

Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function

Published, JBC Papers in Press, October 9, 2012, DOI 10.1074/jbc.R112.418574

Scott Horowitz¹ and Raymond C. Trievel^{5,2}

From the Departments of ¹Biophysics and ⁵Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109

Carbon-oxygen (CH \cdots O) hydrogen bonding represents an unusual category of molecular interactions first documented in biological structures over 4 decades ago. Although CH \cdots O hydrogen bonding has remained generally underappreciated in the biochemical literature, studies over the last 15 years have begun to yield direct evidence of these interactions in biological systems. In this minireview, we provide a historical context of biological CH \cdots O hydrogen bonding and summarize some major advancements from experimental studies over the past several years that have elucidated the importance, prevalence, and functions of these interactions. In particular, we examine the impact of CH \cdots O bonds on protein and nucleic acid structure, molecular recognition, and enzyme catalysis and conclude by exploring overarching themes and unresolved questions regarding unconventional interactions in biomolecular structure.

Introduction and Historical Perspective

Conventional hydrogen bonds (NH \cdots O, OH \cdots O, OH \cdots N, and NH \cdots N) represent fundamental stabilizing forces in biomolecular structure. Traditionally, carbon has not been considered a conventional hydrogen bond donor due to its relatively low electronegativity compared with oxygen and nitrogen. However, several studies have illustrated that even aliphatic carbon atoms are capable of forming weak hydrogen bonds, which are denoted as CH \cdots O hydrogen bonds (1, 2). In contrast, with increased polarization due to adjacent atoms, carbon atoms can theoretically participate in hydrogen bonds as strong as those formed by conventional donors, specifically oxygen or nitrogen (3, 4).

Many authors have pointed out that it is difficult to define a hydrogen bond, as it is a class of interactions that exhibit varied properties and behavior. One useful definition posits that hydrogen bonding occurs between a proton donor group D-H, where D can be any electronegative element, and an acceptor group that is either a lone pair of electrons or a π bond (5). In these interactions, the hydrogen is shared between the donor and acceptor to varying degrees. The extent of this sharing often dictates the properties of the hydrogen bond, leading to a wide range of hydrogen bond strengths and geometries. Experimental evaluation of hydrogen atom sharing in a biomolecular

system is somewhat challenging, so it has become commonplace to use distance and angular criteria to define a hydrogen bond. Hydrogen bonds tend toward linearity and optimal overlap between the lone pair of the hydrogen bond acceptor and the hydrogen atom. Typically, the sharing of the hydrogen atom allows the hydrogen bond acceptor and donor to encroach to within distances that would otherwise cause steric clashes. Thus, the most commonly used method for discovering hydrogen bond interactions is to examine the hydrogen bond length between the donor and acceptor groups (Fig. 1). Distances that equal less than that of the sum of the atoms' van der Waals radii often indicate hydrogen bond formation. Spectroscopic signatures can also be used to characterize hydrogen bonding. Analogous to conventional hydrogen bonds, CH \cdots O bonds cause a substantial downfield ¹H chemical shift change (6). In infrared spectroscopy, these interactions are unusual in that they usually cause a blue infrared shift, indicative of C-H bond shortening as opposed to the typical bond lengthening observed in conventional hydrogen bonds (7, 8). Despite this difference, the literature on this subject has reached a consensus that the CH \cdots O interaction represents a *bona fide* hydrogen bond (9, 10).

The emergence of CH \cdots O hydrogen bonding as an important interaction in biological structure and function stems from research dating back several decades. As the purpose of this minireview is to focus on recent experimental discoveries, we recommend the authoritative book on the history of weak hydrogen bonding by Desiraju and Steiner (11) for a history of the early years of CH \cdots O hydrogen bonding research. With respect to early work on biological CH \cdots O hydrogen bonds, the review by Wahl and Sundaralingam is also recommended (12). However, certain landmark studies leading to our current understanding of biological CH \cdots O bonding merit discussion here. Notably, studies by Ramachandran (13, 14) and Krimm (7, 15) in the 1960s were among the first to illuminate the contributions of these interactions to protein structure. In more recent work, Derewenda *et al.* (16) catalogued the ubiquitous nature of backbone C α donor hydrogen bonds in proteins based on a survey of 13 high-resolution crystal structures. Using van der Waals distance cutoffs, their survey identified that a surprisingly high percentage of C α \cdots O contacts form CH \cdots O hydrogen bonds in these proteins. The mean distance calculated for all C \cdots O contacts in the interactions surveyed was 3.5 Å, well within the van der Waals distance cutoff of 3.7 Å (Fig. 1). Using C \cdots C interactions as a reference, they were able to clearly demonstrate the widespread nature of CH \cdots O hydrogen bonding in proteins (Fig. 2), especially in the standard backbone hydrogen bonding pattern of β -sheets (Fig. 2A). In addition, CH \cdots O hydrogen bonding has been recognized in nucleic acids, dating back to early crystal structures that identified surprisingly close contact distances between purine C8 and pyrimidine C6 atoms and phosphate backbone oxygen atoms (12).

Since these early studies, many theoretical and experimental studies have endeavored to elucidate the breadth, scope, and importance of CH \cdots O hydrogen bonding in biomolecular structures. As some of the recent computational discoveries in this

¹ Supported by a University of Michigan Rackham predoctoral fellowship.

² Supported by National Science Foundation Grant CHE-1213484. To whom correspondence should be addressed. E-mail: rtrievel@umich.edu.

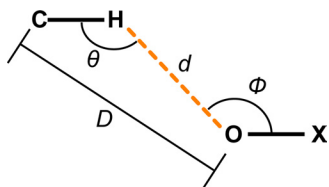


FIGURE 1. Distance and angular parameters used when defining CH \cdots O hydrogen bonds. Typical van der Waals distances d (2.7 Å) and D (3.7 Å) are frequently used as distance cutoffs for hydrogen bond identification.

field have been reviewed recently (3), we will focus on major experimental studies within the past 15 years that have advanced our understanding of biological CH \cdots O hydrogen bonding, with a specific emphasis on its contributions to protein and nucleic acid structure, molecular recognition, and enzyme catalysis.

Contributions to Protein Structure

Characterization of CH \cdots O hydrogen bonds in protein structure spans several decades (7, 13, 16, 18). The aforementioned study by Derewenda *et al.* (16) has led to widespread acceptance of C α -H α \cdots O=C hydrogen bonds, especially main chain interactions within β -sheet structures. More recently, x-ray crystallography and NMR spectroscopy have validated the existence of these interactions in proteins. A number of ultra-high-resolution (<1.0 Å) x-ray structures have allowed direct visualization of hydrogen positions, permitting elucidation of hydrogen bonding patterns within proteins. Many of these studies have attempted to define CH \cdots O bonding patterns within protein structures (19–21) and established unequivocal evidence for CH \cdots O hydrogen bond formation in parallel and antiparallel β -sheets. In fact, it was determined that the idealized position of H α atoms in β -sheet structure was rarely observed, as the H α atom was frequently displaced 0.2–0.3 Å away from its idealized position to increase its CH \cdots O hydrogen bonding potential (21). These findings were further substantiated using NMR spectroscopy through scalar and quadrupolar coupling measurements. In 2003, Grzesiek and colleagues (22) employed long-range scalar coupling experiments to examine C α -H α \cdots O=C interactions in the immunoglobulin-binding domain of protein G. In this study, the authors demonstrated that, analogous to conventional hydrogen bonds, magnetization could be transferred via scalar couplings across CH \cdots O hydrogen bonds in the context of a folded protein, providing direct evidence of hydrogen bond formation (22). Other NMR evidence has more recently come from the measurement of H α quadrupolar coupling constants. The measured constants in ubiquitin revealed variability in the quadrupolar coupling magnitude and that the lowest set of couplings corresponded to residues that were predicted to form C α -H α \cdots O=C bonds based on distance (23). As quadrupolar coupling constants are dictated by the shape of the electron density surrounding the nucleus of interest, the decreased quadrupolar coupling constant magnitudes were attributed to increased electronic symmetry due to hydrogen bonding. These studies have demonstrated with certainty that C α -H α \cdots O=C hydrogen bonds are highly prevalent in protein structure and should be considered a building block of secondary and tertiary structure.

Beyond backbone interactions, there is also experimental evidence for CH \cdots O hydrogen bonds involving amino acid side chains in protein structure. For example, histidine side chains have been implicated in CH \cdots O hydrogen bonding (20, 24) and are predicted to form interactions as strong as conventional hydrogen bonds when the imidazole group is protonated or bound to a metal ion (25). Additional evidence has emerged from neutron crystallography, an evolving technology that holds great promise for directly visualizing CH \cdots O hydrogen bonds. Unlike x-rays, neutrons are diffracted by atomic nuclei, enabling complementary structural information and, most importantly, direct visualization of hydrogen or deuterium atom positions (26). Over the past decade, the number of neutron structures in the Protein Data Bank has increased dramatically, providing many unprecedented snapshots of protein hydrogen bonding. Most notably, one recent study endeavored to analyze the prevalence of CH \cdots O hydrogen bonding in amicyanin, a cupredoxin that binds copper in bacteria. By using joint x-ray and neutron refinement, the investigators were able to obtain a high-resolution structure with excellent visualization of hydrogen atoms. By analyzing the hydrogen positions, the authors observed a remarkable 27 CH \cdots O hydrogen bonds in the copper coordination site (Fig. 2C) in addition to eight conventional hydrogen bonds (27). As neutron crystallography continues to evolve, we anticipate that future studies will illuminate many new and biologically relevant CH \cdots O hydrogen bond networks in protein structure.

CH \cdots O Hydrogen Bonding in Nucleic Acid Structure

Although many recent advances in understanding CH \cdots O hydrogen bonding have arisen from studies of protein structure, these interactions have long been appreciated in nucleic acid structure. In early crystal structures of single nucleotides, it was noted that multiple CH \cdots O hydrogen bonds were apparent, especially between phosphate backbone oxygens and the C6 atom of pyrimidines and the C8 atom of purines, and that these interactions were likely stabilized by the *anti*-conformation (28, 29). NMR and Raman spectroscopy later corroborated the claim that the acidity of the C8 atom in purine rings poises it for CH \cdots O hydrogen bonding in RNA duplexes (30). Finally, a survey of high-resolution RNA structures (<2.0 Å) illustrated that short CH \cdots O contacts are ubiquitous and appear to stabilize RNA tertiary structure (31). Despite these crystallographic findings and multiple studies implicating CH \cdots O hydrogen bond formation in nucleic acids (32–35), the quantity of direct evidence of CH \cdots O hydrogen bonding in nucleic acids is comparatively smaller than that in proteins. In part, this dearth is due to a general lack of neutron or high-resolution x-ray crystal structures of RNA. Of those few RNA structures solved to sufficiently high resolution to visualize hydrogen atoms (<1.0 Å), only in one study of an RNA tetraplex do the authors explore the possibility of CH \cdots O hydrogen bonding (Fig. 3C) (36). However, this structure does not address many common RNA secondary and tertiary structural elements. Although the formation of CH \cdots O hydrogen bonds in A-T base pairs, both in Watson-Crick and Hoogsteen conformation, has been debated in the literature (37–40), experiments have yet to definitively resolve whether these interactions represent true hydrogen

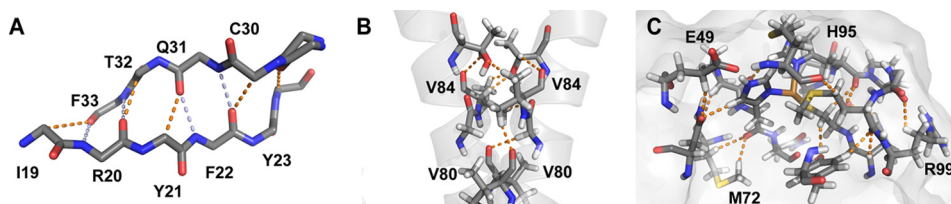


FIGURE 2. Examples of CH...O hydrogen bonds (orange dashes) in proteins. A, typical hydrogen bonding pattern in β -sheets (21). B and C, extensive CH...O hydrogen bonding in packed α -helices (72) and the copper coordination site in amicyanin (27), respectively.

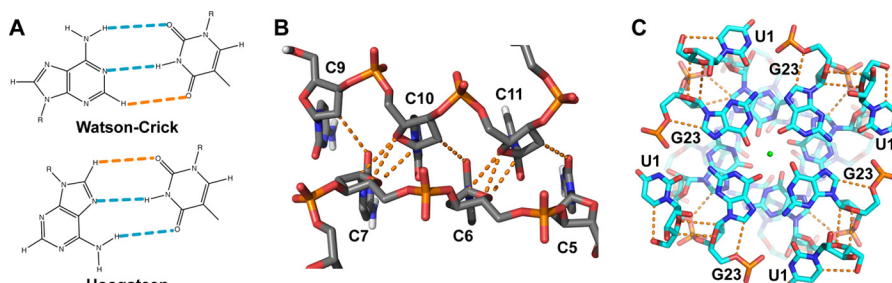


FIGURE 3. Examples of CH...O hydrogen bonds (orange dashes) in nucleic acids. A, hydrogen bonding patterns in Watson-Crick (upper) and Hoogsteen (lower) adenine-thymidine base pairs. B and C, extensive CH...O hydrogen bonding in the DNA i-motif (73) and the RNA tetraplex (36), respectively.

bonds that stabilize DNA and RNA structure. Corollary studies to those in proteins aimed at directly probing CH...O hydrogen bond formation in base pairs, backbone interactions, and tertiary structure represent a promising area of future research.

CH...O Hydrogen Bonding in Molecular Recognition

Evidence and importance of CH...O hydrogen bonding in molecular recognition are also a current focus of research. These hydrogen bonds have been implicated in many intermolecular interactions, including those involving protein-protein, protein-ligand, and protein-nucleic acid complexes. Although relatively few studies have analyzed CH...O hydrogen bonding in protein-protein complexes, a recently determined x-ray structure at 0.9 Å resolution of the ATRX ADD domain bound to a histone H3 peptide bearing a trimethylated Lys-9 in which the methyl hydrogen density is clearly visible illustrates that the lysine trimethylammonium cation is specifically recognized through a network of methyl CH...O hydrogen bonds (Fig. 4A) (41). Similarly, in a recently determined neutron structure of transthyretin, the authors noted that the ratio of CH...O to conventional hydrogen bonds in the interface between the A and D subunits is >2:1 (42), potentially providing substantial binding energy.

With respect to protein-ligand binding, Klaholz and Moras (43) were among the first to use x-ray crystallography to analyze the nuclear retinoic acid receptor- γ bound to retinoid SR11254 to identify CH...O hydrogen bonds involved in ligand recognition. Along with several protein-protein CH...O interactions, they discovered numerous CH...O bonds between different moieties of the ligand with hydrogen bond donors and acceptors at several different positions, suggesting that multiple CH...O interactions may significantly contribute to ligand binding affinity (43). Similar to neutron studies of protein-protein interactions, neutron diffraction studies have also provided evidence for these interactions between proteins and water and ligand molecules. As observed in the neutron structure of xylulose-isomerase, a CH...O hydrogen bond between the protein

His-220 C1 and O2 of xylulose may facilitate xylulose recognition (44). To address the possibility of protein-solvent CH...O bonds, one neutron crystallographic study analyzed the protein-solvent interface of lysozyme. Surprisingly, large sections of its surface interacted with water exclusively through CH...O hydrogen bonding (45), suggesting important roles for these interactions in protein solubility and folding.

Protein-nucleic acid recognition is fundamental to myriad biological processes, particularly those involving DNA transactions. Due to the relatively fixed nature of the DNA duplex, both conventional and CH...O hydrogen bond interactions are formed in concert in relatively predictable patterns. In a survey of protein-DNA crystal structures, it was found that thymine and cytosine appear to consistently form CH...O hydrogen bonds when in complex (46). In total, the distribution of C...O distances in protein-nucleic acid complexes was more consistent with the distribution of conventional N...O hydrogen bond distances than control C...C distances. Interestingly, the thymine methyl group is the most frequent CH...O hydrogen bond donor, despite its lower polarization compared with the nucleobase aromatic carbon atoms. Despite the promising nature of this early work, more neutron and ultra-high-resolution x-ray structures that specifically probe the importance and nature of CH...O hydrogen bonds in protein-nucleic acid complexes are required. Determining the importance and function of CH...O hydrogen bonds in protein-nucleic acid complexes remains an exciting future avenue of research.

Experimental Measurement of CH...O Hydrogen Bond Strength

Although obtaining evidence of CH...O hydrogen bond formation in biomolecular structures is informative, it does not address the strength or biological importance of these interactions. Much of the insight gained thus far on the relevance of these interactions derives from computational studies of the relative strengths of CH...O bonds in protein structures. For details on these studies, we suggest previous reviews on com-

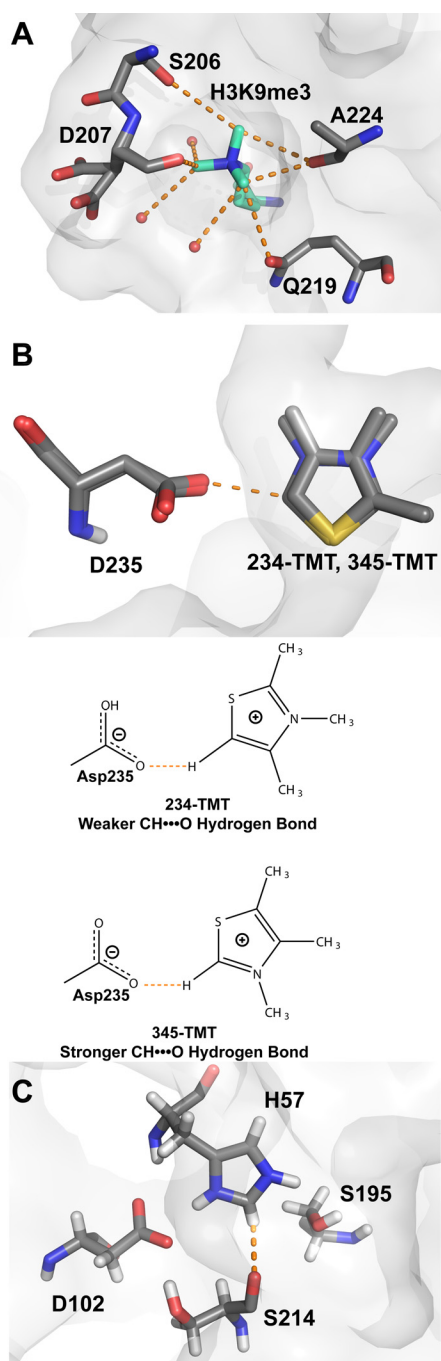


FIGURE 4. Examples of CH \cdots O hydrogen bonds (orange dashes) in molecular recognition and enzyme catalysis. A, recognition of trimethyllysine by CH \cdots O hydrogen bonds in the ATRX ADD domain (41). B, schematic depiction of 234-TMT (upper) and 345-TMT (lower) binding to an engineered protein active site (54). C, conserved serine hydrolase CH \cdots O hydrogen bond between the catalytic histidine and serine residues (17).

putational studies of biological CH \cdots O bonds (3, 47). These studies have generally shown CH \cdots O bonds to be weak interactions, typically exhibiting half the bond energy of a conventional hydrogen bond in gas-phase calculations. However, it is noteworthy that these studies have consistently found that not only are these interactions fundamental in protein structure, but in some instances, they can be as strong as conventional hydrogen bonds (3). In one striking example, it was estimated that 17% of all protein-protein surface interaction energy arises

from CH \cdots O hydrogen bonding and that, in some extreme cases, this percentage can be as high as 40–50% (48). Although CH \cdots O hydrogen bonds are often weak, they frequently occur in greater quantities than their conventional counterparts and thus may contribute significantly to protein-protein interaction energies.

In addition to a solid computational foundation, a few studies have provided direct experimental evidence of CH \cdots O hydrogen bond strength and importance in proteins. By a combination of high-pressure infrared spectroscopy and x-ray crystallography, CH \cdots O bonds were found to be crucial in artificial β -sheet-like network formation (49), and it is possible that these interactions are of similar importance in true β -sheets. In a contrasting example, Bowie and co-workers (50) determined that a C α -H α \cdots O hydrogen bond in bacteriorhodopsin did not provide any protein stabilization, although subsequent studies demonstrated that conventional hydrogen bonding in bacteriorhodopsin was also surprisingly weak (51). However, a separate concurrent study using infrared spectroscopy discovered a substantially stronger protein CH \cdots O hydrogen bond, 0.9 kcal/mol (18), comparable with a weak conventional hydrogen bond. Correlatively, Kallenbach and co-workers (52) analyzed helical peptides using circular dichroism to show that a single CH \cdots O hydrogen bond contributed \sim 0.5 kcal/mol helix stabilization energy, but only with certain side chain sequences and orientations. Thus, from experimental evidence, the exact energetic stabilization contributed by CH \cdots O hydrogen bonds to protein folding remains ambiguous. As in the case of conventional interactions, the contribution of a single CH \cdots O hydrogen bond to protein stabilization depends on its hybridization and polarization as well as its role and position in protein folding (53). Further studies are needed to precisely define the energetic values of CH \cdots O hydrogen bonds and their contributions to protein folding.

In comparison with the foregoing examples, relatively few studies have experimentally explored CH \cdots O hydrogen bond strengths in protein-ligand binding. One notable exception is a study that analyzed the binding of two related compounds, 2,3,4-trimethylthiazole (234-TMT)³ and 3,4,5-trimethylthiazole (345-TMT), to a cytochrome *c* peroxidase mutant (Fig. 4B) (54). These compounds are both cationic, containing a formal positive charge on the nitrogen atom. The only difference between them involves the relative placement of the nitrogen in the aromatic ring system. These nitrogens are methylated and thus cannot directly participate in hydrogen bonding. However, a single carbon atom in the aromatic ring in each compound is in appropriate geometry to form a CH \cdots O hydrogen bond to Asp-235 in the enzyme. The only difference between these two molecules is that, in 234-TMT, the nitrogen is two positions removed from this hydrogen-bonded C–H group, whereas in 345-TMT, it is only one position removed, similar to histidine residues. Due to the proximity of the positively charged nitrogen to the CH \cdots O hydrogen bond that should enhance the carbon polarization, 345-TMT was computationally predicted to

³ The abbreviations used are: 234-TMT, 2,3,4-trimethylthiazole; 345-TMT, 3,4,5-trimethylthiazole; KMT, lysine methyltransferase; AdoMet, S-adenosylmethionine; KDM, lysine demethylase.

form a substantially stronger $\text{CH}\cdots\text{O}$ bond than 234-TMT, accounting for the difference in binding energy between the two ligands. By solving crystal structures of the two complexes, the authors verified that the ligand binding modes were essentially identical (Fig. 4B). Upon measuring the binding constants for each ligand, the authors determined that the binding energy of 345-TMT was 1.2 kcal/mol greater than that of 234-TMT (54), substantiating their computational predictions.

Within nucleic acids, the $\text{CH}\cdots\text{O}$ hydrogen bond strength within the *i*-motif has undergone uniquely in-depth study. This motif is a quadruplex-like structure consisting of four strands containing intercalated $\text{C}\cdot\text{CH}^+$ base pairs and linked loops (Fig. 3B) that may form at the end of telomeres (55). Initial structures of this motif revealed extensive $\text{CH}\cdots\text{O}$ hydrogen bond networks, suggesting that these interactions may promote *i*-motif formation (56, 57). To substantiate these observations, Gueron and co-workers (57) attempted to determine the average strength of $\text{C1}'\text{-H1}'\cdots\text{O}$ bonds using different intercalation topologies that a single sequence could form as a function of ionic strength and temperature. Based on these experiments, the average $\text{C1}'\text{-H1}'\cdots\text{O}$ hydrogen bond strength was measured to be 0.6 kcal/mol. Cumulatively, these interactions likely contribute significant stabilization energy in the *i*-motif.

Given their prevalence in RNA, it is likely that $\text{CH}\cdots\text{O}$ hydrogen bonds contribute substantially to the specificity of RNA folding. Indeed, these interactions have been introduced recently as a parameter in RNA tertiary structure predictions (58), but they have yet to be experimentally evaluated. Similarly, determining the contribution (if any) of $\text{CH}\cdots\text{O}$ hydrogen bonds to DNA base pairing would expand our understanding of their energetic contributions in both base pair separation and formation in processes such as transcription and DNA replication, repair, and recombination. Finally, the ~ 1 kcal/mol hydrogen bond found by Arbely and Arkin (18) is similar in strength to that determined by the previous study of 345-TMT, supporting the notion that these interactions are important not only in protein structure and folding but also in ligand binding.

$\text{CH}\cdots\text{O}$ Hydrogen Bonding in Enzyme Catalysis

In addition to macromolecular structure, $\text{CH}\cdots\text{O}$ hydrogen bonds have been implicated either directly or indirectly in the catalytic mechanisms of several classes of enzymes. One well studied case is serine hydrolases bearing a His-(Asp/Glu)-Ser catalytic triad, including but not limited to serine proteases, lipases, and thioesterases. An early survey of these enzymes by Derewenda *et al.* (59) identified short histidine-oxygen distances between the catalytic histidine C1 and an adjacent carbonyl oxygen that were indicative of $\text{CH}\cdots\text{O}$ hydrogen bonding. This hydrogen bond was proposed to stabilize the imidazolium cation and to potentially facilitate a ring-flipping mechanism (24, 59). This observation of Derewenda *et al.* was further supported by Bachovchin and co-workers (24), who used NMR chemical shift to show that the downfield chemical shift of the histidine H1 proton was consistent with $\text{CH}\cdots\text{O}$ hydrogen bonding. Corroboratively, Hunter and co-workers (60) proposed, based on crystallographic distances, that a $\text{CH}\cdots\text{O}$ bond formed between the catalytic triad histidine and a substrate carbonyl group in trypanothione reductase would presumably

stabilize the positive charge formed on the histidine side chain, facilitating an electronic induced fit mechanism.

Recent studies have also identified functions for $\text{CH}\cdots\text{O}$ hydrogen bonding in methyl transfer reactions. Structural and functional studies of the lysine methyltransferases (KMTs) belonging to the SET domain family revealed that *S*-adenosylmethionine (AdoMet) methyl $\text{CH}\cdots\text{O}$ hydrogen bonds are conserved in these enzymes, indicating a potential role in cofactor binding and catalysis (61). We consider it likely that these interactions aid not only in AdoMet recognition but also in transition state stabilization. These findings have been corroborated by recent NMR spectroscopic studies using ^1H chemical shift and quantum mechanical calculations that quantitatively established $\text{CH}\cdots\text{O}$ hydrogen bonding between the AdoMet methyl group and oxygen atoms in the active site of the human KMT SET7/9 (62). In addition, $\text{CH}\cdots\text{O}$ hydrogen bonds appear to be important in methyllysine binding by both SET domain KMTs and JmjC (Jumonji-C) lysine demethylases (KDMs), analogous to the ATRX ADD domain (Fig. 4A). In SET domain KMTs, structural and mutagenic data indicate that $\text{CH}\cdots\text{O}$ hydrogen bonds are important for the repositioning of the ϵ -amine group to enable lysine multiple methylation (63, 64). In the JmjC KDM JMJD2A, mutational and structural evidence suggests that $\text{CH}\cdots\text{O}$ hydrogen bonds are important in distinguishing between di- and trimethylated lysine. These interactions have also been reported in other JmjC KDMs, including UTX (65), PHF8 (66) and JMJD3 (67). Together, these studies imply that $\text{CH}\cdots\text{O}$ hydrogen bonds are fundamental to every phase of lysine methylation and demethylation.

$\text{CH}\cdots\text{O}$ hydrogen bonding has also been implicated in enzyme acid/base-catalyzed reactions involving carbon atoms. Typically, proton abstraction is preceded by hydrogen bond formation. In the case of proton abstraction from a carbon, a $\text{CH}\cdots\text{O}$ hydrogen bond would therefore be formed prior to proton transfer. For example, in acyl-CoA dehydrogenases, the initial step of catalysis involves proton abstraction from the $\text{C}\alpha$ atom of the acyl-CoA substrate by a glutamate base. Computational investigation of this reaction indicated that a strong ionic $\text{CH}\cdots\text{O}$ hydrogen bond forms along the reaction coordinate prior to proton abstraction from the $\text{C}\alpha$ position in the acyl chain (68). Given the ubiquitous nature of acid/base chemistry in enzyme catalysis, this reaction constitutes a mechanism of particular interest for future experimental investigation that will likely uncover additional roles for $\text{CH}\cdots\text{O}$ hydrogen bonding in many enzyme mechanisms.

Conclusions and Future Directions

In the past 15 years, experiments have provided new insights into many facets regarding the biological functions of $\text{CH}\cdots\text{O}$ hydrogen bonding. With respect to proteins, these interactions have fundamental roles in mediating ligand recognition, enzyme-catalyzed reactions, and macromolecular interactions. Additionally, several studies have yielded experimental data regarding the strength of biological $\text{CH}\cdots\text{O}$ hydrogen bonds, emphasizing their energetic contributions to macromolecular structure. We are only beginning to appreciate the breadth of functions of these interactions in protein structure and function, and many fundamental questions concerning $\text{CH}\cdots\text{O}$

hydrogen bonding in nucleic acids, enzyme catalysis, and protein folding remain unresolved, representing fruitful avenues for future research. Moreover, we expect that several practical applications will emerge from these studies, most notably improved methods for structure-based drug design and optimization and revised computational models for RNA and protein folding that take into account CH \cdots O hydrogen bonding.

Finally, a very recently determined ultra-high-resolution neutron structure of the protein crambin emphasizes that CH \cdots O hydrogen bonds represent only one class of a large category of underappreciated interactions in biomolecules (69). In addition to many CH \cdots O hydrogen bonds, the authors directly observed many π acceptor hydrogen bonds. Other unconventional interactions, including CH \cdots N and CH \cdots S (70), as well as $n\cdots\pi^*$ interactions that are typified in proline residues (71), are even less understood. Similar to CH \cdots O hydrogen bonds, it is conceivable that these interactions also play important roles in macromolecular structure and function that have yet to be fully explored and understood.

Acknowledgments—We thank Samuel Krimm and Noemi Mirkin for reading the manuscript and providing insightful comments.

REFERENCES

- Steiner, T., and Desiraju, G. R. (1998) Distinction between the weak hydrogen bond and the van der Waals interaction. *Chem. Commun.* 891–892
- Steiner, T. (2002) The hydrogen bond in the solid state. *Angew. Chem. Int. Ed. Engl.* **41**, 48–76
- Scheiner, S. (2011) Weak H-bonds. Comparisons of CH \cdots O to NH \cdots O in proteins and PH \cdots N to direct P \cdots N interactions. *Phys. Chem. Chem. Phys.* **13**, 13860–13872
- Cannizzaro, C. E., and Houk, K. N. (2002) Magnitudes and chemical consequences of R₃N⁺–C–H \cdots O=C hydrogen bonding. *J. Am. Chem. Soc.* **124**, 7163–7169
- Gilli, G., and Gilli, P. (2009) *The Nature of the Hydrogen Bond*, Oxford University Press, New York
- Scheiner, S., Gu, Y., and Kar, T. (2000) Evaluation of the H-bonding properties of CH \cdots O interactions based upon NMR spectra. *J. Mol. Struct. Theochem* **500**, 441–452
- Krimm, S., and Kuroiwa, K. (1968) Low temperature infrared spectra of polyglycines and C–H \cdots O=C hydrogen bonding in polyglycine II. *Biopolymers* **6**, 401–407
- Scheiner, S. (2009) Identification of spectroscopic patterns of CH \cdots O H-bonds in proteins. *J. Phys. Chem. B* **113**, 10421–10427
- Gu, Y. L., Kar, T., and Scheiner, S. (1999) Fundamental properties of the CH \cdots O interaction: is it a true hydrogen bond? *J. Am. Chem. Soc.* **121**, 9411–9422
- Desiraju, G. R. (1996) The C–H \cdots O hydrogen bond: structural implications and supramolecular design. *Acc. Chem. Res.* **29**, 441–449
- Desiraju, G. R., and Steiner, T. (1999) *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, New York
- Wahl, M. C., and Sundaralingam, M. (1997) C–H \cdots O hydrogen bonding in biology. *Trends Biochem. Sci.* **22**, 97–102
- Ramachandran, G. N., and Chandrasekharan, R. (1968) Interchain hydrogen bonds via bound water molecules in collagen triple helix. *Biopolymers* **6**, 1649–1658
- Ramachandran, G. N., Sasisekharan, V., and Ramakrishnan, C. (1966) Molecular structure of polyglycine II. *Biochim. Biophys. Acta* **112**, 168–170
- Krimm, S. (1967) Hydrogen bonding of C–H \cdots O=C in proteins. *Science* **158**, 530–531
- Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) The occurrence of C–H \cdots O hydrogen bonds in proteins. *J. Mol. Biol.* **252**, 248–262
- Fuhrmann, C. N., Daugherty, M. D., and Agard, D. A. (2006) Subangstrom crystallography reveals that short ionic hydrogen bonds, and not a His-Asp low-barrier hydrogen bond, stabilize the transition state in serine protease catalysis. *J. Am. Chem. Soc.* **128**, 9086–9102
- Arbely, E., and Arkin, I. T. (2004) Experimental measurement of the strength of a Ca–H \cdots O bond in a lipid bilayer. *J. Am. Chem. Soc.* **126**, 5362–5363
- Esposito, L., Vitagliano, L., Sica, F., Sorrentino, G., Zagari, A., and Mazzarella, L. (2000) The ultrahigh resolution crystal structure of ribonuclease A containing an isoaspartyl residue: hydration and stereochemical analysis. *J. Mol. Biol.* **297**, 713–732
- Sandalova, T., Schneider, G., Käck, H., and Lindqvist, Y. (1999) Structure of dethiobiotin synthetase at 0.97 Å resolution. *Acta Crystallogr. D* **55**, 610–624
- Addlagatta, A., Krzywda, S., Czapinska, H., Otlewski, J., and Jaskolski, M. (2001) Ultrahigh-resolution structure of a BPTI mutant. *Acta Crystallogr. D* **57**, 649–663
- Cordier, F., Barfield, M., and Grzesiek, S. (2003) Direct observation of C α –H \cdots O=C hydrogen bonds in proteins by interresidue ^{h3}J_{C α C'}} scalar couplings. *J. Am. Chem. Soc.* **125**, 15750–15751
- Sheppard, D., Li, D. W., Godoy-Ruiz, R., Brüschweiler, R., and Tugarinov, V. (2010) Variation in quadrupole couplings of α deuterons in ubiquitin suggests the presence of C α –H \cdots O=C hydrogen bonds. *J. Am. Chem. Soc.* **132**, 7709–7719
- Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G., and Bachovchin, W. W. (2000) Unusual ¹H NMR chemical shifts support (His) C ϵ^1 –H \cdots O=C H-bond: proposal for reaction-driven ring flip mechanism in serine protease catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10371–10376
- Schmiedekamp, A., and Nanda, V. (2009) Metal-activated histidine carbon donor hydrogen bonds contribute to metalloprotein folding and function. *J. Inorg. Biochem.* **103**, 1054–1060
- Blakeley, M. P., Langan, P., Niimura, N., and Podjarny, A. (2008) Neutron crystallography: opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600
- Sukumar, N., Mathews, F. S., Langan, P., and Davidson, V. L. (2010) A joint x-ray and neutron study on amicyanin reveals the role of protein dynamics in electron transfer. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6817–6822
- Shefter, E., Barlow, M., Sparks, R., and Trueblood, K. (1964) The crystal and molecular structure of β -adenosine-2'- β -uridine-5'-phosphoric acid. *J. Am. Chem. Soc.* **86**, 1871–1874
- Sussman, J. L., Seeman, N. C., Kim, S. H., and Berman, H. M. (1972) Crystal structure of a naturally occurring dinucleoside phosphate: uridylyl 3',5'-adenosine phosphate model for RNA chain folding. *J. Mol. Biol.* **66**, 403–421
- Benevides, J. M., and Thomas, G. J. (1988) A solution structure for poly(rA)poly(dT) with different furanose pucker and backbone geometry in rA and dT strands and intrastrand hydrogen bonding of adenine 8CH. *Biochemistry* **27**, 3868–3873
- Brandl, M., Lindauer, K., Meyer, M., and Suhnel, J. (1999) C–H \cdots O and C–H \cdots N interactions in RNA structures. *Theor. Chem. Acc.* **101**, 103–113
- Li, F., Pallan, P. S., Maier, M. A., Rajeev, K. G., Mathieu, S. L., Kreutz, C., Fan, Y., Sanghvi, J., Micura, R., Rozners, E., Manoharan, M., and Egli, M. (2007) Crystal structure, stability and *in vitro* RNAi activity of oligoribonucleotides containing the ribo-difluorotoluyl nucleotide: insights into substrate requirements by the human RISC Ago2 enzyme. *Nucleic Acids Res.* **35**, 6424–6438
- Liu, H., Matsugami, A., Katahira, M., and Uesugi, S. (2002) A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops. *J. Mol. Biol.* **322**, 955–970
- Berger, I., and Egli, M. (1997) The role of backbone oxygen atoms in the organization of nucleic acid tertiary structure: zippers, networks, clamps, and C–H \cdots O hydrogen bonds. *Chem. Eur. J.* **3**, 1400–1404
- Duszczyk, M. M., Wutz, A., Rybin, V., and Sattler, M. (2011) The Xist RNA A-repeat comprises a novel AUCG tetraloop fold and a platform for multimerization. *RNA* **17**, 1973–1982
- Deng, J., Xiong, Y., and Sundaralingam, M. (2001) X-ray analysis of an RNA tetraplex (UGGGGU)₄ with divalent Sr²⁺ ions at subatomic resolu-

- tion (0.61 Å). *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13665–13670
37. Ghosh, A., and Bansal, M. (1999) C–H⁺⋯O hydrogen bonds in minor groove of A-tracts in DNA double helices. *J. Mol. Biol.* **294**, 1149–1158
 38. Castellano, R. K., Gramlich, V., and Diederich, F. (2002) Rebeck imides and their adenine complexes: preferences for Hoogsteen binding in the solid state and in solution. *Chem. Eur. J.* **8**, 118–129
 39. Quinn, J. R., Zimmerman, S. C., Del Bene, J. E., and Shavitt, I. (2007) Does the A·T or G·C base pair possess enhanced stability? Quantifying the effects of CH⁺⋯O interactions and secondary interactions on base-pair stability using a phenomenological analysis and *ab initio* calculations. *J. Am. Chem. Soc.* **129**, 934–941
 40. Zhou, P. P., and Qiu, W. Y. (2009) Red-shifted hydrogen bonds and blue-shifted van der Waals contact in the standard Watson-Crick adenine-thymine base pair. *J. Phys. Chem. A* **113**, 10306–10320
 41. Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P. W., Cochrane, J. C., Allis, C. D., Picketts, D. J., Patel, D. J., Li, H., and Shi, Y. (2011) ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat. Struct. Mol. Biol.* **18**, 769–776
 42. Yokoyama, T., Mizuguchi, M., Nabeshima, Y., Kusaka, K., Yamada, T., Hosoya, T., Ohhara, T., Kurihara, K., Tomoyori, K., Tanaka, I., and Nimura, N. (2012) Hydrogen-bond network and pH sensitivity in transthyretin: neutron crystal structure of human transthyretin. *J. Struct. Biol.* **177**, 283–290
 43. Klaholz, B., and Moras, D. (2002) C–H⁺⋯O hydrogen bonds in the nuclear receptor RAR γ —a potential tool for drug selectivity. *Structure* **10**, 1197–1204
 44. Kovalevsky, A. Y., Katz, A. K., Carrell, H. L., Hanson, L., Mustyakimov, M., Fisher, S. Z., Coates, L., Schoenborn, B. P., Bunick, G. J., Glusker, J. P., and Langan, P. (2008) Hydrogen location in stages of an enzyme-catalyzed reaction: time-of-flight neutron structure of D-xylose isomerase with bound D-xylulose. *Biochemistry* **47**, 7595–7597
 45. Bon, C., Lehmann, M. S., and Wilkinson, C. (1999) Quasi-Laue neutron-diffraction study of the water arrangement in crystals of triclinic hen egg-white lysozyme. *Acta Crystallogr. D* **55**, 978–987
 46. Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L., and Zhurkin, V. B. (1998) A role for CH⁺⋯O interactions in protein-DNA recognition. *J. Mol. Biol.* **277**, 1129–1140
 47. Castellano, R. K. (2004) Progress toward understanding the nature and function of C–H⁺⋯O interactions. *Curr. Org. Chem.* **8**, 845–865
 48. Jiang, L., and Lai, L. (2002) CH⁺⋯O hydrogen bonds at protein-protein interfaces. *J. Biol. Chem.* **277**, 37732–37740
 49. Lee, K. M., Chang, H. C., Jiang, J. C., Chen, J. C., Kao, H. E., Lin, S. H., and Lin, I. J. (2003) C–H—O hydrogen bonds in β -sheetlike networks: combined x-ray crystallography and high-pressure infrared study. *J. Am. Chem. Soc.* **125**, 12358–12364
 50. Yohannan, S., Faham, S., Yang, D., Grosfeld, D., Chamberlain, A. K., and Bowie, J. U. (2004) A C α –H⁺⋯O hydrogen bond in a membrane protein is not stabilizing. *J. Am. Chem. Soc.* **126**, 2284–2285
 51. Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Woods, V. L., and Bowie, J. U. (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* **453**, 1266–1270
 52. Shi, Z., Olson, C. A., Bell, A. J., Jr., and Kallenbach, N. R. (2002) Non-classical helix-stabilizing interactions: C–H⁺⋯O H-bonding between Phe and Glu side chains in α -helical peptides. *Biophys. Chem.* **101**, 267–279
 53. Strop, P., and Mayo, S. L. (2000) Contribution of surface salt bridges to protein stability. *Biochemistry* **39**, 1251–1255
 54. Musah, R. A., Jensen, G. M., Rosenfeld, R. J., McRee, D. E., Goodin, D. B., and Bunte, S. W. (1997) Variation in strength of an unconventional CH–O hydrogen bond in an engineered protein cavity. *J. Am. Chem. Soc.* **119**, 9083–9084
 55. Gehring, K., Leroy, J. L., and Guéron, M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* **363**, 561–565
 56. Berger, I., Egli, M., and Rich, A. (1996) Inter-strand C–H⁺⋯O hydrogen bonds stabilizing four-stranded intercalated molecules: stereoelectronic effects of O4' in cytosine-rich DNA. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12116–12121
 57. Leroy, J. L., Snoussi, K., and Guéron, M. (2001) Investigation of the energetics of C–H⁺⋯O hydrogen bonds in the DNA i-motif via the equilibrium between alternative intercalation topologies. *Magn. Reson. Chem.* **39**, S171–S176
 58. Das, R., Karanicolas, J., and Baker, D. (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **7**, 291–294
 59. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) (His)C ϵ –H⁺⋯O=C hydrogen bond in the active sites of serine hydrolases. *J. Mol. Biol.* **241**, 83–93
 60. Bond, C. S., Zhang, Y., Berriman, M., Cunningham, M. L., Fairlamb, A. H., and Hunter, W. N. (1999) Crystal structure of *Trypanosoma cruzi* trypanothione reductase in complex with trypanothione, and the structure-based discovery of new natural product inhibitors. *Structure* **7**, 81–89
 61. Couture, J. F., Hauk, G., Thompson, M. J., Blackburn, G. M., and Trievel, R. C. (2006) Catalytic roles for carbon-oxygen hydrogen bonding in SET domain lysine methyltransferases. *J. Biol. Chem.* **281**, 19280–19287
 62. Horowitz, S., Yesselman, J. D., Al-Hashimi, H. M., and Trievel, R. C. (2011) Direct evidence for methyl group coordination by carbon-oxygen hydrogen bonds in the lysine methyltransferase SET7/9. *J. Biol. Chem.* **286**, 18658–18663
 63. Couture, J. F., Dirk, L. M., Brunzelle, J. S., Houtz, R. L., and Trievel, R. C. (2008) Structural origins for the product specificity of SET domain protein methyltransferases. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20659–20664
 64. Del Rizzo, P. A., Couture, J. F., Dirk, L. M., Strunk, B. S., Roiko, M. S., Brunzelle, J. S., Houtz, R. L., and Trievel, R. C. (2010) SET7/9 catalytic mutants reveal the role of active site water molecules in lysine multiple methylation. *J. Biol. Chem.* **285**, 31849–31858
 65. Sengoku, T., and Yokoyama, S. (2011) Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *Gene Dev.* **25**, 2266–2277
 66. Horton, J. R., Upadhyay, A. K., Qi, H. H., Zhang, X., Shi, Y., and Cheng, X. (2010) Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* **17**, 38–43
 67. Kruidenier, L., Chung, C. W., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., Eberhard, D., Hutchinson, S., Jones, E., Katso, R., Leveridge, M., Mander, P. K., Mosley, J., Ramirez-Molina, C., Rowland, P., Schofield, C. J., Sheppard, R. J., Smith, J. E., Swales, C., Tanner, R., Thomas, P., Tumber, A., Drewes, G., Oppermann, U., Patel, D. J., Lee, K., and Wilson, D. M. (2012) A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408
 68. Bach, R. D., Thorpe, C., and Dmitrenko, O. (2002) C–H⁺⋯carboxylate oxygen hydrogen bonding in substrate activation by acyl-CoA dehydrogenases: synergy between the H-bonds. *J. Phys. Chem. B* **106**, 4325–4335
 69. Chen, J. C., Hanson, B. L., Fisher, S. Z., Langan, P., and Kovalevsky, A. Y. (2012) Direct observation of hydrogen atom dynamics and interactions by ultrahigh resolution neutron protein crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15301–15306
 70. Westler, W. M., Lin, I. J., Perczel, A., Weinhold, F., and Markley, J. L. (2011) Hyperfine-shifted ¹³C resonance assignments in an iron-sulfur protein with quantum chemical verification: aliphatic C–H⁺⋯S 3-center-4-electron interactions. *J. Am. Chem. Soc.* **133**, 1310–1316
 71. Bartlett, G. J., Choudhary, A., Raines, R. T., and Woolfson, D. N. (2010) $n \rightarrow \pi^*$ interactions in proteins. *Nat. Chem. Biol.* **6**, 615–620
 72. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The C α –H⁺⋯O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9056–9061
 73. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1994) Crystal structure of intercalated four-stranded d(C₃T) at 1.4 Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11636–11640