidazolium species carry a +1 charge.

Analysis of over five million interatomic "non-bonded" distances from 603,297 crystals within the Cambridge Structural Database (CSD) led Alvarez to propose a consistent set of van der Waals radii for most naturally occurring elements.[53] Reported values therein for carbon and oxygen are 1.77 and 1.50 Å respectively. In the neutron diffraction structure analysis in Table 1, 71% of listings have C to O lengths ≤ 3.3

Å. The optimized O - - - C lengths of the model complexes match well with crystallographic data, with an average percent difference of only (4 ± 3)% across all 17 listings in Table 1. H-bond angle data (for the C–H···O geometric arrangement) is also reported in Table 2a. IUPAC notes that "the X–H···Y hydrogen bond angle tends toward 180˚ and should preferably be above 110˚."[51] All optimized complexes have C–H···O angles well above this 110˚ threshold, as do all the experimental listings in Table 1. The larger average percent difference between model complexes and crystallographic data for bond angles (17 ± 12)% is not surprising. For unconstrained optimizations, the angle clearly tends toward 180˚. However, when crystal packing forces are a factor, Nanda and Schmiedekamp[11] note that C–H···X interactions in neutral histidine are not statistically biased toward a linear configuration. According to the crystallographic data in Table 1, this observation appears to be true for C–H···X interactions of both histidine and histidinium, with many experimental angle listings well below 180˚, but above the 110˚ IUPAC value. According to crystallographic data in Table 1, the Cε–H···X and Cδ–H···X interactions in neutral histidine and histidinium show a similar propensity toward linearity.

Table 2a. *Geometric data for the model complexes. H-bond length (Å), O - - - C distance (Å), C–H···O angle (degrees).*

Complex Name	BHandHLYP/6-31+G(d)	BHandHLYP/6-31+G(d,p)	M06-2X/6-31+G(d,p)	MP2/6-31+G(d)	MP2/aug-cc-pVTZ
tau-5-methylimidazole/C4-H···water	2.30, 3.37, 179	2.30, 3.37, 176	2.24, 3.31, 172	2.25, 3.33, 179	2.30, 3.37, 172
tau-5-methylimidazole/C2-H···water	2.26, 3.34, 178	2.26, 3.34, 179	2.23, 3.31, 178	2.24, 3.32, 179	2.29, 3.37, 178
pi-5-methylimidazole/C4-H···water	2.39, 3.46, 178	2.39, 3.47, 178	2.35, 3.39, 161	2.33, 3.42, 179	2.43, 3.46, 160
pi-5-methylimidazole/C2-H···water	2.26, 3.34, 179	2.27, 3.34, 179	2.24, 3.32, 180	2.24, 3.33, 179	2.29, 3.37, 179
5-methylimidazolium/C4-H···water	2.14, 3.21, 176	2.14, 3.22, 176	2.12, 3.20, 180	2.12, 3.21, 179	2.14, 3.22, 175
5-methylimidazolium/C2-H···water	2.03, 3.10, 179	2.03, 3.11, 179	2.03, 3.11, 179	2.05, 3.13, 178	2.06, 3.14, 178
tau-5-methylimidazole/C4-H···acetone	2.35, 3.41, 173	2.34, 3.41, 178	2.44, 3.33, 139	*	*
tau-5-methylimidazole/C2-H···acetone	2.31, 3.38, 174	2.30, 3.38, 177	2.27, 3.35, 175	*	*
pi-5-methylimidazole/C4-H···acetone	2.48, 3.56, 179	2.46, 3.54, 179	2.52, 3.38, 136	*	*
pi-5-methylimidazole/C2-H···acetone	2.31, 3.38, 175	2.30, 3.38, 176	2.30, 3.35, 164	*	*
5-methylimidazolium/C4-H···acetone	2.18, 3.24, 169	2.17, 3.24, 170	2.25, 3.10, 134	*	*
5-methylimidazolium/C2-H···acetone	2.05, 3.12, 173	2.04, 3.12, 175	2.06, 3.14, 177	*	*

Table 2b. ΔC–H bond length (in pm) upon complex formation, ΔH–O_{water} (or ΔC=$O_{acetone}$) bond length (in pm) upon complex formation.

tau-5-methylimidazole/C4-H···water	0.214, 0.019	0.202, 0.031	0.262, 0.024	0.133, 0.109	0.166, 0.033
tau-5-methylimidazole/C2-H···water	0.206, 0.022	0.200, 0.031	0.273, 0.028	0.115, 0.100	0.152, 0.029
pi-5-methylimidazole/C4-H···water	0.076, 0.030	0.097, 0.026	0.116, 0.027	0.022, 0.101	0.048, 0.035
pi-5-methylimidazole/C2-H···water	0.196, 0.021	0.195, 0.031	0.271, 0.023	0.111, 0.099	0.150, 0.030
5-methylimidazolium/C4-H···water	0.441, 0.041	0.429, 0.040	0.496, 0.025	0.339, 0.120	0.381, 0.038
5-methylimidazolium/C2-H···water	0.734, 0.059	0.713, 0.054	0.792, 0.046	0.602, 0.105	0.644, 0.045
tau-5-methylimidazole/C4-H···acetone	0.105, 0.043	0.131, 0.040	0.126, 0.084	*	*
tau-5-methylimidazole/C2-H···acetone	0.081, 0.051	0.094, 0.052	0.151, 0.085	*	*
pi-5-methylimidazole/C4-H···acetone	0.036, 0.018	0.047, 0.019	-0.027, 0.046	*	*
pi-5-methylimidazole/C2-H···acetone	0.110, 0.053	0.129, 0.055	0.311, 0.129	*	*
5-methylimidazolium/C4-H···acetone	0.331, 0.123	0.388, 0.107	0.160, 0.170	*	*
5-methylimidazolium/C2-H···acetone	0.665, 0.211	0.704, 0.215	0.753, 0.243	*	*

** Unable to geometrically optimize*

Intramolecular bond length variations for a species participating in an H-bonded complex may be interpreted through the lens of Gutmann's rules for electron pair donor-acceptor interactions.[54] Upon complex formation (X–H···Y–Z), respective intramolecular geometric changes occur within X–H and :Y–Z, not simply at the sites directly participating in the donor-acceptor interaction, but at more remote sites as well. Gutmann's first bond-length variation rule predicts that as the intermolecular distance between H_2O:→H–C (or C=O:→H–C) becomes smaller, the induced lengthening of the H–O_{water} and H–C bonds becomes larger. If the H_2O:→H–C interaction is a true donor-acceptor pair, bond lengthening of H–O_{water} (or C=O in acetone) and H–C will be observed. In terms of induced C–H bond lengthening, the bond-length variation rule appears to hold quite well for the imidazolium complexes of water and acetone; these complexes are calculated to have

the shortest H-bonds and also the greatest magnitude of induced C–H bond lengthening. A notable outlier in Table 2b appears to be the pi-C4–H···acetone complex at the M06-2X/6-31+G(d,p) level, which predicts a ΔC–H < 0. It should be noted that, overall, the BHandHLYP functional with a double-zeta quality basis set gave optimized geometries in good agreement with MP2/aug-cc-pVTZ geometries for the water complexes. While the bond-length variation rule does not appear to be strictly obeyed across the board, a cursory glance at Tables 2a and 2b reveals that it is still a useful rule of thumb.

Studies of changes in fractional charges on atoms in electron pair donor-acceptor species upon adduct formation were pioneered decades ago. For example, in the $H_3N{:}\rightarrow F_{(1)}–F_{(2)}$ interaction, calculations revealed a charge transfer of 0.0483e from ammonia to fluorine.[55] Changes in fractional charges (**Δ**e) upon adduct formation were as follows: each H atom, +0.0282; N, -0.0362; $F_{(1)}$, +0.0181; and $F_{(2)}$, -0.0664. Calculations by Leroy and Louterman-Leloup on the $H_2O{:}\rightarrow Cl_{(1)}–Cl_{(2)}$ interaction indicated a charge transfer from donor to acceptor of 0.003e. Changes in fractional charges upon adduct formation were as follows: each H atom, +0.186; O, -0.369; $Cl_{(1)}$, +0.023; and $Cl_{(2)}$, -0.026.[56] In these two specific examples and many others highlighted by Gutmann,[54] within an adduct, electron density increase is observed at the Lewis basic donor atom, described as the "pileup effect" and electron density decrease is observed at the Lewis acidic acceptor atom, described as the "spillover effect".[57] These studies showed that atoms adjacent to the donor atom become more positively charged while atoms adjacent to the acceptor atom become more negatively charged. The aforementioned charge transfer studies[55, 56] are relatively old, and so geometric optimization and NPA calculations on these adducts in the gas phase were carried out at the levels of theory relevant to this current study, the results of which are shown in Table 3.

Table 3. *Total NPA charge transfer (units of e) from electron pair donor to acceptor and change in fractional NPA charge (also units of e) on respective atoms upon adduct formation. The change in fractional NPA charge for a given atom is calculated using: ΔNPA charge = NPA charge of atom in adduct - NPA charge of atom in isolated species.*

		change in fractional NBO charge on each respective atom						
$H_3N{:}\rightarrow F_{(1)}–F_{(2)}$	Total charge transfer	H	N	O	$F_{(1)}$	$F_{(2)}$	$Cl_{(1)}$	$Cl_{(2)}$
BHandHLYP/6-31+G(d)	0.027	0.006	0.010	-	0.021	-0.048	-	-
BHandHLYP/6-31+G(d,p)	0.025	0.005	0.010	-	0.021	-0.046	-	-
M06-2X/6-31+G(d,p)	0.030	0.006	0.010	-	0.017	-0.047	-	-
MP2/6-31+G(d)	0.495	0.045	0.359	-	-0.280	-0.215	-	-
MP2/aug-cc-pVTZ	0.457	0.037	0.345	-	-0.266	-0.191	-	-
$H_2O{:}\rightarrow Cl_{(1)}–Cl_{(2)}$								
BHandHLYP/6-31+G(d)	0.023	0.010	-	0.002	-	-	0.020	-0.043
BHandHLYP/6-31+G(d,p)	0.019	0.010	-	0.000	-	-	0.022	-0.041
M06-2X/6-31+G(d,p)	0.022	0.011	-	0.000	-	-	0.019	-0.041
MP2/6-31+G(d,p)	0.016	0.010	-	-0.003	-	-	0.021	-0.037
MP2/aug-cc-pVTZ	0.013	0.007	-	-0.001	-	-	0.018	-0.031

As seen in Table 3, the pileup effect as described by Gutmann is observed in two cases, for $H_2O{:}\rightarrow Cl_{(1)}–Cl_{(2)}$ at the MP2 level. The spillover effect is observed for both adducts for all DFT methods, and the expected charge decrease on $F_{(2)}$ and $Cl_{(2)}$ in their respective adducts is seen at all levels. Also, the expected charge increase on hydrogen is seen at all levels of theory. This NPA charge study was extended to the Z–Y:→H–X hydrogen bonding interaction. In the Z–Y:→H–X interaction, Y: donates its electron pair to the Lewis acidic hydrogen and an electron density transfer (EDT) between the donor and acceptor species occurs. Numerical values for the EDT for each respective complex are shown in Table 4a. From the results displayed in Table 3 for the $H_3N{:}\rightarrow F_{(1)}–F_{(2)}$ and $H_2O{:}\rightarrow Cl_{(1)}–Cl_{(2)}$ adducts, the following reasonable predictions can be made for NPA charge studies of the 5-methylimidazole complexes using the computational methods relevant to this study:

1.) In terms of NPA charge, an EDT from the electron pair donor (hydrogen bond acceptor) to the electron pair acceptor

(hydrogen bond donor) will be observed.

2.) The pileup effect on the oxygen atom of water or acetone may not always be observed using NPA charge calculations.

3.) The H–O:→H–C (or C=O:→H–C) interaction will likely display the following properties: a decrease in NPA charge on the methylimidazole carbon, an increase in NPA charge for the hydrogen atom involved in the H-bond, and an increase in the NPA charge of the hydrogens of water (or the carbonyl carbon of acetone).

Table 4a. *Total NPA charge (units of e) transferred from water (or acetone) to the respective 5-methylimidazole isomer.*

Complex Name	BHandHLYP /6-31+G(d)	BHandHLYP /6-31+G(d,p)	M06-2X/6-31+G(d,p)	MP2/6-31+G(d)	MP2/aug-cc-pVTZ
tau-5-methylimidazole/C4-H···water	0.009	0.005	0.007	0.010	0.007
tau-5-methylimidazole/C2-H···water	0.008	0.007	0.009	0.010	0.007
pi-5-methylimidazole/C4-H···water	0.004	0.004	0.006	0.007	0.007
pi-5-methylimidazole/C2-H···water	0.008	0.006	0.008	0.010	0.007
5-methylimidazolium/C4-H···water	0.013	0.008	0.013	0.015	0.012
5-methylimidazolium/C2-H···water	0.018	0.016	0.018	0.022	0.017
tau-5-methylimidazole/C4-H···acetone	0.004	0.005	0.005	*	*
tau-5-methylimidazole/C2-H···acetone	0.006	0.006	0.007	*	*
pi-5-methylimidazole/C4-H···acetone	0.004	0.003	0.012	*	*
pi-5-methylimidazole/C2-H···acetone	0.004	0.006	0.004	*	*
5-methylimidazolium/C4-H···acetone	0.011	0.011	0.012	*	*
5-methylimidazolium/C2-H···acetone	0.019	0.020	0.020	*	*

Table 4b. *NPA charge study of electron pair donor component (water or acetone). Change in calculated NPA charge (units of e) on H_{water} or carbonyl carbon, change in calculated NPA charge on oxygen atom participating in the H-bond. The change in fractional NPA charge for a given atom is calculated using: ΔNPA charge = NPA charge of atom in complex - NPA charge of atom in isolated species.*

tau-5-methylimidazole/C4-H···water	0.003, 0.003	-0.001, 0.007	0.003, 0.000	0.004, 0.003	0.003, 0.002
tau-5-methylimidazole/C2-H···water	0.003, 0.002	-0.001, 0.008	0.003, 0.001	0.004, 0.003	0.003, 0.002
pi-5-methylimidazole/C4-H···water	0.001, 0.002	-0.002, 0.007	0.002, 0.001	0.002, 0.004	0.002, 0.004
pi-5-methylimidazole/C2-H···water	0.003, 0.002	-0.001, 0.007	0.003, 0.001	0.004, 0.003	0.003, 0.002
5-methylimidazolium/C4-H···water	0.007, -0.001	0.002, 0.005	0.008, -0.004	0.007, 0.002	0.006, 0.001
5-methylimidazolium/C2-H···water	0.010, -0.002	0.005, 0.006	0.011, -0.004	0.009, 0.005	0.008, 0.002
tau-5-methylimidazole/C4-H···acetone	0.005, -0.003	0.005, -0.003	0.003, 0.003	*	*
tau-5-methylimidazole/C2-H···acetone	0.006, -0.004	0.006, -0.005	0.008, -0.007	*	*
pi-5-methylimidazole/C4-H···acetone	0.002, 0.000	0.002, 0.000	-0.108, 0.129	*	*
pi-5-methylimidazole/C2-H···acetone	0.004, -0.002	0.005, 0.000	0.005, 0.000	*	*
5-methylimidazolium/C4-H···acetone	0.009, -0.005	0.009, -0.005	0.012, -0.010	*	*
5-methylimidazolium/C2-H···acetone	0.015, -0.008	0.015, -0.009	0.015, -0.008	*	*

Table 4c. *NPA charge study of electron pair acceptor component (5-methylimidazole). Change in calculated NPA charge (units of e) on hydrogen atom participating in the H-bond, change in calculated NPA charge on C4 or C2 carbon. The change in fractional NPA charge for a given atom is calculated using: ΔNPA charge = NPA charge of atom in complex - NPA charge of atom in isolated species.*

tau-5-methylimidazole/C4-H···water	0.014, -0.006	0.040, 0.058	0.014, -0.006	0.013, -0.008	0.016, -0.015
tau-5-methylimidazole/C2-H···water	0.015, -0.007	0.037, -0.031	0.015, -0.008	0.014, -0.009	0.016, -0.011
pi-5-methylimidazole/C4-H···water	0.014, -0.006	0.035, 0.045	0.013, -0.006	0.014, -0.005	0.013, -0.014
pi-5-methylimidazole/C2-H···water	0.015, -0.008	0.037, -0.030	0.015, -0.007	0.015, -0.012	0.017, -0.012
5-methylimidazolium/C4-H···water	0.013, -0.006	0.045, 0.066	0.013, -0.007	0.012, -0.007	0.018, -0.018
5-methylimidazolium/C2-H···water	0.012, -0.007	0.040, -0.026	0.012, -0.007	0.011, -0.009	0.019, -0.016
tau-5-methylimidazole/C4-H···acetone	0.011, -0.005	0.011, -0.004	0.009, -0.003	*	*
tau-5-methylimidazole/C2-H···acetone	0.012, -0.005	0.012, -0.005	0.010, -0.006	*	*

pi-5-methylimidazole/C4-H···acetone	0.010, -0.005	0.011, -0.005	-0.076, -0.015	*	*
pi-5-methylimidazole/C2-H···acetone	0.012, -0.006	0.011, -0.006	0.007, -0.005	*	*
5-methylimidazolium/C4-H···acetone	0.011, -0.006	0.012, -0.006	0.008, -0.003	*	*
5-methylimidazolium/C2-H···acetone	0.010, -0.007	0.010, -0.007	0.007, -0.008	*	*

** Unable to geometrically optimize*

As seen in Table 4a, a net EDT from water (or acetone) to 5-methylimidazole is calculated for all complexes across all levels of theory. Table 4c reveals that the spillover effect is observed for all water complexes except for the C4–H···water complexes at the BHandHLYP/6-31+G(d,p) level, and for all acetone complexes except for the pi-C4–H···acetone complex at the M06-2X/6-31+G(d,p) level. Both the pileup and the spillover effect are observed for the imidazolium complexes using DFT with the exception of the two water complexes at the BHandHLYP/6-31+G(d,p) level. While Gutmann's rules were established some 40 years ago, Denmark and Beutner invoke a Gutmann analysis several times in a relatively recent (2008) review of base-catalyzed organic synthesis, demonstrating the usefulness and modern applicability of the concept.[58]

Wiberg and Rablen recently conducted a study of atomic charges using seven different charge calculation methods, including both the NPA and Hirshfeld methods.[59] They examined the calculated charges on hydrogen for 23 different small organic molecules and also for an H-bonded methanol dimer at the MP2/aug-cc-pVTZ level and noted that the Hirshfeld method gave the most satisfactory results of the seven methods. Table 5 presents results of the Hirshfeld method at the level used by Wiberg and Rablen. The hydrogen charges at the C2–H (Cε–H) position are nearly the same for both the pi and tau-5-methylimidazole tautomers. However, the hydrogen charge at the C4–H (Cδ–H) position is 42% larger in the tau tautomer as compared to the pi tautomer; this aligns qualitatively with electrostatic potential data (Table 10), in which the electrostatic potential is 153% larger at the C4 site in the τ-tautomer as compared to the π-tautomer, while the potential at the C2 site is identical in the two tautomers. The Hirshfeld charges on C2–H and C4–H hydrogens in 5-methylimidazolium are approximately double that of the two neutral tautomers at the respective hydrogen positions. The charges on hydrogens in each complex are also shown in Table 5 to emphasize the fact that the charges on the C–H (and N–H) hydrogens not participating in the H-bond are only slightly perturbed upon complex formation (on the order of 0.001-0.003 e); from a strictly electrostatic point of view, this would seem to suggest that formation of the first H-bond does not diminish the molecule's capacity for further H-bond formation at the other H-bond donor sites. Such multiple H-bonding capacity can be seen in Table 1 and Figure 2 with structure 5GX9 (mutant E46Q PYP), in which the Histidinium108 is H-bonded at both the C2–H and C4–H sites.

Table 5. *Hirshfeld charges on hydrogen atoms (units of e) calculated at MP2/aug-cc-pVTZ level on MP2/aug-cc-pVTZ geometries. Refer to Figure 1 for hydrogen atom locations.*

Compound/Complex Name	N1-H	C2-H	N3-H	C4-H
tau-5-methylimidazole	-	**0.0671**	**0.1766**	**0.0674**
tau-5-methylimidazole/C4-H···water	-	0.0657	0.1748	0.0424
tau-5-methylimidazole/C2-H···water	-	0.0408	0.1746	0.0662
pi-5-methylimidazole	**0.1758**	**0.0664**	-	**0.0475**
pi-5-methylimidazole/C4-H···water	0.1750	0.0651	-	0.0286
pi-5-methylimidazole/C2-H···water	0.1739	0.0401	-	0.0461
5-methylimidazolium	**0.2090**	**0.1223**	**0.2119**	**0.1001**
5-methylimidazolium/C4-H···water	0.2078	0.1207	0.2093	0.0642
5-methylimidazolium/C2-H···water	0.2061	0.0781	0.2088	0.0982

According to IUPAC, "The length of the X–H bond usually increases on hydrogen bond formation leading to a red shift in the infrared X–H stretching frequency and an increase in the infrared absorption

cross-section for the X–H stretching vibration."[51] Table 6 summarizes the results of IR calculations, including shifts in the respective C–H stretching frequencies and concomitant changes in band intensity upon complex formation. Most complexes analyzed in this study display red-shifted C–H bonds, with notable exceptions underlined in Table 6. The C2–H complexes at all levels display negative $\Delta\nu_{C-H}$ values with concomitant increases in $\Delta\nu_{C-H}$ intensity. Hobza and Havlas note that "formation of the X–H···Y hydrogen bond results in weakening of the X–H bond. This weakening is accompanied by a bond elongation and a concomitant decrease of the X–H stretch vibration frequency compared to the noninteracting species."[60] Hobza and Havlas were describing a classic red-shifting H-bonding interaction, which they prefer to describe simply as an "H-bond", in contrast to a blue-shifting H-bond, which they prefer to describe as an "improper, blue-shifting" H-bond. The π-C4–H···acetone complex at the BHandHLYP/6-31+G(d) level displays both a blue-shifted C–H stretching mode and a band intensity decrease, which is evidence for an improper, blue-shifting H-bond.[60 61] It would be expected that the C–H bond in this complex would be shortened relative to the isolated 5-methylimidazole, but curiously, the ΔC–H bond length for this complex is +0.036 pm as seen in Table 2b.

Table 6. *IR Data. $\Delta\nu_{C-H}$ (cm^{-1}), $\Delta\nu_{C-H}$ intensity. A negative value of $\Delta\nu_{C-H}$ and a positive value of $\Delta\nu_{C-H}$ intensity indicates evidence for a red-shifting H-bond interaction. Underlined data indicates outliers, which have either a positive $\Delta\nu_{C-H}$ or a negative $\Delta\nu_{C-H}$ intensity (or both). $\Delta\nu_{C-H}$ calculated from $\Delta\nu_{C-H} = \nu_{C-H\ in\ complex} - \nu_{C-H\ in\ separated\ species}$ and similarly for $\Delta\nu_{C-H}$ intensity.*

Complex Name	BHandHLYP/6-31+G(d)	BHandHLYP/6-31+G(d,p)	M06-2X/6-31+G(d,p)	MP2/6-31+G(d)	MP2/aug-cc-pVTZ
tau-5-methylimidazole/C4-H···water	-26.0, 56.4	-23.2, 54.2	-24.5, 77.4	-14.7, 64.3	-18.9, 54.2
tau-5-methylimidazole/C2-H···water	-23.6, 51.4	-22.4, 52.6	-44.5, 70.6	-9.5, 60.1	-15.8, 57.4
pi-5-methylimidazole/C4-H···water	-3.4, -0.3	-6.7, 3.3	-22.9, 10.2	3.1, 12.3	-1.3, 2.6
pi-5-methylimidazole/C2-H···water	†	†	-10.6, 73.9	-8.1, 62.9	-14.6, 63.7
5-methylimidazolium/C4-H···water	-61.3, 166.5	-59.3, 156.2	-61.1, 186.3	-44.9, 168.0	-52.8, 169.3
5-methylimidazolium/C2-H···water	-113.6, 347.8	-112.3, 350.6	-99.3, 364.0	-91.1, 335.9	-101.2, 363.9
tau-5-methylimidazole/C4-H···acetone	-5.9, 44.1	-11.8, 51.5	12.3, 20.2	*	*
tau-5-methylimidazole/C2-H···acetone	-1.0, 37.6	-4.3, 44.9	-22.0, 59.2	*	*
pi-5-methylimidazole/C4-H···acetone	4.2, -4.1	1.4, 0.4	-2.8, -4.0	*	*
pi-5-methylimidazole/C2-H···acetone	-7.0, 45.4	-11.3, 57.8	†	*	*
5-methylimidazolium/C4-H···acetone	-43.8, 177.6	-54.0, 190.9	†	*	*
5-methylimidazolium/C2-H···acetone	-102.3, 417.7	-111.6, 449.2	-76.9, 431.7	*	*

**Unable to geometrically optimize*
†Unable to visually distinguish ν_{C2-H} and ν_{C4-H} stretching modes in GaussView 5.0

The imidazolium complexes display the largest C–H frequency shifts and intensity increases. IUPAC notes that upon formation of an H-bond, "new vibrational modes associated with the formation of the H···Y bond are generated."[51] Indeed, new low energy IR vibrational modes are calculated for all complexes at all levels of theory (for full listing, see Table S2 in supporting information). As an arbitrary example, in the imidazolium-C4–H···water complex at the MP2/aug-cc-pVTZ level, seven new low energy vibrational modes associated with the complex are calculated at 143, 132, 110., 58.5, 50.3, 39.0, and 27.2 cm^{-1}. GaussView 5.0 software allowed for visual identification of these new modes. The new mode at 110. cm^{-1} is a very pronounced H···O_{water} stretch of the C–H···O_{water} hydrogen bond.

According to IUPAC, "the X–H···Y–Z hydrogen bond leads to characteristic NMR signatures that typically include pronounced proton deshielding for H in X–H, through hydrogen bond spin-spin couplings between X and Y, and nuclear Overhauser enhancements."[51] Table 7 reports changes in calculated isotropic shielding (σ) tensors upon complex formation for the hydrogen atom participating in the H-bond. The isotropic shielding tensors decrease (i.*e.* predicted downfield shift in the NMR spectrum) on the respective hydrogen atoms involved in H-bonding upon complex formation, with the most pronounced changes occurring for the imidazolium complexes. Two methods for proton NMR shift prediction, WP04/aug-cc-pVDZ[62] and PBE0/cc-pVTZ[63], were recommended in the literature. For neutral histidine, both methods give very similar predictions. The differences between methods for the imidazolium complexes are larger, but the qualitative trends are similar.

Table 7. *Change in calculated isotropic shielding (Δσ in units of ppm) on hydrogen atom participating in H-bond. Δ isotropic shielding = isotropic shielding of atom in complex - isotropic shielding of atom in separated species. Calculations performed on BHandHLYP/6-31+G(d,p) geometries.*

Complex Name	WP04/aug-cc-pVDZ*	PBE0/cc-pVTZ†
tau-5-methylimidazole/C4-H···water	-1.07	-1.11
tau-5-methylimidazole/C2-H···water	-1.08	-1.06
pi-5-methylimidazole/C4-H···water	-0.83	-0.86
pi-5-methylimidazole/C2-H···water	-1.06	-1.05
5-methylimidazolium/C4-H···water	-1.61	-1.86
5-methylimidazolium/C2-H···water	-2.06	-2.33
tau-5-methylimidazole/C4-H···acetone	-1.41	-1.41
tau-5-methylimidazole/C2-H···acetone	-1.38	-1.33
pi-5-methylimidazole/C4-H···acetone	-1.12	-1.16
pi-5-methylimidazole/C2-H···acetone	-0.98	-0.95
5-methylimidazolium/C4-H···acetone	-2.03	-2.19
5-methylimidazolium/C2-H···acetone	-2.51	-2.81

*These calculations were performed using the GIAO method and with IEF-PCM(Water)[64] implicit solvation. WP04: BLYP IOp (3/76=1000001189, 3/77 = 0961409999, 3/78 = 0000109999).

†These calculations were performed using the CSGT method in the gas phase.

IUPAC states that "The Gibbs energy of formation for the hydrogen bond should be greater than the thermal energy of the system for the hydrogen bond to be detected experimentally."[51] Tables 8a and 8b do not display ΔG values, but rather the change in electronic energy upon complex formation. In a similar study, gas-phase counterpoise corrected calculations using MP2(frozen core approximation)/6-31+G(d,p) by Scheiner and coworkers found the following interaction energies: for imidazole, the Cε–H···OH_2 complex ΔE = −2.4 kcal/mol and the Cδ–H···OH_2 complex ΔE = −2.3 kcal/mol; for imidazolium, the Cε–H···OH_2 complex ΔE = −11.3 kcal/mol and the Cδ–H···OH_2 complex ΔE = −9.5 kcal/mol.[65] Results in the present study on 5-methylimidazole (and 5-methylimidazolium) complexes most comparable to Scheiner's results are gas phase MP2 calculations shown in Table 8b below. Although a different basis set was used in the present study, values in Table 8b are similar to those from Scheiner and coworkers.[65] Values for ΔE_{MP2} between these two studies for complex formation with water are within about 1 kcal/mol. Also reported in Figure 8b are ΔE_{MP2} values for complexes involving the N–H groups acting as H-bond donors, and complexes with the C=O group of N-methylacetamide (the "amide" complexes) acting as an H-bond acceptor. The optimized structures and geometric parameters for these complexes can be seen in Figures S1 and S2. As expected, complexes where N–H acts as the H-bond donor group have more negative ΔE_{MP2} values than the comparable C–H complexes. The complexes with N-methylacetamide have more negative ΔE_{MP2} values than the corresponding complexes involving acetone; this can perhaps be explained simply by examining the resonance structures for N-methylacetamide. Two resonance structures of N-methylacetamide place a negative formal charge on the oxygen atom, as opposed to acetone, where only one resonance structure places a formal negative charge on the oxygen atom.

Table 8a. *ΔE of complex formation (kcal/mol). $\Delta E_{BHandHLYP}$, $\Delta E_{M06\text{-}2X}$, and ΔE_{MP2}, respectively. Energies taken from output files of geometric optimizations (with C-PCM implicit solvation in water). ΔE = E (complex) - Σ E (separated species).*

Complex Name	BHandHLYP /6-31+G(d)	BHandHLYP/6-31+G(d,p)	M06-2X/6-31+G(d,p)	MP2/6-31+G(d)	MP2/aug-cc-pVTZ
tau-5-methylimidazole/C4-H···water	-0.88	-0.95	-1.39	-1.66	-1.37
tau-5-methylimidazole/C2-H···water	-1.05	-1.11	-1.43	-1.63	-1.33
pi-5-methylimidazole/C4-H···water	-0.44	-0.52	-0.89	-1.16	-0.99
pi-5-methylimidazole/C2-H···water	-1.03	-1.09	-1.37	-1.63	-1.32
5-methylimidazolium/C4-H···water	-1.96	-1.98	-2.40	-2.75	-2.27
5-methylimidazolium/C2-H···water	-3.12	-3.11	-3.38	-3.53	-2.95

tau-5-methylimidazole/C4-H···acetone	-0.69	-0.74	-2.15	*	*
tau-5-methylimidazole/C2-H···acetone	-0.81	-0.86	-1.06	*	*
pi-5-methylimidazole/C4-H···acetone	-0.36	-0.40	-0.94	*	*
pi-5-methylimidazole/C2-H···acetone	-0.78	-0.85	-1.74	*	*
5-methylimidazolium/C4-H···acetone	-1.60	-1.68	-2.34	*	*
5-methylimidazolium/C2-H···acetone	-2.68	-2.81	-3.06	*	*

Table 8b. *ΔE_{MP2} (kcal/mol) of complex formation at MP2/aug-cc-pVTZ level using BHandHLYP/6-31+G(d,p) optimized geometries. Solvents of different dielectrics (given in parenthesis) were tested using the C-PCM solvation model. $\Delta E_{MP2} = E_{MP2}$ (complex) - ΣE_{MP2} (separated species). Gas phase calculations are included for reference.*

Complex Name	water (78.36)	DMSO (46.83)	acetonitrile (35.69)	ethanol (24.85)	acetone (20.49)	dichloro-methane (8.93)	heptane (1.91)	gas phase (1.00)
tau-5-methylimidazole/C4-H···water	-1.35	-1.36	-1.37	-1.40	-1.41	-1.52	-2.12	-2.64
tau-5-methylimidazole/C2-H···water	-1.32	-1.33	-1.34	-1.36	-1.37	-1.46	-1.98	-2.48
pi-5-methylimidazole/C4-H···water	-0.97	-0.97	-0.98	-0.98	-0.98	-1.01	-1.20	-1.47
pi-5-methylimidazole/C2-H···water	-1.32	-1.33	-1.34	-1.36	-1.37	-1.46	-1.98	-2.48
5-methylimidazolium/C4-H···water	-2.28	-2.34	-2.39	-2.47	-2.53	-2.99	-5.86	-9.06
5-methylimidazolium/C2-H···water	-2.94	-3.01	-3.06	-3.16	-3.22	-3.72	-6.83	-10.26
tau-5-methylimidazole/C4-H···acetone	-1.82	-1.84	-1.85	-1.87	-1.88	-1.99	-2.57	-3.05
tau-5-methylimidazole/C2-H···acetone	-1.71	-1.72	-1.73	-1.75	-1.76	-1.85	-2.39	-2.90
pi-5-methylimidazole/C4-H···acetone	-1.42	-1.42	-1.42	-1.42	-1.43	-1.44	-1.61	-1.89
pi-5-methylimidazole/C2-H···acetone	-1.79	-1.80	-1.82	-1.84	-1.86	-1.98	-2.70	-3.35
5-methylimidazolium/C4-H···acetone	-2.88	-2.96	-3.02	-3.13	-3.20	-3.77	-7.44	-11.60
5-methylimidazolium/C2-H···acetone	-3.58	-3.68	-3.75	-3.88	-3.97	-4.63	-8.83	-13.49
tau-5-methylimidazole/C4-H···amide	-2.15	-2.18	-2.20	-2.23	-2.25	-2.42	-3.31	-4.03
tau-5-methylimidazole/C2-H···amide	-2.25	-2.27	-2.28	-2.30	-2.32	-2.43	-3.10	-3.75
pi-5-methylimidazole/C4-H···amide	-1.65	-1.65	-1.65	-1.66	-1.66	-1.68	-1.94	-2.34
pi-5-methylimidazole/C2-H···amide	-2.11	-2.13	-2.14	-2.17	-2.19	-2.35	-3.21	-4.01
5-methylimidazolium/C4-H···amide	-4.02	-4.12	-4.20	-4.33	-4.43	-5.13	-9.57	-14.51
5-methylimidazolium/C2-H···amide	-4.83	-4.96	-5.05	-5.23	-5.35	-6.25	-11.77	-17.67
tau-5-methylimidazole/N3-H···water	-4.93	-4.95	-4.97	-4.99	-5.01	-5.13	-5.81	-6.39
pi-5-methylimidazole/N1-H···water	-5.04	-5.06	-5.07	-5.09	-5.11	-5.23	-5.90	-6.46
5-methylimidazolium/N1-H···water	-7.37	-7.44	-7.50	-7.60	-7.68	-8.22	-11.68	-15.54
5-methylimidazolium/N3-H···water	-7.31	-7.39	-7.45	-7.56	-7.64	-8.21	-11.79	-15.75
tau-5-methylimidazole/N3-H···acetone	-5.83	-5.86	-5.88	-5.92	-5.94	-6.13	-7.13	-7.95
pi-5-methylimidazole/N1-H···acetone	-5.97	-6.00	-6.02	-6.06	-6.09	-6.27	-7.27	-8.06
5-methylimidazolium/N1-H···acetone	-8.40	-8.51	-8.59	-8.75	-8.86	-9.66	-14.76	-20.50
5-methylimidazolium/N3-H···acetone	-8.28	-8.39	-8.48	-8.64	-8.75	-9.57	-14.75	-20.49

**Unable to geometrically optimize*

Hydrogen bond interaction energies depend on dielectric environment, and amino acids within a protein can experience a range of dielectric values depending on location (exterior vs. interior) and microenvironment. According to Anslyn and Dougherty, "Often an 'effective dielectric' constant anywhere in the range of 4-37 is ascribed to the interior of a protein. However, we are no longer in a relatively homogeneous medium like in a pure solvent, and so any such approximation must be considered fairly crude."[66] Refinements to modeling the dielectric environment of proteins have been examined.[67] It is clear from Table 8b that the dielectric of the environment impacts interaction energies, especially for the charged imidazolium complexes; the ΔE_{MP2} value for a given complex becomes more negative as the solvent dielectric (ε) is decreased using the C-PCM model. In a plot of ΔE_{MP2} vs. solvent dielectric for pi and tau complexes of acetone and water (see Figures S3 and S4), $d\Delta E_{MP2}/d\varepsilon$ has a small positive sign and is essentially constant

from ε=10 to ε=78. Below 10, $d\Delta E_{MP2}/d\varepsilon$ noticeably increases for the π- and τ-complexes. For the protonated complexes, the slope $d\Delta E_{MP2}/d\varepsilon$ is slightly larger relative to the π- and τ- complexes and is constant from about 20 to 78. Below 20, $d\Delta E_{MP2}/d\varepsilon$ dramatically increases for imidazolium complexes, and ΔE_{MP2} becomes quite negative when ε < ~10. The overall trend for the acetone and water complexes is that as solvent dielectric increases, ΔE_{MP2} becomes more positive and $d\Delta E_{MP2}/d\varepsilon$ decreases. Therefore, generally, C–H hydrogen bonds of histidine in the protein interior can be expected to be stronger than equivalent ones at the protein surface.

Bond critical points (bcp) were analyzed using both the HF and MP2 electron densities. According to IUPAC, "analysis of the electron density topology of hydrogen bonded systems usually shows a bond path connecting H and Y and a (3, −1) bond critical point between H and Y".[51] It is expected that for C–H···O hydrogen bonding complexes that the Laplacian of the HF charge density ($\nabla^2\varrho$) at the bcp will be positive and fall roughly between 0.024 and 0.139, as small variation is observed across different basis sets.[52] For a set of four H-bonding complexes (formaldehyde···chloroform, acetone···chloroform, benzene···formaldehyde, and 1,1-dichloroethane···acetone) involving C–H···O interactions, HF calculations revealed a range of electron densities (ϱ) at the bond critical points from 0.0039 to 0.0205 au.[52] All complexes of interest to this study have computed $\nabla^2\varrho$ and ϱ values that fall within these established ranges, as shown in Table 9, with the imidazolium complexes having larger $\nabla^2\varrho$ and ϱ values than the comparable neutral complexes.

Table 9. *bcp analysis of the complexes. Electron density (ϱ) at bcp and Laplacian of the charge density at the bcp ($\nabla^2\varrho$) using MP2 densities. Values using HF densities are shown in parentheses to enable direct comparison with reference 49.*

Complex Name	*ϱ (au)*	*$\nabla^2\varrho$*	*BCP distance from hydrogen (Å)*
tau-5-methylimidazole/C4-H···water	0.01101 (0.01102)	0.05468 (0.05450)	0.889 (0.889)
tau-5-methylimidazole/C2-H···water	0.01193 (0.01192)	0.05569 (0.05563)	0.868 (0.868)
pi-5-methylimidazole/C4-H···water	0.01101 (0.00896)	0.05468 (0.04314)	0.889 (0.948)
pi-5-methylimidazole/C2-H···water	0.01193 (0.01184)	0.05569 (0.05460)	0.868 (0.876)
5-methylimidazolium /C4-H···water	0.01521 (0.01524)	0.07455 (0.07409)	0.807 (0.807)
5-methylimidazolium /C2-H···water	0.01982 (0.01983)	0.09230 (0.09213)	0.735 (0.735)
tau-5-methylimidazole/C4-H···acetone	0.00914 (0.00913)	0.04640 (0.04656)	0.931 (0.931)
tau-5-methylimidazole/C2-H···acetone	0.00962 (0.00960)	0.05101 (0.05140)	0.907 (0.907)
pi-5-methylimidazole/C4-H···acetone	0.00696 (0.00696)	0.03642 (0.03631)	0.998 (0.997)
pi-5-methylimidazole/C2-H···acetone	0.01016 (0.01015)	0.05181 (0.05177)	0.953 (0.907)
5-methylimidazolium/C4-H···acetone	0.01394 (0.01395)	0.06866 (0.06845)	0.825 (0.825)
5-methylimidazolium/C2-H···acetone	0.01906 (0.01907)	0.09114 (0.09107)	0.742 (0.742)

Electrostatic potential surfaces were generated for tau-5-methylimidazole, pi-5-methylimidazole, and 5-methylimidazolium. The electrostatic potential was mapped onto the 0.002 au electron density isosurface in GaussView 5.0, as the 0.002 au contour may be used to reasonably represent the size and shape of a molecule in the condensed phase.[68] Numerical values of electrostatic potential are displayed in Table 10 and the electrostatic potential maps are displayed in Figure 5.

Table 10. *Electrostatic potential (kcal/mol) on the 0.002 au isodensity surface. See text for identification of the respective sites.*

	N1 site	*C2 site*	*N3 site*	*C4 site*
tau-5-methylimidazole	-62	35	80	38
pi-5-methylimidazole	80	35	-65	15
5-methylimidazolium	173	149	172	122

To define the sites listed in Table 10, a specific example is given as follows: For the C2 site, the line representing the C2–H bond path is extrapolated and followed out and away from the molecule to the 0.002 au isodensity surface; the potential is recorded at this point. A similar process is followed for the other three sites (see Figure S5 for an explicit pictorial representation of the site locations). This process was carried out in GaussView 5.0. Clearly, the electrostatic potential at each hydrogen bond donor or hydrogen bond acceptor site is dependent upon both tautomeric and protonation state. The electrostatic potential at the C4 site of tau-5-methylimidazole is over twice that of pi-5-methylimidazole, while the C2 site in both tautomers is

equivalent in terms of the potential. In 5-methylimidizolium, the potential at the C2 site approaches that of the N1 and N3 sites. The electrostatic potential values trend with the hydrogen bond strengths indicated by ΔE values. In summary, electronic structure calculations show that histidinium forms the strongest C–H hydrogen bonds, with somewhat stronger hydrogen bonds for C2 than C4. The tau tautomer forms the next strongest hydrogen bonds, with nearly equal H-bond strength at the C2 and C4 sites. The pi tautomer forms the weakest hydrogen bond at the C4 site, while its C2–H hydrogen bond is near in strength to those in the tau tautomer.

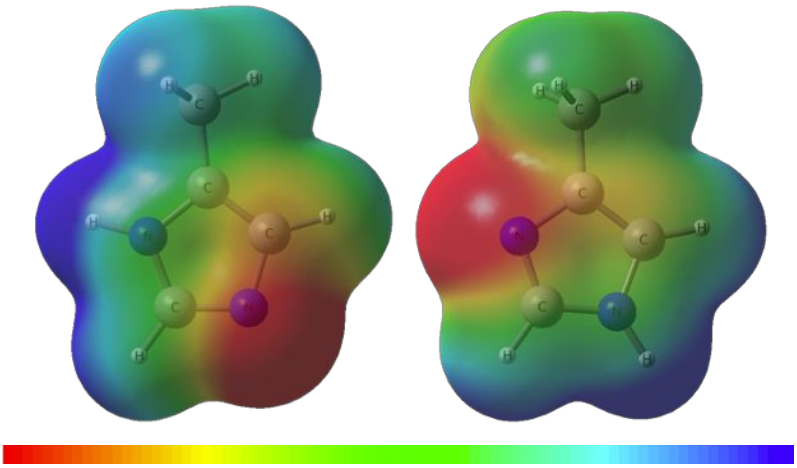

Figure 5. Electrostatic potential mapped on the 0.002 au isodensity surface in GaussView 5.0. Color bar ranges from -38 kcal/mol (most red) to +38 kcal/mol (most blue). Left: π-5-methylimidazole. Right: τ-5-methylimidazole.

5. DISCUSSION

The observations of C–H hydrogen bonds from histidine in the active sites of enzymes sheds new light on several highly-studied systems, including trypsin, human carbonic anhydrase II, myoglobin, and RNase A. Some of these interactions involve coordination of water, while others involve interactions with key residues, such as that between His57 and Ser214 of trypsin. In the crystal structure, we observe Cε–H of cationic His57 coordinated to the main chain of Ser214, which may function to orient the His57 for the critical side chain interaction between the Nε1 of histidine with the O–H of Ser195 to stabilize the tetrahedral intermediate transition state.[29] Mechanistic studies using QM-MM simulations indicate that His57 undergoes a subtle reorientation, rather than a ring flip, as it abstracts a proton from the weak acid side chain.[69] The authors note that His57 "needs to be ideally positioned to deprotonate [Ser195]" in the tetrahedral intermediate during acylation. Zhou and Zhang observed that the H-bond between the carboxylate of Asp102 and the proton on Nδ2 shortens during the transition state.[69] Our identification of the hydrogen bond between Cε–H of His57 and Ser214 invites a revisitation of mechanistic data to understand to what extent this interaction also influences His57 positioning and reorientation during the transition state, and also how it plays a role in the electrostatic microenvironment around His57 that facilitates charge build up and proton transfer steps from Ser195 to His57 and then from His57 to substrate.

Meanwhile, in myoglobin, the His93 Cε–H bond to the heme carboxylate group must affect the rotational barrier of this proximal histidine coordinated to the heme iron. A hydrogen bond between Ser92 and His93 (Nε–H) has previously been identified to influence histidine rotation.[70] It is known that energetic barriers for internal bond rotation are influenced by hydrogen bonds, but not restricted entirely.[71] In myoglobin, after release of O_2/CO ligand, the protein undergoes a conformational change involving change in the Fe–N bond, which was described as "proteinquake" relaxation in seminal work on protein dynamics.[72] The contributions of the carboxylate-Cε–H interaction to the proteinquake rearrangement, also described as heme doming,[38] could be an area for future investigation.

In RNase A, cationic His12 is seen to coordinate water molecules via both the Cε–H and Cδ–H bonds, which are in turn coordinated to water H-bonded to the adjacent N(ε1)–H group. The solvation shell attenuates pKa and tautomeric equilibria of histidine side chains, although it remains an open question whether the same water coordination network observed in the crystal structure exists in solution phase. Given the open questions about tautomeric form in the mechanism of RNase A, [13 30 31 32] the structure of His12 and His119 solvent shells merits further investigation. However, for MD simulations to effectively do so,

special attention to methods (and tautomeric form) is needed to capture C–H bonding effects in histidine. This could involve, for instance, use of QM/MM simulations or custom charges on active site residues with fixed-charge force fields (as was done by Merz and co-workers in HCA II, PDB ID 4Q49, discussed below).[73] In summary, the present results suggest water coordination by multiple atoms in the histidine side chain may participate in stabilization or facilitation of the charge transfer/proton transfer steps. Furthermore, atomic charge calculations indicate that participation in a C–H hydrogen bond does not reduce polarity, in line with the observation of multiple hydrogen bonds by many of the histidine residues observed in multiple neutron diffraction structures (Table 1), including β-lactamase and RNase A.

The active site C–H hydrogen bond between His119 and Ser214 in HCA II provides new information about a hydrogen bond network that has been shown to be critical in enzymatic function. His119 is one of three histidine residues coordinated to zinc, which binds H_2O/OH^-, surrounded by an ordered water network mediated by hydrophilic residues.[74] Simulations by Merz and co-workers, in which custom charges were calculated for the histidine residues by QM methods show that this region of the active site has low RMSF (high rigidity). They concluded that the low flexibility in the cavity around the active site leads to an optimal geometry for enzymatic activity.[73] In fact, His119 was observed to have the lowest RMSF value among the three histidine residues bound to zinc. The H-bond between Glu106 and Cε–H of His119 must contribute to the low flexibility of this particular histidine, as well as contributing to the rigidity of Glu106.[73] This glutamate residue has been shown to dictate ligand preferences at the HCAII active site.[33] The Glu106 H-bonded to Cε–H of His119 is in turn also H-bonded to Thr199. This latter H-bond between Thr199 and Glu106 was shown previously in MD simulations to be long-lived, locked into catalytically-favorable geometries through a rigid H-bond network.[73] Thus, the His119 C–H bond in the active site of HCA II may make an underappreciated contribution to the enzymatic selectivity and efficiency of the enzyme.

Although we relied on neutron diffraction data in this study, as it provides more accurate proton locations that aid in particular with assignment of protonation/tautomeric state, it is possible to identify histidine Cδ-H and Cε-H hydrogen bonds in X-ray crystal structures. Therefore, the interactions identified herein can be looked for within past, current, and future systems of interest regardless of diffraction method. As a proof of principle, we arbitrarily selected 6 X-ray diffraction structures of protein crystals. Using a heavy atom distance cut-off of 3.6 Å, numerous histidine C-H hydrogen bonds were identified. These are presented in Table S6 and Figures S6-S17. The average heavy-atom bond distance for C-H hydrogen bonds in Table S6 is 3.2 ± 0.3 Å.

Most of the heavy atom distances observed for histidine C–H bonds in the neutron and X-ray diffraction structures are longer than many "traditional" N-H and O-H hydrogen bonds, ranging from 3.1-3.4 Å in most cases, concomitant with the larger van der Waals radius of carbon. A recent survey by Merski et al. of heavy atom distances in protein H-bonds found the average for C–H hydrogen bonds to be 3.3 Å.[75] A longer H-bond distance typically indicates a weaker interaction, but even longer and weaker hydrogen bonds can still be critical: recent work shows that the synergy of multiple hydrogen bonds in peptides shapes the conformational landscape.[5 76 77 78] It is notable that comparison of C-H---O and N-H---O interactions in an *ab initio* study by Scheiner indicated that the hydrogen bonds are essentially isoenergetic in parallel β-sheets while N-H---O interactions are slightly stronger in antiparallel β-sheets.[79] It is also interesting to note that the newly-identified Cδ–H hydrogen bond of histidine shows redshifted vibrational frequencies, a common hallmark of "traditional" H-bonds, in contrast with blueshifting Cε–H hydrogen bonds.

It appears that the Cδ–H group, particularly, can act as a hook that reaches back in toward a carbonyl hydrogen bond acceptor in the protein structure to anchor the histidine side chain in a specific orientation, including in the active site. Meanwhile, Cδ and Cε groups were observed to serve as a bridge "reaching across" in secondary structures like β sheets, and this was especially prominent for Cδ–H groups, in spite of the steric restrictions inherent in the Cδ position on the histidine aromatic ring. It has been previously shown that side chain O–H interactions with backbone carbonyls seem to have an important effect on protein structural stability.[80] Both types of C–H bonds in the histidine ring were observed to coordinate water, and protein-water hydrogen bonds influence solvent shell structure and dynamics,[81 82 83 84] which in turn play a role in both the stability and flexibility/conformational dynamics of the proteins.[85 86] In summary, the finding and characterization of Cδ–H hydrogen bonds in histidine adds to the growing knowledge of the diversity of hydrogen bonds, which are crucially involved in myriad ways in protein structure-function-dynamics.

6. CONCLUSIONS

For hydrogen bonds in proteins, strength lies in numbers. Collectively, hydrogen bonds inform the folded shape, the structure and dynamics of the hydration layer, and how substrate interacts and reacts at the active site of enzymes. Hydrogen bonds vary in strength, but taken all together they are integral to protein structure-function-dynamics. Our previous analysis of electronic structure calculations of histidine suggested, from bond polarity, that the C4(Cδ)–H bond of histidine is a viable hydrogen bond donor.[9 12 13 14] Investigation of neutron diffraction structures, allowing for identification of tautomeric/protonation state, yielded experimental evidence for Cδ–H hydrogen bonds in proteins, in addition to C2(Cε)–H hydrogen bonds, which have been identified previously as a type of hydrogen bond. Common motifs for histidine Cδ–H hydrogen bonds in crystal structures include coordinated water molecules and interactions in beta sheets and loops that presumably help stitch together the secondary structure. Notable was the observation of a double C–H hydrogen bond (at both the C2/Cε and C4/Cδ sites) at His108 of photoactive yellow protein and in the active site residue His12 in RNase A. Computational analysis has characterized the properties of the C–H···O hydrogen bonds in methylimidazole–water, –acetone, and –N-methylacetamide complexes in various protonation/tautomeric states of the imidazole ring. IUPAC H-bonding criteria were used to make this characterization systematic.

The Cδ–H···O complexes within the crystal structures shown in Figure 2 meet the IUPAC bond angle criteria for H-bonding (C–H···O angle > 110°), and many of the experimental listings meet an older empirical van der Waals criteria (C to O length < 3.3 Å), although the latter criterion is not recommended by IUPAC. Based on the model compounds used for electronic structure calculations, there is ample evidence that all complexes presented here can be characterized as H-bonded complexes, and the Cδ–H bonds in histidine and histidinium should be viewed as viable H-bond donor groups. The model systems show that, if there is sufficient configurational freedom for all atoms in a real system, that the equilibrium geometries fulfill the overwhelming majority of IUPAC criteria for H-bonding, as seen in Table 11 for Cε–H···O and Cδ–H···O model complexes with water and acetone. Moreover, from Hirshfeld charge data, multiple H-bond formation at one histidine residue is viable from a purely electrostatic viewpoint.

Table 11. *The following IUPAC H-bonding criteria[51] were examined with respect to the BHandHLYP/6-31+G(d,p) optimized geometries. Whether the criteria are met for a given methylimidazole···water or acetone complex is indicated by "Yes", "No", or "Inconclusive".*

IUPAC Criteria	τ-Cδ-H water, acetone	τ-Cε-H water, acetone	π-Cδ-H water, acetone	π-Cε-H water, acetone	imidazolium-Cδ-H water, acetone	imidazolium-Cε-H water, acetone
ΔG for H-bond formation > thermal energy of the system	Not calculated	Not calculated	Not calculated	Not calculated	Not calculated	Not calculated
(3,-1) Bond Critical Point found along C–H···O bond path	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes
Charge transfer from water or acetone, respectively, to the methylimidazole species	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes
$\Delta E_{complex}$ shows good correlation with extent of charge transfer between donor and acceptor	Yes (R^2 = 0.934) , Yes (R^2 = 0.996) - See tables S3A and S3B					
Nuclear deshielding of proton due to C–H···O interaction	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes
Red-shifted C–H IR stretching frequency	Yes, Yes	Yes, Yes	Yes, No	Inconclusive, Yes	Yes, Yes	Yes, Yes
Increased C–H IR stretching absorption intensity	Yes, Yes	Yes, Yes	Yes, Yes	Inconclusive, Yes	Yes, Yes	Yes, Yes
New vibrational modes associated with C–H···O interaction	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes
C–H···O angle > 110°	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes

ΔC–H bond length > 0	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes	Yes, Yes

Compared to its neutral form, 5-methylimidazolium forms Cε–H···O and Cδ–H···O hydrogen bonding complexes roughly twice as strong (or more) than the respective neutral tau-5-methylimidazole or π-5-methylimidazole complexes. The τ- and the π-tautomers form Cε–H···O complexes of very similar strength, the τ-Cδ–H···O interaction is approximately twice as strong as the π-Cδ–H···O interaction, and the π-Cδ–H···O interaction forms the weakest complexes. The data indicates that in the neutral form, tautomeric state matters, which can perhaps be most vividly demonstrated by examination of the electrostatic potential maps in Figure 5 and examination of Hirshfeld charges (Table 5) at the Cδ–H site in the π- and τ-tautomers. Stronger H-bonds are formed with the C=O group of acetone than with the H–O group of water for each respective methylimidazole species, and stronger H-bonds are formed with the C=O group of N-methylacetamide than with the C=O group of acetone. When identifying C–H···O interactions of histidine in proteins, if possible, the tautomeric and protonation state should be characterized to help gauge the relative strength of the interaction.

Recognition of the histidine Cδ(C4)–H hydrogen bond motif in proteins is important for protein structure-function analysis, as well as biomolecular modeling. Neutron diffraction structures indicate C4–H (Cδ–H) is capable of coordinating water, thus contributing to solvation layer structure and dynamics. In the active site, the hydrogen bond is anticipated to be capable of participating in catalytic steps, such as proton transfer reactions, hydrolysis and hydrogen bond network formation. Histidine is present in ~50% of enzyme active sites.[8] Furthermore, the C4–H hydrogen bond of histidine (and particularly histidinium) needs to be considered in modeling. Currently, the commonly used force fields in molecular modeling/dynamics do not reflect the C–H bond polarity calculated for histidine, histidinium, and values implied by structural data. Previous QM/MM studies with custom charges on histidine have shown critical interactions of histidine in the active sites, but could be re-evaluated for more information about the C–H hydrogen bonds identified in this work. The Cδ–H hydrogen bond to a main chain carbonyl in an active site histidine of photoactive yellow protein, in addition to Cε–H interactions observed in the Ser-Asp-His catalytic triad of trypsin, and the active sites of human carbonic anhydrase II, RNase A, and myoglobin, suggest that C-H hydrogen bonding interactions in the chemically versatile histidine side chain may make underappreciated contributions to active site protein dynamics, positioning of residues, and formation of hydrogen bonding networks that are critical for enzymatic activity.

7. DATA AVAILABILITY

All data supporting the findings of this study are available within the article and in the ESI. Geometric (x,y,z) coordinates for equilibrium geometries are provided in the ESI.

8. AUTHOR CONTRIBUTIONS

RMS, CK, and KRMK devised the above methodology to investigate these H-bonding interactions. RMS and CK performed the computational investigation using the software packages listed above. MEH, CK, and KRMK visualized neutron diffraction structures using PyMOL software and compiled the H-bonding data for relevant histidine residues. KRMK and RMS curated data, performed the formal analysis, and wrote the original draft of this manuscript. All authors have contributed to the review and editing process and have approved the submitted version of this manuscript. KRMK helped acquire resources and funding for this project.

9. CONFLICTS OF INTEREST

There are no conflicts of interest to declare.

10. ACKNOWLEDGEMENTS

We would like to thank Jonathan M. Ellis for his insightful comments during the preparation of this manuscript. We would also like to thank Kevin Langenwalter for our in-house computer assembly and G09

installation. Lastly, we would like to thank WSU High Performance Computing Center (HiPeCC) and WSU BeoShock High Performance Computing for providing the necessary resources for electronic structure calculations. This material is based upon work supported by National Science Foundation under Grant CHE-1665157.

11. REFERENCES

[1] Shefter, E.; Barlow, M.; Sparks, R.; Trueblood, K. The crystal and molecular structure of β-adenosine-2'-β-uridine-5'-phosphoric acid. *J. Am. Chem. Soc.* **1964**, 86, 1871–1874.

[2] Sussman, J. L.; Seeman, N. C.; Kim, S. H.; Berman, H. M. Crystal structure of a naturally occurring dinucleoside phosphate: uridylyl 3',5'-adenosine phosphate model for RNA chain folding. *J. Mol. Biol.* **1972**, 66, 403– 421.

[3] Brandl, M.; Lindauer, K.; Meyer, M.; Sühnel, J. C-H···O and C-H···N interactions in RNA structures. *Theor. Chem Acc.* **1999**, 101, 103-113.

[4] Derewenda, Z. S.; Lee, L.; and Derewenda, U. The occurrence of C-H···O hydrogen bonds in proteins. *J. Mol. Biol.* **1995**, 252 (2), 248-262.

[5] Newberry, R. W.; Raines, R. T. A Prevalent Intraresidue Hydrogen Bond Stabilizes Proteins. *Nat. Chem. Biol.* **2016**, 12**,** 1084–1088.

[6] Horowitz, S.; Trievel, R. C. Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function. *J. Biol. Chem.* **2012**, 287, 41576-41582.

[7] Hubbard, R. E.; Haider, M. K. Hydrogen bonds in proteins: role and strength. *eLS* **2010**.

[8] Shimba, N.; Serber, Z.; Ledwidge, R.; Miller, S. M.; Craik, C. S.; Dötsch, V. Quantitative identification of the protonation state of histidines in vitro and in vivo. *Biochemistry* **2003**, 42 (30), 9227-9234.

[9] Kasireddy, C; Ellis, J. M.; Bann, J. G.; Mitchell-Koch, K. R. The Biophysical Probes 2-fluorohistidine and 4-fluorohistidine: Spectroscopic Signatures and Molecular Properties. *Scientific Reports* **2017**, 7, 42651.

[10] Buncel, E.; Clement, O.; Onyido, I. Metal Ion Effects in Isotropic Hydrogen Exchange in Biologically Important Heterocycles. *Acc. Chem. Res.* **2000**, 33, 672-678.

[11] Nanda, V.; Schmiedekamp, A. Are aromatic carbon donor hydrogen bonds linear in proteins?. *Proteins* **2008**, 70, 489-497.

[12] Dahanayake, J. N.; Kasireddy, C.; Ellis, J. M.; Hildebrandt, D.; Hull, O. A.; Karnes, J. P.; Morlan, D.; Mitchell-Koch, K. R. Evaluating Electronic Structure Methods for Accurate Calculation of ^{19}F Chemical Shifts in Fluorinated Amino Acids. *J. Comput. Chem* **2017**, 38 (30), 2605-2617.

[13] Kasireddy, C; Ellis, J. M.; Bann, J. G.; Mitchell-Koch, K. R. Tautomeric Stabilities of 4-fluorohistidine Shed New Light on Mechanistic Experiments with Labeled Ribonuclease A. *Chem. Phys. Lett.* **2016**, 666, 58-61.

[14] Dahanayake, J. N.; Kasireddy, C.; Karnes, J. P.; Verma, R.; Steinert, R. M.; Hildebrandt, D.; Hull, O. A.; Ellis, J. M.; and Mitchell-Koch, K. R. Progress in Our Understanding of ^{19}F Chemical Shifts. Chapter 5 in *Annual Reports on NMR Spectroscopy, Vol. 93*, Graham A. Webb, editor, **2018**, 282-374.

[15] Reynolds, W. F.; Peat, I. R.; Freedman, M. H.; Lyerla, J. R. Determination of Tautomeric Form of the Imidazole Ring of L-Histidine in Basic Solution by Carbon-13 Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **1973**, 95, 2, 328-331.

[16] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr.; J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K.

Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc.; Wallingford CT, **2016**.

[17] Dennington, R.; Keith, T. A.; Millam, J. M. GaussView, Version 5.0 Semi-chem Inc.; Shawnee Mission KS, **2009**.

[18] Zhao, Y; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, 120, 215-241.

[19] Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, 98 (2), 1372-1377.

[20] Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **1998**, 102, 1995-2001.

[21] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comp. Chem.* **2003**, 24, 669-681.

[22] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, 77, 18, 3865-3868.

[23] NBO Version 3.1, E. D. Glendening, A. E. Reed, J. E. Carpenter, and F. Weinhold.

[24] Hirshfeld, F. L. Bonded Atom Fragments for Describing Molecular Charge Distributions. *Theor. Chim. Acc.* **1977**, 44, 129-138.

[25] Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient Implementation of the Gauge-Independent Atomic Orbital Method for NMR Chemical Shift Calculations. *J. Am. Chem. Soc.* **1990**, 112, 8251-8260.

[26] Keith, T. A.; Bader, R. F. W. Calculation of the magnetic response properties using a continuous set of gauge transformations. *Chem Phys Lett*, **1993**, 210, 223–231.

[27] Vega, D.; Almeida, D. AIM-UC: An application for QTAIM analysis. *J. Comput. Methods Sci. Eng.* **2014**, 14, 131-136.

[28] Goerigk, L.; Mehta, N. A Trip to the Density Functional Theory Zoo: Warnings and Recommendations for the User. *Aust. J. Chem.* **2019**, 72, 563-573.

[29] Ishida T.; Kato S. Role of Asp102 in the Catalytic Relay System of Serine Proteases: A Theoretical Study. *J. Am. Chem. Soc.* **2004**, 126 (22), 7111-7118.

[30] Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science*, **1994**, 266 (5183), 243-247.

[31] Cuchillo, C. M.; Nogués, M. V.; Raines, R. T. Bovine Pancreatic Ribonuclease: Fifty Years of the First Enzymatic Reaction Mechanism. *Biochemistry*, **2011,** 50 (37), 7835-7841.

[32] Tanokura, M. ^{1}H-NMR Study On The Tautomerism of the Imidazole Ring of Histidine Residues: II. Microenvironments of Histidine-12 and Histidine-119 of Bovine Pancreatic Ribonuclease A. *Biochim. et Biophys. Acta* **1983**, 742 (3), 586-596.

[33] Lindskog, S. Structure and Mechanism of Carbonic Anhydrase. *Pharmacol. Ther.* **1997,** 74 (1), 1-20.

[34] Collman, J. P.; Fu, L. Synthetic Models for Hemoglobin and Myoglobin. *Acc. Chem. Res.* **1999,** 32 (6), 455-463.

[35] Schmiedekamp, A.; Nanda, V. Metal-activated Histidine Carbon Donor Hydrogen Bonds Contribute to Metalloprotein Folding and Function. *J. Inorg. Biochem.* **2009**, 103, 1054-1060.

[36] Lockwood, C. W.; Burlat, B.; Cheesman, M. R.; Kern, M.; Simon, J.; Clarke, T. A.; Richardson, D. J.; and Butt, J. N. Resolution of Key Roles for the Distal Pocket Histidine in Cytochrome C Nitrite Reductases. *J. Am. Chem. Soc.* **2015**, 137, 3059- 3068.

[37] Ross, M. R.; White, A. M.; Yu, F.; King, J. T.; Pecoraro, V. L.; Kubarych, K. J. Histidine Orientation Modulates the Structure and Dynamics of a De Novo Metalloenzyme Active Site. *J. Am. Chem. Soc.* **2015,** 137 (32), 10164-10176.

[38] Levantino, M.; Lemke, H. T.; Schirò, G.; Glownia, M.; Cupane, A.; Cammarata, M. Observing Heme Doming in Myoglobin with Femtosecond X-Ray Absorption Spectroscopy. *Struct. Dyn.* **2015**, 2 (4), 041713.

[39] Langan, P.; Sangha, A. K.; Wymore, T.; Parks, J. M.; Yang, Z. K.; Hanson, B. L.; Fisher, Z.; Mason, S. A.; Blakeley, M. P.; Forsyth, V. T.; Glusker, J. P.; Carrell, H. L.; Smith, J. C.; Keen, D. A.; Graham, D. E.; Kovalevsky, A. L-Arabinose Binding, Isomerization, and Epimerization by D-Xylose Isomerase: X-Ray/Neutron Crystallographic and Molecular Simulation Study. *Structure* **2014**, 22, 9, 1287-1300.

[40] Ostermann, A.; Tanaka, I.; Engler, N.; Niimura, N.; Parak, F. G. Hydrogen and deuterium in myoglobin as seen by a neutron structure determination at 1.5 Å resolution. *Biophys. Chem.* **2002**, 95, 183-193.

[41] Cheng, X.; Schoenborn, B. P. Hydration in Protein Crystals. A Neutron Diffraction Analysis of Carbonmonoxymyoglobin. *Acta. Cryst.* **1990**, B46, 195-208.

[42] Michalczyk, R.; Unkefer, C. J.; Bacik, J. P.; Schrader, T. E.; Ostermann, A.; Kovalevsky, A. Y.; McKenna, R.; Fisher, S. Z. Joint Neutron Crystallographic and NMR Solution Studies of Tyr Residue Ionization and Hydrogen Bonding: Implications for Enzyme-mediated Proton Transfer. *Proc. Natl. Acad. Sci. USA* **2015**, 112 (18), 5673-5678.

[43] Knihtila, R.; Holzapfel, G.; Weiss, K.; Meilleur, F.; Mattos, C. Neutron Crystal Structure of RAS GTPase Puts in Question the Protonation State of the GTP gamma-Phosphate. *J. Biol. Chem.* **2015**, 290, 31025-31036.

[44] Vandavasi, V. G.; Weiss, K. L.; Cooper, J. B.; Erskine, P. T.; Tomanicek, S. J.; Ostermann, A.; Schrader, T. E.; Ginell, S. L.; Coates, L. Exploring the Mechanism of Beta-Lactam Ring Protonation in the Class a Beta-Lactamase Acylation Mechanism Using Neutron and X-Ray Crystallography. *J. Med. Chem.* **2016**, 59, 474-479.

[45] Howard, E. I.; Guillot, B.; Blakeley, M. P.; Haertlein, M.; Moulin, M.; Mitschler, A.; Cousido-Siah, A.; Fadel, F.; Valsecchi, W. M.; Tomizaki, T.; Petrova, T.; Claudot, J.; Podjarny, A. High-resolution neutron and X-ray diffraction room-temperature studies of an H-FABP-oleic acid complex: study of the internal water cluster and ligand binding by a transferred multipolar electron-density distribution. *IUCrJ* **2016**, 3, 115-126.

[46] Yamaguchi, S.; Kamikubo, H.; Kurihara, K.; Kuroki, R.; Niimura, N.; Shimizu, N.; Yamazaki, Y.; Kataoka, M. Low-barrier Hydrogen Bond in Photoactive Yellow Protein. *Proc. Natl. Acad. Sci. USA* **2009**, 106 (2), 440-444.

[47] Adachi, M.; Ohhara, T.; Kurihara, K.; Tamada, T.; Honjo, E.; Okazaki, N.; Arai, S.; Shoyama, Y.; Kimura, K.; Matsumura, H.; Sugiyama, S.; Adachi, H.; Takano, K.; Mori, Y.; Hidaka, K.; Kimura, T.; Hayashi, Y.; Kiso, Y.; Kuroki, R. Structure of HIV-1 Protease in Complex with Potent Inhibitor KNI-272 Determined by High-Resolution X-Ray and Neutron Crystallography. *Proc Natl Acad Sci USA* **2009**, 106 (12), 4641-4646.

[48] Yagi, D.; Yamada, T.; Kurihara, K.; Ohnishi, Y.; Yamashita, M.; Tamada, T.; Tanaka, I.; Kuroki, R.; Niimura, N. A Neutron Crystallographic Analysis of Phosphate-Free Ribonuclease A at 1.7 A Resolution. *Acta. Cryst.* **2009**, D65, 892-899.

[49] Yonezawa, K.; Shimizu, N.; Kurihara, K.; Yamazaki, Y.; Kamikubo, H.; Kataoka, M. Neutron Crystallography of Photoactive Yellow Protein Reveals Unusual Protonation State of Arg52 in the Crystal. *Sci. Rep.* **2017**, 7, 9361.

[50] Schiebel, J.; Gaspari, R.; Sandner, A.; Ngo, K.; Gerber, H. D.; Cavalli, A.; Ostermann, A.; Heine, A.; Klebe, G. Charges Shift Protonation: Neutron Diffraction Reveals that Aniline and 2-Aminopyridine Become Protonated Upon Binding to Trypsin. *Angew. Chem. Int. Ed. Engl.* **2017**, 56, 4887-4890.

[51] Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; Kjaergaard, H. G.; Legon, A. C.; Mennucci, B.; Nesbitt, D. J. Definition Of The

Hydrogen Bond - Recommendation submitted by the IUPAC task group (2004-026-2-100). *Pure Appl. Chem.* **2011,** 83 (8), 1637-1641.

[52] Koch, U. and Popelier, P. L. A. Characterization of C-H-O Hydrogen Bonds on the Basis of Charge Density. *J. Phys. Chem.* **1995**, 99, 9747-9754.

[53] Alvarez, S. A Cartography of the van der Waals Territories. *Dalton Trans.* **2013**, 42, 8617-8636.

[54] Gutmann, V. *The Donor-Acceptor Approach to Molecular Interactions*, Plenum Press, New York, **1978**.

[55] Carreira, L. A.; Person, W. B. Some INDO Calculations of Properties of Molecular Complexes. *J. Am. Chem. Soc.* **1972**, 94, 1485-1495.

[56] Leroy, G.; Louterman-Leloup, G. Étude théorique des liaisons intermoleculaires par "transfert de charge" II. La liaison eau β chlore. *J. Mol. Structure* **1975**, 28, 33-44.

[57] Gutmann, V. *Rev. Chim. Roum.* **1977**, 22, 679.

[58] Denmark, S. E.; Beutner, G. L. Lewis Base Catalysis in Organic Synthesis. *Angew. Chem. Int. Ed.* **2008**, 47, 1560–1638.

[59] Wiberg, K. B.; Rablen, P. R. Atomic Charges. *J. Org. Chem.* **2018**, 83, 15463-15469.

[60] Hobza, P.; Havlas, Z. Blue-Shifting Hydrogen Bonds. *Chem. Rev.* **2000**, 100, 4253-4264.

[61] Chang, X.; Zhang, Y.; Weng, X.; Su, P.; Wu, W.; Mo, Y. Red-Shifting versus Blue-Shifting Hydrogen Bonds: Perspective from Ab Initio Valence Bond Theory. *J. Phys. Chem. A* **2016,** 120, 2749-2756.

[62] Pierens, G. K. ^{1}H and ^{13}C NMR Scaling Factors for the Calculation of Chemical Shifts in Commonly Used Solvents Using Density Functional Theory. *J. Comp. Chem.* **2014**, 35, 18, 1388-1394.

[63] Toomsalu, E.; Burk, P. Critical test of some computational methods for prediction of NMR ^{1}H and ^{13}C chemical shifts. *J Mol Model* **2015**, 21, 244, 1-21.

[64] Tomasi, J.; Mennucci, B.; Cancés, E. The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *Journal of Molecular Structure: THEOCHEM* **1999**, 464, 211-226.

[65] Scheiner, S.; Kar, T.; Pattanayak, J. Comparison of Various Types of Hydrogen Bonds Involving Aromatic Amino Acids. *J. Am. Chem. Soc.* **2002**, 124, 13257-13264.

[66] Anslyn, E. V.; Dougherty, D. A.; *Modern Physical Organic Chemistry*, University Science Books, Melville, NY, **2006**, 165.

[67] Li, L.; Li, C.; Zhang, Z; Alexov, E. On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *J. Chem. Theory Comput.* **2013**, 9 (4), 2126-2136.

[68] Wong, M. W.; Wiberg, K. B; Frisch, M. J. Ab Initio Calculation of Molecular Volumes: Comparison with Experiment and Use in Solvation Models. *J. Comput. Chem.* **1995**, 16, 385-394.

[69] Zhou Y; Zhang Y. Serine protease acylation proceeds with a subtle re-orientation of the histidine ring at the tetrahedral intermediate. *Chem. Commun*. **2011,** 47 (5), 1577-1579.

[70] Shiro, Y.; Iizuka, T.; Marubayashi, K.; Ogura, T.; Kitagawa, T.; Balasubramanian, S.; Boxer, S. G. Spectroscopic study of Ser92 mutants of human myoglobin: hydrogen bonding effect of Ser92 to proximal His93 on structure and property of myoglobin. *Biochemistry* **1994,** 33 (50), 14986-92.

[71] Némethy, G.; Steinberg, I. Z.; Scheraga, H. A. Influence of water structure and of hydrophobic interactions on the strength of side-chain hydrogen bonds in proteins. *Biopolymers* **1963**, 1 (1), 43-69.

[72] Ansari, A.; Berendzen, J.; Bowne, S.F.; Frauenfelder, H.; Iben, I.E.; Sauke, T.B.; Shyamsunder, E.; Young, R.D.; 1985. Protein states and proteinquakes. *Proc. Natl. Acad. Sci.* **1985**, 82 (15), 5000-5004.

[73] Toba, S.; Colombo, G.; Merz, K.M. Solvent dynamics and mechanism of proton transfer in human carbonic anhydrase II. *J. Am. Chem. Soc.* **1999**, 121 (10), 2290-2302.

[74] Merz Jr, K.M. Determination of pKas of ionizable groups in proteins: the pKa of Glu 7 and 35 in hen eggs white lysozyme and Glu 106 in human carbonic anhydrase II. *J. Am. Chem. Soc.* **1991,** 113 (9), 3572-3575.

[75] Merski M.; Skrzeczkowski J.; Roth J.K.; Górna M.W. A geometric definition of short to medium range hydrogen-mediated interactions in proteins. *Molecules* **2020***,* 25, 5326.

[76] Mundlapati, V. R.; Imani, Z.; D'mello, V. C.; Brenner, V.; Gloaguen, E.; Baltaze, J. P.; Robin, S.; Mons, M.; Aitken, D. J. N–H⋯ X interactions stabilize intra-residue C5 hydrogen bonded conformations in heterocyclic α-amino acid derivatives. *Chem. Sci.* **2021**, 12 (44), 14826-14832.

[77] Imani, Z.; Mundlapati, V. R.; Goldsztejn, G.; Brenner, V.; Gloaguen, E.; Guillot, R.; Baltaze, J. P.; Le Barbu-Debus. K.; Robin, S.; Zehnacker, A.; Mons, M.; Aitken, D.J. Conformation control through concurrent N-H⋯S and N-H⋯O=C hydrogen bonding and hyperconjugation effects. *Chem Sci.* **2020**, 11 (34), 9191-9197.

[78] Kumar, S.; Mishra, K. K.; Singh, S. K.; Borish, K.; Dey, S.; Sarkar, B.; Das, A. Observation of a weak intra-residue C5 hydrogen-bond in a dipeptide containing Gly-Pro sequence. *J Chem Phys*. **2019** 151 (10), 104309.

[79] Scheiner, S. Contributions of NH⋯O and CH⋯O Hydrogen Bonds to the Stability of β-Sheets in Proteins

J. Phys. Chem. B 2006, 110 (37), 18670–18679.

[80] Langkilde, A.; Kristensen, S. M.; Lo Leggio, L.; Mølgaard, A.; Jensen, J. H.; Houk, A. R.; Navarro Poulsen, J. C.; Kauppinen, S.; Larsen, S. Short strong hydrogen bonds in proteins: a case study of rhamnogalacturonan acetylesterase. *Acta Cryst. D* **2008,** 64 (8), 851-863.

[81] Jana, B.; Pal, S.; Bagchi, B. Hydration dynamics of protein molecules in aqueous solution: Unity among diversity. *J. Chem. Sci.* **2012,** 124 (1), 317-325.

[82] Adhikari, A.; Park, W. W.; Kwon, O. H. Hydrogen-Bond Dynamics and Energetics of Biological Water. *ChemPlusChem*, **2020,** 85 (12), 2657-2665.

[83] Sterpone, F.; Stirnemann, G.; Hynes, J. T.; Laage, D. Water hydrogen-bond dynamics around amino acids: the key role of hydrophilic hydrogen-bond acceptor groups. *J. Phys. Chem. B*, **2010,** 114 (5), 2083-2089.

[84] Charkhesht, A.; Regmi, C. K.; Mitchell-Koch, K. R.; Cheng, S.; Vinh, N. Q. High-precision megahertz-to-terahertz dielectric spectroscopy of protein collective motions and hydration dynamics. *J. Phys. Chem. B*, **2018,** 122 (24), 6341-6350.

[85] Tarek, M.; Tobias, D. J. Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys. Rev. Lett.* **2002,** 88 (13), 138101.

[86] Dahanayake, J. N.; Mitchell-Koch, K. R. How does solvation layer mobility affect protein structural dynamics?. *Front. Mol. Biosci.* **2018**, 5, 65.