The Functions and Importance of CH···O Bonds in SET Domain Methyltransferases

by

Scott A. Horowitz

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in The University of Michigan
2013

Doctoral Committee:

Professor Hashim M. Al-Hashimi, Co-Chair
Associate Professor Raymond C. Trievel, Co-Chair
Professor Charles L. Brooks III
Assistant Professor Tomek Cierpicki
Professor Anna K. Mapp

2013

Acknowledgements

I would like to first and foremost thank Profs. Ray Trievel and Hashim Al-Hashimi. I consider myself lucky to have been able to learn two different approaches to a scientific problem from two tremendous scientists. I would also like to thank my dissertation committee for their many useful suggestions over the years.

I would like to thank everyone from the Trievel and Al-Hashimi groups, many of whom have aided and taught me over the years. In particular, Paul Del Rizzo taught me much of what I know about working with proteins and biochemistry, and Jennifer Nimtz was extremely helpful in protein crystallization. Alex Hansen has given me enormous support learning advanced NMR concepts, and most importantly, Joseph Yesselman has been an enthusiastic collaborator on many facets of this work, without whom I would not have been able to complete much of what is presented here.

Also, many collaborators from outside the university have been instrumental in this work. Most notably, Lynnette Dirk and Bob Houtz performed the radiometric assays, and gladly took up each new challenge as it came. Also, the advice of Ryan Mehl provided essential advice on making protein with unnatural amino acids and the help of Sam Butcher and Kirk van der Meulen helped us considerably with ITC analysis. And last but not least, Doug Markham was responsible for helping all of this get started in the first place with his gift of the AdoMet synthetase enzyme.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Methylation is a ubiquitous reaction in biology, particularly with respect to metabolism, signal transduction, and gene regulation. This reaction is most often catalyzed through a convergently evolved S_N2 mechanism that utilizes the cofactor S-adenosylmethionine (AdoMet). Despite its importance, little is known of the determinants of this conserved reaction. The following dissertation describes the discovery of one such determinant, unconventional CH···O hydrogen bonds. A new methodology utilizing NMR spectroscopy is first developed to identify biological CH···O hydrogen bonds, and applied to the AdoMet methyl group within the active site of model methyltransferase SET7/9. Using a structural survey, the universality of these interactions in the classes of AdoMet-dependent methyltransferases is then demonstrated. The functional importance of CH···O hydrogen bonds hydrogen within the active site of a model methyltransferase are verified by structural, functional, and dynamic analyses of SET7/9, using a combination of x-ray crystallography, mutagenesis with natural and unnatural amino acids, quantum chemistry calculations, and NMR spin-relaxation experiments. The CH···O hydrogen bonds are shown to aid in cofactor binding as well as catalysis by stabilizing both the reactant and transition states of methyl transfer, and optimizing the catalytic geometry of the AdoMet methyl group. Together, these studies elucidate a universal mechanism of AdoMet-dependent methyl transfer, providing a molecular framework for this integral biological reaction.

Chapter 1: **Introduction**

1.1 Mechanism and Structure of AdoMet Dependent Methyltransferases

1.1.1 Mechanism of Methylation

S-Adenosylmethionine (AdoMet) is the second most utilized substrate in the cell, after ATP (*1*). AdoMet comprises adenine and methionine molecules that are covalently linked through their C5′ and sulfur atoms, respectively, creating a sulfonium cation (Figure 1.1). The most common AdoMet-dependent reaction in eukaryotes is methylation, in which the single electrophilic AdoMet methyl group is transferred to a nucleophilic acceptor atom (N, O, S, or C). Cellular methyl acceptors are numerous, including proteins, RNA, DNA, natural products, amino acids, lipids, and hormones (*2*) In addition to metabolic functions, AdoMet-dependent methyltransferases also participate in signal transduction through, RNA (*3*), DNA, and protein methylation (*4*), as numerous proteins have evolved the function of "reading" the methylation status of specific residues. Dysfunction or misregulation of AdoMet-dependent methylation has been implicated in many diseases, including multiple forms of cancer (*5-8*), Parkinson's Disease (*9, 10*), Alzheimer's Disease (*11, 12*), HIV (*13*), and antibiotic-resistant bacterial infections (*14-16*). Due to its importance to many cellular functions as well as disease, determining the mechanism and specificity of AdoMet-dependent methylation has been a subject of study for decades.

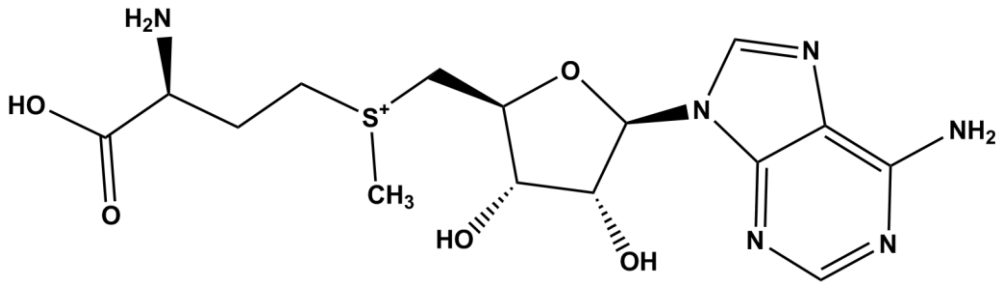

Figure 1.1 AdoMet chemical structure

In the 1970s, several groups attempted to determine the mechanism of AdoMet-dependent methyl transfer. Many of the early studies used sulfonium-containing small molecules to model the methyl transfer reaction by mimicking the reactivity of the AdoMet sulfonium group. By analyzing the kinetic and thermodynamic properties of sulfonium containing molecules and reaction properties as a function of solvent and temperature, and comparing these parameters to known reactive compounds, these studies discovered that the reaction proceeds through an S_N2 mechanism in which the methyl group is transferred directly to the acceptor atom, passing through a planar sp2 transition state (*17, 18*) (Figure 1.2). Schowen and coworkers expanded upon these investigations in the enzyme catechol-o-methyltransferase (COMT), utilizing direct kinetic isotope effect (KIE) measurements to confirm the proposed S_N2 mechanism. In these experiments, the authors used AdoMet containing either a ^{12}C or ^{13}C methyl group, and examined the maximal enzymatic rate in each case. The authors measured large kinetic isotope effect on the transferred group, indicating that methyl transfer was the rate limiting step of the reaction in COMT (*19, 20*). As a result of these studies, rate-limiting S_N2 transfer became the paradigm for the mechanism and kinetics of methylation (*21*).

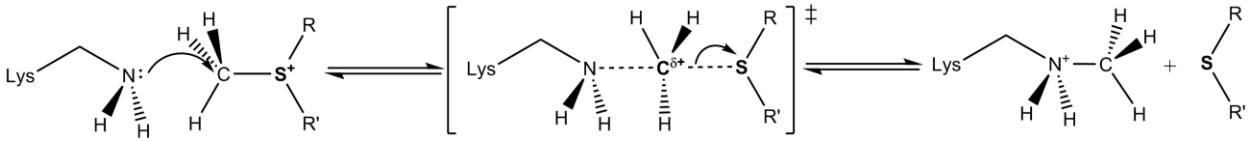

Figure 1.2: SN2 mechanism of AdoMet-dependent methyl transfer.
Methyl acceptors can be O, S, C or N atoms.

In addition to determining the rate-limiting step and general chemical mechanism of methylation, the aforementioned studies examined how methyltransferases could achieve their catalytic power. Although many catalyzed methyl transfer reactions are slow by enzymatic standards (usually slower than 1 s^{-1}), this still accounts for a 10^{16} rate enhancement over the uncatalyzed reaction in solution (*22*). The sulfonium model studies discussed above were the first to suggest a mode for the acceleration, as they observed that the transfer reaction was faster when performed in a nonpolar environment, suggesting that a nonpolar active site aided in catalysis (*17*). The KIE studies refined this model by proposing additional details. Schowen and coworkers observed an unusually large inverse deuterium KIE for COMT (*20*), in which ^{2}D-methyl AdoMet caused a 20% increase in reaction rate compared to ^{1}H-methyl AdoMet, suggesting that methyltransferases utilize a "tight" transition state (Figure 1.3), in which the enzyme compresses, or shortens, the distance between the AdoMet sulfonium donor group and the oxygen acceptor, thereby creating a chemically distinct transition state that is preferentially bound by the enzyme (*22, 23*). After being verified by early computational models (*23, 24*) compression of the donor-acceptor distance in a nonpolar active site became the standard model for the AdoMet methyl transfer reaction, and stood unchallenged for over a decade.

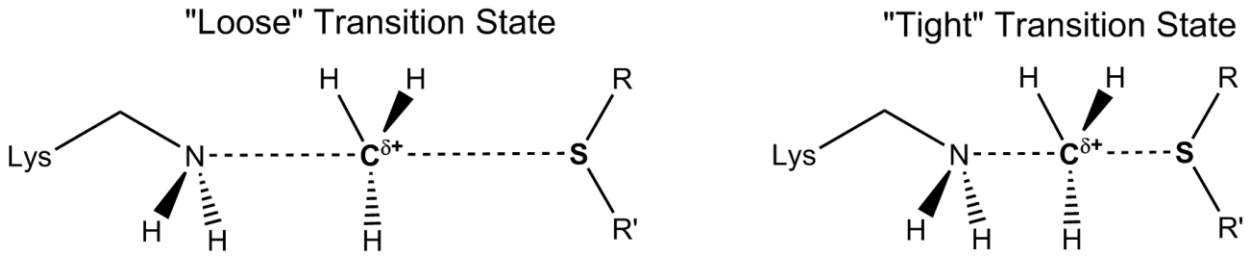

Figure 1.3: Example of "loose" and "tight" transition states.
Methyltransferases were proposed to compress the donor-acceptor distances to create the "tight" transition state, shown right.

By 1993, the computational methods used to support compression hypothesis were under fire (*25*). As computational quantum chemistry methods improved, a new wave of studies revisited the original KIE experiments by Schowen to try and determine the S_N2 transition state structure in greater detail. These studies especially focused on factors that could have led to the inverse KIE, and to test the compression model. Updated QM/MM models often reversed previous thinking, concluding that the inverse KIE could be achieved without any active site compression (*26, 27*). Still, the compression hypothesis has not been entirely ruled out. A recent experimental study of COMT confirmed Schowen's KIE measurements, and implied that active site compaction is involved in the methyltransferase mechanism. The authors questioned the validity of the QM/MM simulations on the basis of lingering computational limitations (*28*).

As the debate over the compression hypothesis continues, other theories have risen to challenge it. One alternative proposal to compression posits that the primary purpose of a methyltransferase is to orient the reactants in an ideal methyl transfer position, termed the near-attack conformation (NAC), and that this pre-orientation is responsible for rate enhancement (*29, 30*). This hypothesis is also not without critics. Warshel and coworkers have contended that the near attack conformation in S_N2

reactions is merely a byproduct of the enzyme's attempts to stabilize the transition state by providing an electrostatically favorable environment that is more selective for the transition state than the surrounding solution (*31*). This contention has been supported by computations suggesting that the primary purpose of the enzyme is to selectively recognize and stabilize the transition state over the reactant states through hydrogen bonding and favorable active site electrostatics (*26*). Complicating matters, computational studies by Williams and coworkers have shown that sampling of different dynamic reactant states in the enzyme can lead to multiple different possible attack conformations that averaged together can yield the experimentally-determined KIEs, questioning the notion that the experimental transition state can be accurately described by any single structure (*32*). Although these studies have questioned the long-held notion of active site compression and provide interesting and plausible alternatives, difficulties in experimentally observing the properties and structure of the transition state have prevented them from being experimentally tested to date. As a result, the methyl transfer community is still attempting to answer a major question that was first addressed over thirty years ago: how does the methyltransferase enhance the methyl transfer rate?

1.1.2 Structure of AdoMet-Dependent Methyltransferases

At the outset of methyltransferase structural investigation, the primary theory was that methyltransferases were not structurally similar given their low sequence conservation. Thus, it was a surprise when the first crystal structures of methyltransferases, Hhal in 1993 (*33*) and COMT (*34*) in 1994, revealed that these enzymes fold into highly similar structures despite low sequence homology. They each contain a Rossmann-like fold, consisting of a seven strand β-sheet buttressed by six flanking alpha helices. As more methyltransferase crystal structures were solved, this

fold was repeatedly observed, leading it to be named the "consensus methyltransferase fold" (*35*). Despite maintaining the same fold, AdoMet conformation, and AdoMet binding position, the sequence conservation in the active site of these enzymes remained surprisingly low, even in the AdoMet binding pocket. It was hypothesized that this low sequence homology was required to allow methylation of a variety of substrates while maintaining the same structure (*35*). This was unexpected, because often function is preserved through conserved sequence. Further, the low sequence conservation obscured mechanistic insights of methyltransferases, as it became difficult to attribute functions to specific functional groups in the enzyme active site.

Despite the title of "consensus methyltransferase fold", this conclusion was based on a limited number of structures, and it was not long before other AdoMet-dependent methyltransferase architectures were discovered. By 1996, a second class of methyltransferase had been found that was structurally distinct from the Rossmann-like Class I enzymes (*36*). In the intervening years, many more structural classes have been discovered and come to the fore. As of today, there are as many as nine unrelated structural classes of AdoMet-dependent methyltransferases (*37*). This expanding set of structural classes was first categorized by Cheng et al. in 2003, shortly after the discovery of the fifth class, the SET domain superfamily. Cheng proposed the well-accepted theory that AdoMet-dependent methylation evolved convergently, with different structural classes emerging to methylate novel substrates as organisms developed the need (*38*). Despite having no homology in sequence, structure, AdoMet binding mode, or substrate specificity, it is generally thought that these enzymes all catalyze methylation using the same S_N2 mechanism, (in the case of the radical SAM methyltransferases, as the first part of a multi-step mechanism (*39*)) with generally the

same properties (*40, 41*). Despite this similarity in catalytic mechanism, no common structural properties of all methyltransferase classes have yet been defined.

After Class I, the most studied methyltransferase class is Class V, or the SET domain class. This class is primarily responsible for lysine methylation on both histone and non-histone substrates in eukaryotes, with over fifty separate family members in humans that regulate functions such as gene transcription, cell cycle progression, and epigenetic inheritance (*42, 43*). These enzymes all contain a highly conserved SET domain that adopts a pseudo-knot-type structure consisting of a β-sheet fold bifurcated by an inserted region of variable size and structure, termed the iSET motif (Figure 1.4) (*44, 45*). AdoMet binds in a pocket with its methyl group protruding into a narrow pore. The protein substrate binds on the opposite face of the SET domain, with the lysine side chain of the protein substrate extending through a channel traversing the core of the protein, terminating at the AdoMet methyl transfer pore on the other side. This arrangement places the lysine ε-amine group and AdoMet methyl group into close proximity and in appropriate geometry for S_N2 transfer (*46*). Among the methyltransferases in the SET domain family, SET7/9, which catalyzes lysine monomethylation in multiple sequence contexts, has emerged as a model system for study of these enzymes, as it was among the first of these enzymes to be structurally, dynamically, and kinetically characterized (*40, 47-55*).

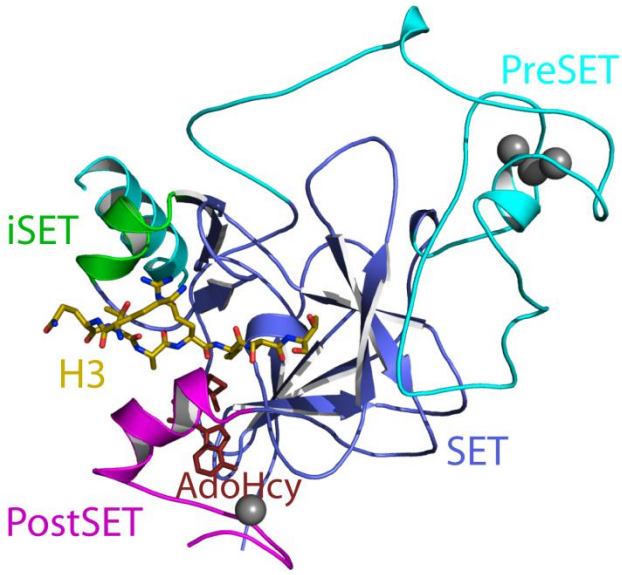

Figure 1.4: SET domain architecture.
SET domain in complex with product AdoHcy (red) and H3K9me1 peptide (yellow carbon atoms).

Early structures from the Hurley and Trievel groups identified that the AdoMet methyl transfer pores of SET domain methyltransferases are curiously lined with oxygen atoms (Figure 1.5) (*49, 56*). This finding was unusual, as it was expected that as the methyl group is often considered a hydrophobic moiety, it would be found within a hydrophobic pore. The distances and positions of these oxygens relative to the AdoMet methyl group suggested that it may form unconventional CH···O hydrogen bonds with multiple oxygen atoms in the SET domain active site. These hydrogen bonds are still little known or understood in biology, but were hypothesized to aid in AdoMet binding as well as catalysis by stabilizing the S_N2 transition state and aligning the electrophile for attack (*49*). Studying these interactions and how they are involved in methyl transfer has formed the bulk of my dissertation research. The following section introduces the concept of CH···O hydrogen bonding and its roles in biology to provide context for investigating their importance and roles in methylation.

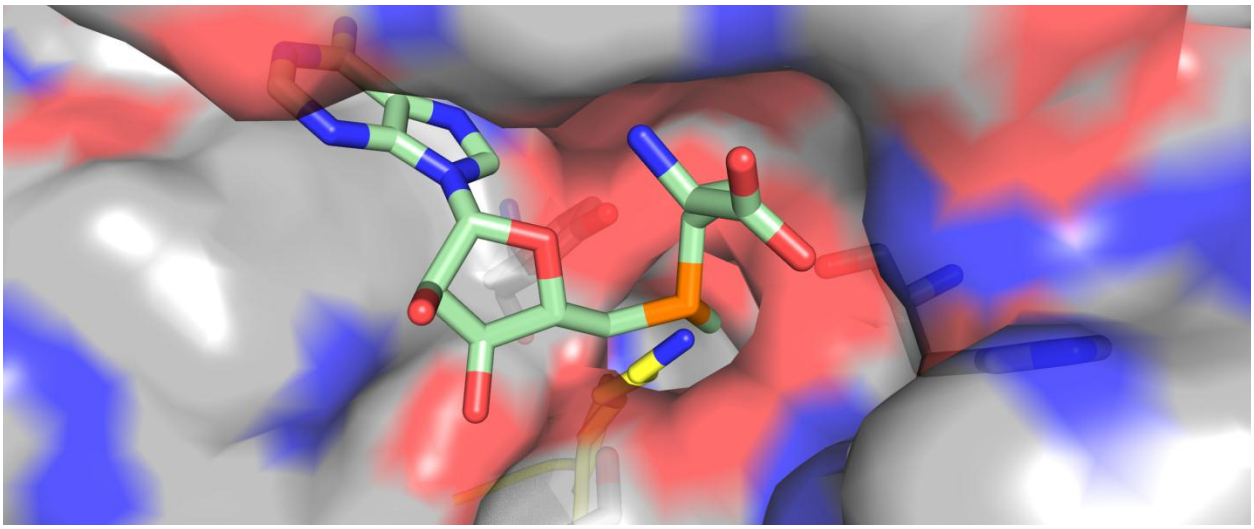

Figure 1.5: Oxygen-lined pore of SET domain methyltransferase.
Lysine acceptor is depicted opposite AdoMet.

1.2 CH···O Hydrogen Bonds in Biology

1.2.1 Introduction and Historical Perspective

Conventional hydrogen bonds (NH···O, OH···O, OH···N, and NH···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 to oxygen and nitrogen. However, several studies have illustrated that aliphatic carbon atoms are capable of forming weak hydrogen bonds, which are denoted as CH···O hydrogen bonds (*57*, *58*). In contrast, with increased polarization due to adjacent atoms, carbons can theoretically participate in hydrogen bonds as strong as those formed by conventional donors, specifically oxygen or nitrogen (*59*, *60*).

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 A-H,

where A can be any electronegative element, and an acceptor group that is either a lone pair of electrons or a π bond (*61*). In these interactions, the hydrogen is shared between the donor and acceptor to varying degrees, and charge transfer occurs across this bond. 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, and furthermore, not all chemists and physicists agree that charge transfer is even a requirement for hydrogen bond formation (*62*). Hydrogen bonds are further distinguished from van der waals contacts by the presence of electrostatic attraction, which is not a requirement of the van der waals interaction that need only consist of dispersive attraction and exchange repulsion (*63*). Recently, The Union of Pure and Applied Chemistry (IUPAC) commissioned a task force to update the definition of the hydrogen bond. The resulting document very specifically addresses how a hydrogen bond is defined, and I refer the reader to it for details (*64*). The overarching theme of this new definition is that it provides a set of criteria (both theoretical and empirical) for evaluation of a particular interaction as to whether or not to consider it a hydrogen bond. As discussed in this document, there are many different means to identify a hydrogen bond, including using structural characteristics such as angles and distances that indicate hydrogen bond formation. This approach is especially useful in biological systems, for which many high-resolution crystal structures have been solved and provide a rich source of hydrogen bonding information. As such, within biochemistry and biophysics, it has become commonplace to use distance and angular criteria to define a hydrogen bond.

Hydrogen bonds tend towards linearity and optimal overlap between the lone pair of the hydrogen bond acceptor and the hydrogen atom. Typically, the sharing of

the hydrogen atom and electrostatic attraction between the donor and acceptor allows them 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 (Figure 1.6). 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···O bonds cause a substantial downfield ^{1}H chemical shift change (*65*). In infrared (IR) spectroscopy, these interactions are unusual in that they usually cause a blue IR shift, indicative of C-H bond shortening, as opposed to the typical bond lengthening observed in conventional hydrogen bonds (*66, 67*). Despite this difference, the literature on this subject has reached a consensus that the CH···O interaction represents a bona-fide hydrogen bond (*68, 69*).

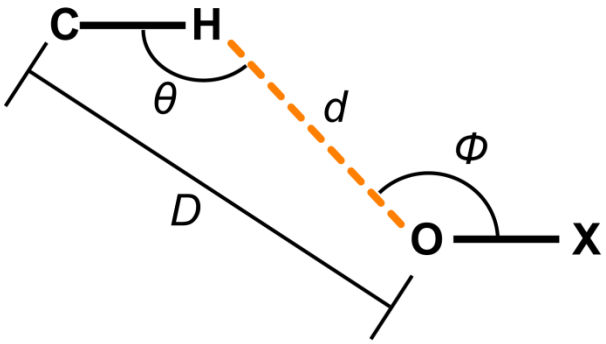

Figure 1.6: Distance and angular parameters used when defining CH···O hydrogen bonds.
Typical van der Waals distances *d* (2.7 Å) and *D* (3.7 Å) are frequently used as distance cutoffs for hydrogen bond identification.

The emergence of CH···O hydrogen bonding as an important interaction in biological structure and function stems from research dating back several decades. Although the first proposals of carbon-donor hydrogen bonding by Kumler (*70*) and

Glasstone (*71*) date to the mid-1930s (Figure 1.7), the first notable studies in proteins by Ramachandran (*72, 73*) and Krimm (*66, 74*) in the 1960s raised the possibility that these interactions contribute significantly to protein structure. In more recent work, Derewenda et al. (*75*) catalogued the ubiquitous nature of backbone Cα donor hydrogen bonds in proteins based on a survey of 13 high-resolution protein crystal structures, using methodology established by Taylor and Kennard in their landmark Cambridge Structural Database survey in 1982 (*76*). Using van der Waals distance cutoffs, Derewenda's survey identified that a surprisingly high percentage of Cα···O contacts form CH···O hydrogen bonds in these proteins. The mean distance calculated for all C···O contacts in the interactions surveyed was 3.5 Å, well within the van der Waals distance cutoff of 3.7 Å (Figure 1.6). Using C···C interactions as a reference, they were able to clearly demonstrate the widespread nature of CH···O hydrogen bonding in proteins (Figure 1.8), especially in the standard backbone hydrogen bonding pattern of β-sheets. In addition, CH···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 (*77*). Since these early studies, many have endeavored to elucidate the breadth, scope, and importance of CH···O hydrogen bonding in biomolecular structures.

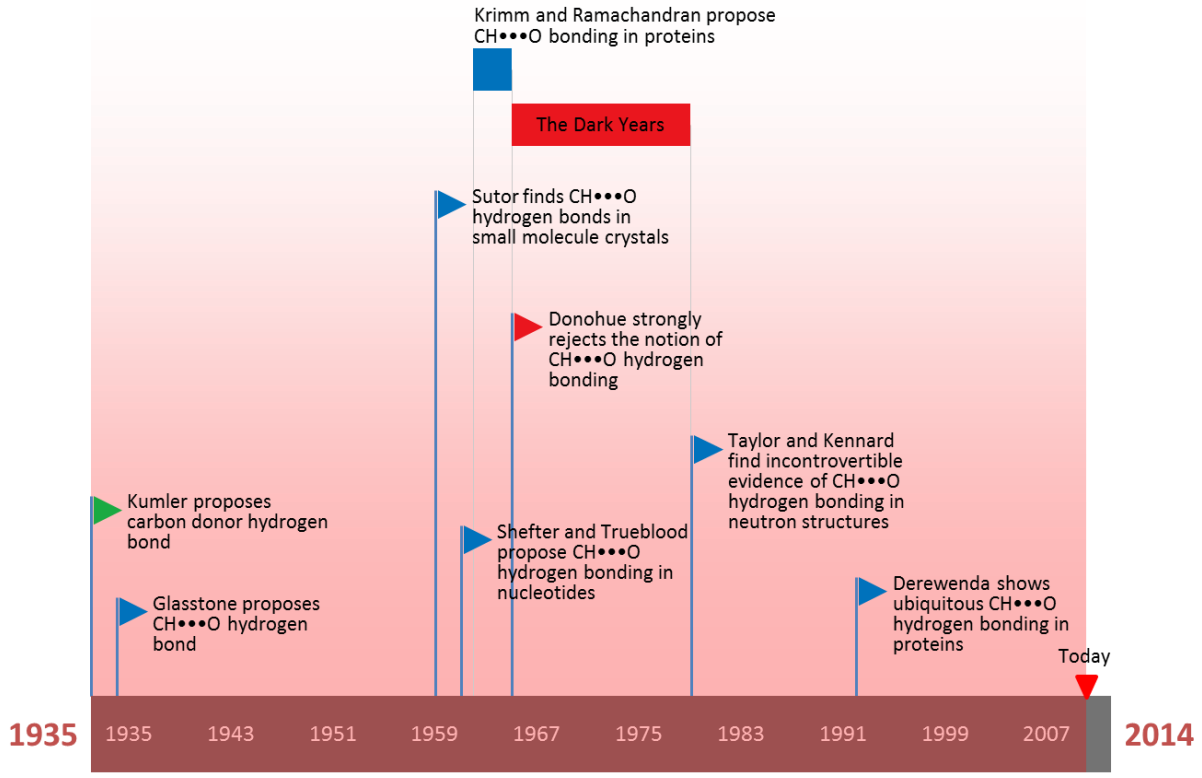

Figure 1.7: Timeline of important events in the history of CH···O hydrogen bonding research.
The "Dark Years" period outlined in red refers to the years following Donohue's rejection of the existence of CH···O hydrogen bonding (red flag), and before Taylor and Kennard's landmark study confirming their existence. For a recount of this episode, please see Schwalbe's review of the subject (*78*).

1.2.2 Contributions to Protein Structure

Characterization of CH···O hydrogen bonds in protein structure spans several decades (*66*, *72*, *75*, *79*). The aforementioned study by Derewenda and colleagues has led to widespread acceptance of Cα-Hα···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 visualization of hydrogen bonding patterns within proteins. Many of these studies have attempted to define CH···O bonding patterns

within protein structures (*80-82*) and established unequivocal evidence for CH···O hydrogen bond formation in parallel and anti-parallel β-sheets. In fact, it was determined that the idealized position of Hα atoms in β-sheet structure was rarely observed, as the Hα was frequently displaced 0.2-0.3 Å away from its idealized position to increase its CH···O hydrogen bonding potential (*82*). These findings were further substantiated using NMR spectroscopy through scalar and quadrupolar coupling measurements. In 2003, Grzesiek and colleagues employed long-range scalar coupling experiments to examine Cα-Hα···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···O hydrogen bonds in the context of a folded protein, providing direct evidence of hydrogen bond formation (*83*). Other NMR evidence has more recently come from the measurement of Hα 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 Cα-Hα···O=C bonds based on distance (*84*). 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 Cα-Hα···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···O hydrogen bonds involving amino acid side chains in protein structure. For example, histidine side chains have been implicated in CH···O hydrogen bonding (*81, 85*), and are predicted to form interactions as strong as conventional hydrogen bonds when the

imidazole group is protonated or bound to a metal ion (*86*). Additional evidence has emerged from neutron crystallography, an evolving technology that holds great promise for directly visualizing CH···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 (*87*). Over the past decade, the number of neutron structures in the PDB has increased dramatically, providing many unprecedented snapshots of protein hydrogen bonding. Most notably, one recent study endeavored to analyze the prevalence of CH···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···O hydrogen bonds in the copper coordination site (Figure 1.8) in addition to eight conventional hydrogen bonds (*88*). As neutron crystallography evolves, unconventional interactions will likely become more accepted as there are observed more commonly in high-resolution structures.

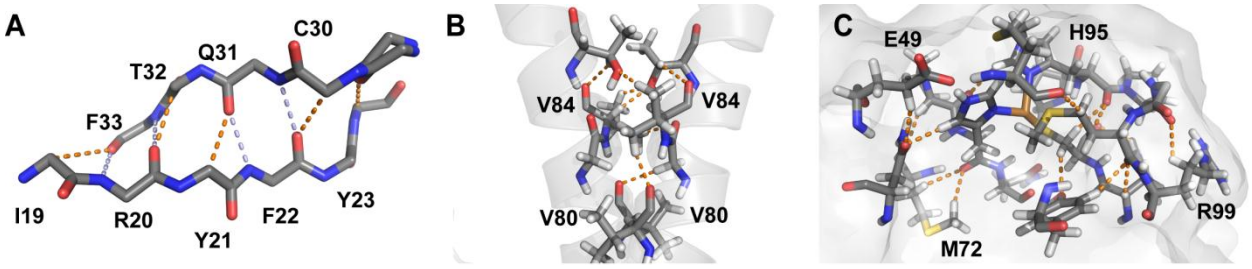

Figure 1.8 Examples of CH···O hydrogen bonds (orange dashes) in proteins.
(A) Typical hydrogen bonding pattern in β-sheet (*82*). Extensive CH···O hydrogen bonding in (B) packed α-helices (*89*) and (C) the copper coordination site in amicyanin (*88*).

1.2.3 CH···O Hydrogen Bonding in Nucleic Acid Structures

Although many recent advances in understanding CH···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···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 (*90, 91*). NMR and Raman spectroscopy later corroborated the claim that the acidity of the C8 atom in purine rings poises it for CH···O hydrogen bonding in RNA duplexes (*92*). Finally, a survey of high-resolution RNA structures (<2.0 Å) illustrated that short CH···O contacts are ubiquitous and appear to stabilize RNA tertiary structure (*93*). Despite these crystallographic findings and multiple studies implicating CH···O hydrogen bond formation in nucleic acids (*94-97*), the quantity of direct evidence of CH···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···O hydrogen bonding (Figure 1.9) (*98*). However, this structure, does not address many common RNA secondary and tertiary structural elements. Although the formation of CH···O hydrogen bonds in A-T base pairs, both in Watson-Crick and Hoogsteen conformation, have been debated in the literature (*99-102*), experiments have yet to definitively resolve whether these interactions represent true hydrogen 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.

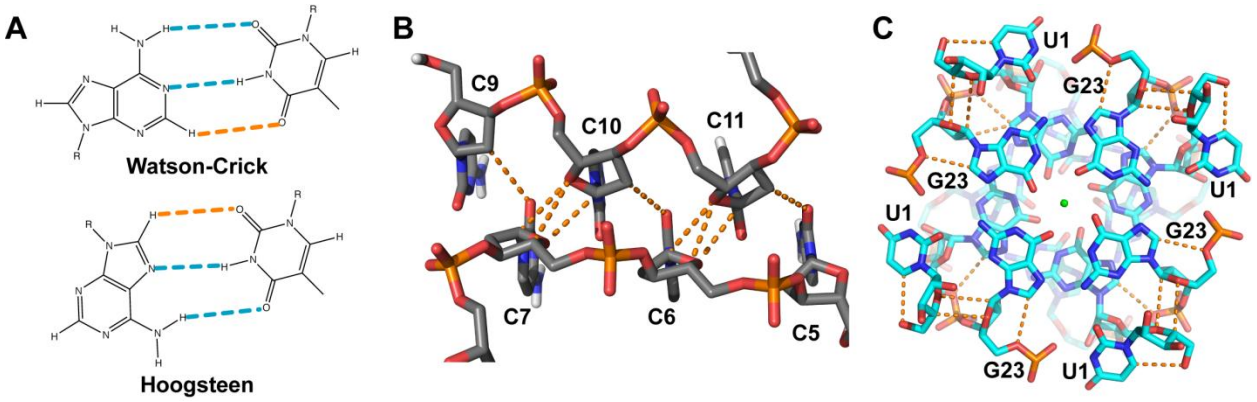

Figure 1.9: Examples of CH···O hydrogen bonds (orange dashes) in nucleic acids.
(A) Hydrogen bonding patterns in Watson-Crick (top) and Hoogsteen (bottom) adenine-thymidine base pairs. Extensive CH···O hydrogen bonding in (B) DNA i-motif (*103*) and (C) RNA tetraplex (*98*).

1.2.4 CH···O Hydrogen Bonding in Molecular Recognition

Evidence and importance of CH···O hydrogen bonding to molecular recognition is 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 (*104*) (Figure 1.10). Similarly, in a recently determined neutron structure of transthyretin, the authors noted that that the ratio of CH···O to conventional hydrogen bonds in the interface between the A and D subunits is greater than a 2:1 (*105*), potentially providing substantial binding energy.

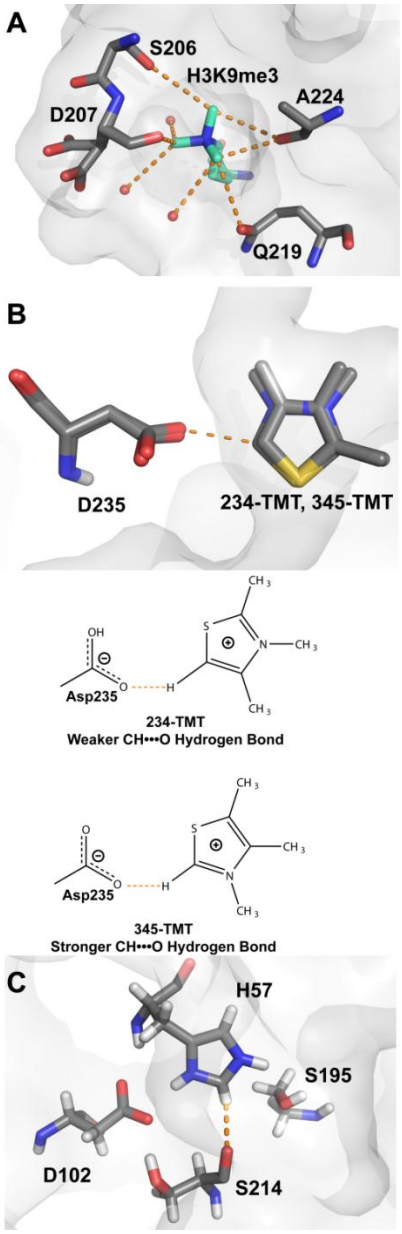

Figure 1.10: Examples of CH···O hydrogen bonds (orange dashes) in molecular recognition and enzyme catalysis.
(A) Recognition of trimethyllysine by CH···O hydrogen bonds in the ATRX ADD domain (41). (B) Schematic depiction of 234TMT (upper) and 345TMT (lower) binding to an engineered protein active site (*106*). (C) Conserved serine hydrolase CH···O hydrogen bond between the catalytic histidine and serine residues (*107*).

With respect to protein•ligand binding, Klaholz et al. was among the first to use X-ray crystallography to analyze the nuclear receptor RARγ 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 (*108*). Similar to neutron studies of protein•protein interactions, neutron diffraction studies also have provided evidence for these interactions between proteins and water and ligand molecules.For example, in the neutron structure of xylulose-isomerase, a CH···O hydrogen bond between the protein His220 C1 and O2 of xylulose may facilitate xylulose recognition (*109*). 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 (*110*), 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 bonding 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 (*111*). In total, the distribution of C···O distances in nucleic acid/protein complexes was more consistent with conventional N···O hydrogen bond distances than that of control C···C distributions. Interestingly, the thymine methyl group is the most frequent CH···O hydrogen bond donor, in spite of its lower polarization compared to 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.

1.2.5 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 (*59*, *112*). These studies have generally shown CH···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 can be as strong as conventional hydrogen bonds (*59*). In one striking example, it was estimated that 17% of all protein•protein surface interaction energy arises from CH···O hydrogen bonding, and that in some extreme cases this percentage can be as high as 40-50% (*113*). Although CH···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···O hydrogen bond strength and importance in proteins. Using a combination of high-pressure infrared spectroscopy and X-ray crystallography, CH···O bonds were found to be crucial in artificial β-sheet-like network formation (*114*), and it is possible that these interactions are of similar importance in

true β-sheets. In a contrasting example, Bowie et al. determined that a Cα-Hα···O hydrogen bond in bacteriorhodopsin did not provide any stabilization protein stabilization energy (*115*), although subsequent studies demonstrated that conventional hydrogen bonding in bacteriorhodopsin was also surprisingly weak (*116*). In contrast, a separate concurrent study using infrared spectroscopy discovered a substantially stronger protein CH···O hydrogen bond, = 0.9 kcal/mol, comparable to a weak conventional hydrogen bond in a protein environment (*79*). Correlatively, Kallenbach et al. analyzed helical peptides using circular dichroism to show that a single CH···O hydrogen bond contributed ~0.5 kcal/mol of helix stabilization energy, but only with certain side chain sequences and orientations (*117*). Thus, from experimental evidence, the exact energetic stabilization contributed by CH···O hydrogen bonds to protein folding remains ambiguous. As in the case of conventional interactions, the contribution of a single CH···O hydrogen bond to protein stabilization depends on its hybridization and polarization as well as its role and position in protein folding (*118*). Further studies are needed to precisely define the energetic values of CH···O hydrogen bonds and their contributions to protein folding.

In comparison to the foregoing examples, relatively few studies have experimentally explored CH···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 (234TMT) and 3,4,5-trimethylthiazole (345TMT), to a cytochrome c peroxidase mutant (Figure 1.10). 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···O hydrogen bond to Asp235 in the enzyme. The only difference between these two molecules is that in 234TMT, the nitrogen is two positions removed from this hydrogen bonded C-H group, whereas in 345TMT, it is only one position removed, similar to histidine residues. Due to the proximity of the positively charged nitrogen to the CH···O hydrogen bond that should enhance the carbon polarization, 345TMT was computationally predicted to form a substantially stronger CH···O bond than 234TMT, 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 (Figure 1.10). Upon measuring the binding constants for each ligand, the authors determined that the binding energy of 345TMT was 1.2 kcal/mol greater than that of 234TMT (*106*), substantiating their computational predictions.

Within nucleic acids, the CH···O hydrogen bonding strength within the i-motif has undergone uniquely in-depth study. This motif is a quadruplex-like structure consisting of four strands containing intercalated C.CH+ base pairs and linked loops (Figure 1.9) that may form at the end of telomeres (*119*). Initial structures of this motif revealed extensive CH···O hydrogen bond networks, suggesting that these interactions may promote i-motif formation (*120, 121*). To substantiate these observations, Gueron et al. attempted to determine the average strength of C1′ CH···O bonds using different intercalation topologies that a single sequence could form as a function of ionic strength and temperature (*121*). Based on these experiments, an average C1′···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 proteins, it is likely that CH···O hydrogen bonding contributes substantially to the specificity of RNA folding. Indeed, these interactions have recently been introduced as a parameter in RNA tertiary structure predictions (*122*), but have yet to be experimentally evaluated. Similarly, determining what, if any, contribution CH···O hydrogen bonding plays in 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 et al. is similar in strength to that determined by the previous study of 345TMT, supporting the notion that these interactions are not only important in protein structure and folding, but also in ligand binding.

1.2.6 CH···O Hydrogen Bonding in Enzyme Catalysis

In addition to macromolecular structure, CH···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. identified short histidine-oxygen distances between the catalytic histidine C1 and an adjacent carbonyl oxygen that were indicative of CH···O hydrogen bonding (*123*). This hydrogen bond was proposed to stabilize the imidazolium cation and to potentially facilitate a ring-flipping mechanism (*85, 123*). Derewenda's observation was further supported by Bachovchin et al., who used NMR chemical shift to show that the downfield chemical shift of the histidine H1 proton was consistent with CH···O hydrogen bonding (*85*). Corroboratively, Hunter et al. proposed based on crystallographic distances that a CH···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 electronically induced fit mechanism (*124*).

CH···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 CH···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 Cα atom of the acyl-CoA substrate by a glutamate base. The reaction was investigated computationally (*125*), indicating that a strong ionic CH···O hydrogen bond forms along the reaction coordinate prior to proton abstraction from the Cα position in the acyl chain. 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 addition roles for CH···O hydrogen bonding in many enzyme mechanisms.

Recent studies have also identified functions for CH···O hydrogen bonding in lysine demethylation and methylation reactions. CH···O hydrogen bonds appear to be important in methyllysine binding both by SET domain methyltransferases and Jumonji-C (JmjC) lysine demethylases (KDMs), analogous to the ATRX ADD domain Figure 1.10.In the JmjC KDM JMJD2A, mutational and structural evidence suggest that CH···O hydrogen bonds are important in distinguishing between di- and trimethylated lysine. These interactions have also been reported in other JmjC KDMS, including UTX (*126*), PHF8 (*127*) and JMJD3 (*128*).

Structural and functional studies of the lysine methyltransferases belonging to the SET domain family revealed that AdoMet methyl CH···O hydrogen bonds are conserved in these enzymes, indicating a potential role in cofactor binding and catalysis (*49*) (Figure 1.11). These hydrogen bonds could potentially be strong interactions due to the electron withdrawing character of the sulfonium cation, polarizing the methyl group of AdoMet. Isothermal calorimetry demonstrated that AdoMet displays higher binding affinity to a SET domain methyltransferase than AdoMet analogues that are incapable of forming CH···O hydrogen bonds (*49*), suggesting the importance of these interactions in substrate binding. It was hypothesized that these interactions not only aid in AdoMet recognition, but also in transition state stabilization by selectively binding to the partially positively charged sp2 carbon in the S_N2 transition state. In addition, structural and mutagenic data indicated that CH···O hydrogen bonds are important for the repositioning of the ε-amine group to enable lysine multiple methylation (*129, 130*). Together, these studies suggested that CH···O hydrogen bonds play an active role in SET domain function, providing the foundation for my dissertation work.

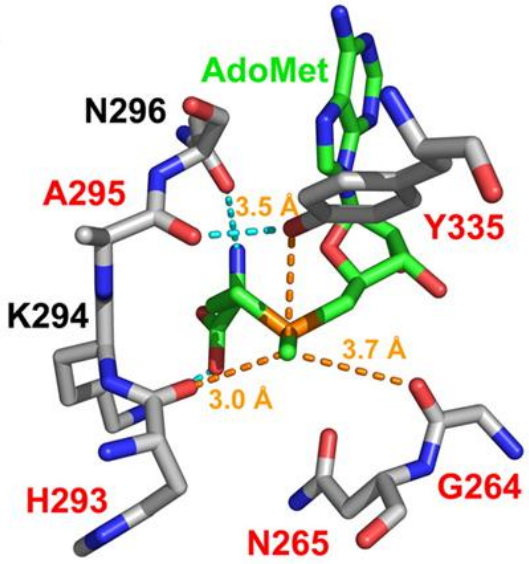

Figure 1.11: Crystal structure of SET7/9 bound to AdoMet (*47*, *49*).
This figure was originally published in The Journal of Biological Chemistry (*49*) © the American Society for Biochemistry and Molecular Biology.

1.3 Objectives of this Dissertation

Despite many recent discoveries pertaining to CH···O hydrogen bonding in biology, to our knowledge, no studies have experimentally demonstrated that CH···O hydrogen bonds are vital to any core biological process. Using AdoMet-dependent methylation as a case study, I intended to directly test whether CH···O hydrogen bonds are fundamentally required for cofactor binding or catalysis in these enzymes. In this dissertation I aim to answer three major questions about the role of CH···O hydrogen bonding in methyl transfer: 1) Do these CH···O hydrogen bonds exist in the methyltransferase active site? 2) Are they important for cofactor binding or catalysis? And 3) How do these interactions facilitate cofactor binding and/or catalysis?

Chapter 2: **Direct Evidence for Methyl Group Coordination by CH···O Hydrogen Bonds in SET Domain Methyltransferases**

2.1 Objective and Approach

As discussed in Chapter 1, CH···O hydrogen bonding has been recognized as an important interaction in proteins and other biological macromolecules dating back over 40 years (*74, 131-133*). Despite their potential importance, experimental characterization of CH···O hydrogen bonds in proteins remains challenging. Current methods for identifying CH···O hydrogen bonds (*83, 84, 88*) are difficult to employ for many proteins, including SET7/9. However, NMR spectroscopy holds promise in identifying CH···O hydrogen bonds in proteins and other macromolecules via ^{1}H chemical shift (*67, 85, 134*). Chemical shift is a sensitive probe of electronic environment. The ^{1}H nucleus is especially sensitive to hydrogen bonding (*65*), as a distinctive downfield chemical shift change accompanies hydrogen bond formation. This change is thought to arise from charge transfer to the hydrogen atom, a property unique to hydrogen bonding that dominates the chemical shift effects of other forces present in many types of interactions (*135*). Using this property, Bachovchin et al. were able to qualitatively confirm that the histidine H1 in serine proteases forms a CH···O hydrogen bond, as discussed in Chapter 1. More recently, it was proposed that CH···O bonds in biomolecules could be detected quantitatively by combining ^{1}H chemical shift measurements with DFT computations (*67*). The feasibility of this approach was recently demonstrated by using a similar strategy to discover CH···S hydrogen bonds to a protein iron sulfur cluster (*136*).

As previously mentioned, CH···O hydrogen bonding in SET domain methyltransferases were identified in X-ray crystal structures, in which C···O distances between the AdoMet methyl carbon and structurally conserved carbonyl oxygen atoms and an invariant tyrosine were appropriate (less than 3.7 Å) for CH···O hydrogen bonding to occur (*49*) (Figure 1.11). In addition, it was postulated that the CH···O hydrogen bonds align the AdoMet methyl group in the requisite linear geometry with the substrate lysine ε-amine group for the S_N2 methyl transfer reaction. (Figure 1.5) (*48, 49, 137-139*). However, as hydrogen atoms are not observed in all but the highest resolution X-ray crystal structures, it is possible that the AdoMet methyl protons occupy positions that preclude CH···O hydrogen bond formation, despite close C···O contacts within SET domain crystal structures. Thus, the presence of CH···O hydrogen bonding can only be inferred from these structures. The fundamental question we address here: Is there direct experimental evidence for CH···O hydrogen bonding between the AdoMet methyl group and the SET domain active site in solution?

2.2 Materials and Methods

2.2.*1* Expression and purification of MetK and SET7/9

A plasmid encoding the AdoMet synthetase (MetK) gene from *Methanococcus jannaschii* was generously provided by George D. Markham. The gene was subcloned into pHT4, a variant on the pET15b vector (Novagen) that contains an N-terminal 6x-His tag with a TEV protease cleavage site to facilitate protein purification. MetK was expressed in *E. coli* BL21 DE3 cells grown in LB media by induction with 0.1 mM isopropylthiogalactoside for 4 hours at 37°. The enzyme was purified on a Talon Co affinity column (Clontech) followed by Superdex 200 (GE Healthcare) gel filtration chromatography. Following gel filtration, MetK was concentrated to approximately 20

mg/mL as determined by its absorbance at 280 nm, flash frozen in liquid nitrogen, and stored at -80^{o} C.

SET7/9 (residues 110-366) was expressed and purified as previously described (*140*), with the following exceptions. The enzyme was purified by denaturation and refolding while immobilized on a Ni-Sepharose column (GE healthcare). SET7/9 was unfolded by washing the column with 7-10 column volumes of 6M guanidinium chloride, refolded with a gradient into the lysis buffer, and subsequently eluted using an imidazole gradient. The denaturation and refolding protocol were necessary to remove AdoMet that can co-purify with the recombinant enzyme expressed in bacteria (*140*). After gel filtration chromatography, SET7/9 was concentrated to at least 40 mg/mL as determined by its absorbance at 280 nm, flash-frozen, and stored at -80^{o} C.

2.2.2 Synthesis and purification of AdoMet

^{13}C-methyl AdoMet was enzymatically synthesized from ATP and ^{13}C-methyl L-methionine (Sigma) using MetK as previously described (*141*) with the following modifications. Reactions were performed in final volume of 5 mL for 5 hours at 298 K. At the end of the reaction, AdoMet was purified using a Source 15S column (GE Healthcare), as previously described (*142*). The purity of the resulting AdoMet was verified by NMR and estimated to be approximately 95% pure. Existing impurities showed neither chemical shift change nor intensity change upon addition of ^{13}C-methyl AdoMet to SET7/9.

2.2.3 Molecular dynamics (MD) simulation and quantum chemistry calculations of AdoMet in water

In order to accurately quantify the chemical shift of the AdoMet methyl group in water, we first determined the number hydrogen bonds formed by the AdoMet methyl group in water, followed by quantum chemical calculations of the AdoMet methyl group's chemical shift. The starting structure for the MD simulation was derived from a B3LYP/6-311+G(2d,p) optimized AdoMet structure. Partial charges were assigned to the starting AdoMet structure using a B3LYP/6-311+G(2d,p) CHELP calculation (*143*). All density functional theory (DFT) computations were accomplished using Gaussian 03 (*144*). The MD simulation was run using the CHARMM36 forcefield (*145*) in (35.2 x 25.5 x 22.2) Å TIP3P water box (*146*) for 10 ns, using a 2 fs timestep. The simulation was conducted at 298K using a Nose-Hoover Thermostat (*147*) using periodic boundary conditions conducted using the particle mesh Ewald method (*148*). Other AdoMet (bond, angle, dihedral) parameters derived from the CHARMM General Force Field parameter set (*149*). Analysis of the MD trajectory in 20 fs steps revealed that the AdoMet methyl group forms, on average, 0.36 hydrogen bonds in solution per methyl group. Therefore, to correctly model the solution state of the AdoMet methyl group, one weak CH···O hydrogen bond was included in the chemical shift calculations.

The coordinates of the aqueous AdoMet solution structure were provided by George D. Markham (*150*). The methyl proton geometry was optimized using B3LYP/6-311+G(2d,p) (*151, 152*) with implicit water solvation by the Polarizable Continuum Model (*153*) with all other atoms frozen. The optimized geometry of the AdoMet methyl group formed one intra-molecular CH···O hydrogen bond with the ribose 3′-hydroxyl group (Figure 2.2). The geometry of this interaction (C-H-O angle = 131° and H···O

interaction distance = 2.5 Å) was consistent with a weak hydrogen bond. Thus, this interaction satisfied our condition to include one weak CH···O hydrogen bond in the quantum chemistry calculations to correctly represent AdoMet in water, as determined by the MD simulation described above. Chemical shifts were calculated using the Gauge Independent Atomic Orbitals (*154*) method in implicit water for the reasons described above. The chemical shift of the AdoMet methyl group in this conformation yielded a chemical shift of 3.0 ppm, well within 0.1 ppm of experiment. Breaking the CH···O hydrogen bond by rotating the AdoMet methyl group yielded a chemical shift of 2.9 ppm, confirming that the suboptimal geometry of the hydrogen bond resulted in a weak interaction (*155*). As a reference, tetramethylsilane was optimized in implicit water and chemical shifts were calculated with the same methods as above.

2.2.4 Quantum chemistry calculations on the SET7/9•AdoMet complex

The model SET7/9 active site included all atoms within 5 Å of the AdoMet methyl group from the SET7/9•AdoMet binary complex (*47*). The AdoMet molecule was truncated at the carbon positions adjacent to the sulfonium cation. Protons were added automatically using Chimera (*156*), and a single chlorine atom was placed in the position of the AdoMet carboxylic acid group to neutralize the system. All added hydrogen positions were optimized using B3LYP/3-21G* (*157, 158*), and methyl protons were subsequently optimized and chemical shifts calculated with B3LYP/6-311+G(2d,p) as described for free AdoMet. Using implicit solvent with a lower dielectric constant to reflect the hydrophobic core of proteins (ε=4.9) for chemical shift calculations yielded no change in ^{1}H chemical shift of the AdoMet methyl group as compared to implicit water solvation. Figures were rendered in PyMol (Schrödinger LLC). All calculated and measured chemical shifts are shown in Table S1. Methyl rotamers for chemical shift

calculations were created by manually rotating the AdoMet methyl group. Structural figures were rendered using PyMOL (Schrödinger, LLC).

2.2.5 NMR spectroscopy

All NMR experiments were performed on a 600 MHz Avance Bruker NMR spectrometer equipped with a triple resonance cryoprobe at 298 K. Spectra were referenced using the water signal. Data was processed and analyzed using NMRPipe/NMRDraw, and Sparky, respectively (*159, 160*). The AdoMet methyl group's assignment was confirmed by recording control ^{13}C 2D- Heteronuclear Single Quantum Coherence (HSQC) spectra of three separate samples: SET7/9 with no added AdoMet, SET7/9 plus stoichiometric quantities of unlabeled AdoMet, and ^{13}C-AdoMet with no added SET7/9. The resonance assigned as the bound AdoMet methyl group was not present in any of these three control spectra, whereas it was apparent in the ^{13}C-methyl-AdoMet•SET7/9 sample. Upon addition of slight stoichiometric excess of SET7/9 to ^{13}C-labeled AdoMet, the peak corresponding to the free AdoMet ^{13}C-methyl group completely shifted to the enzyme bound-peak in the 2D-HSQC spectrum, indicating that the AdoMet was fully saturated by the enzyme. Spectral overlays of SET7/9 in cofactor bound and free states were previously reported (*51, 52*).

2.2.6 Isothermal Titration Calorimetry (ITC)

ITC experiments were performed using a Microcal VP-ITC (GE Healthcare). All experiments were performed at 293 K in 20 mM Sodium Phosphate, pH=7.0 and 100 mM sodium chloride. Varying concentrations of ligand (0.07-7.7 mM) and protein (0.006-0.193 mM) were used due to the large range of binding constants measured. Data analysis was accomplished using Microcal Origin (GE Healthcare). All binding curves had N-values between 0.8-1.0. Errors reported in Table 1 are from curve fitting errors.

Previous crystallographic studies have shown that the binding modes of AdoMet, Sinefungin and S-adenosylhomocysteine (AdoHcy) to SET7/9 are highly homologous (Figure 2.5).

2.3 Results

To probe CH···O hydrogen bonding between SET7/9 and the AdoMet methyl group, we examined the NMR chemical shift produced by the AdoMet methyl group while bound to SET7/9. To measure the bound ^{1}H chemical shift for the AdoMet methyl group and distinguish it from resonances arising from the enzyme, we synthesized ^{13}C-methyl labeled AdoMet using AdoMet synthase and ^{13}C-methyl-methioine (*141, 161, 162*). We recorded 2D-HSQC spectra of ^{13}C-methyl-AdoMet in the presence of stoichiometric quantities of the unlabeled catalytic domain of SET7/9 (Figure 2.1). The ^{1}H chemical shift of the AdoMet methyl group was recorded as 3.8 ppm. This chemical shift was unusually far downfield for a methyl group and also represented a large downfield change relative to the reported chemical shift for AdoMet methyl group free in solution (3.0 ppm) (*163*). Unlike ^{13}C, downfield ^{1}H chemical shift changes often qualitatively indicate hydrogen bond formation (*65, 85, 135, 164, 165*), thus we sought to verify our experimental chemical shift change using quantum chemistry calculations. While this combination of techniques has been used to identify CH···O hydrogen bonds in small organic molecules (*134, 166, 167*) and in computational biology (*65*), it has to our knowledge not yet been applied experimentally in biological macromolecules. Using this combination of techniques, we reasoned that it should be possible to solve for the hydrogen positions, and thus hydrogen bonding patterns, of the AdoMet methyl group within the SET7/9 active site.

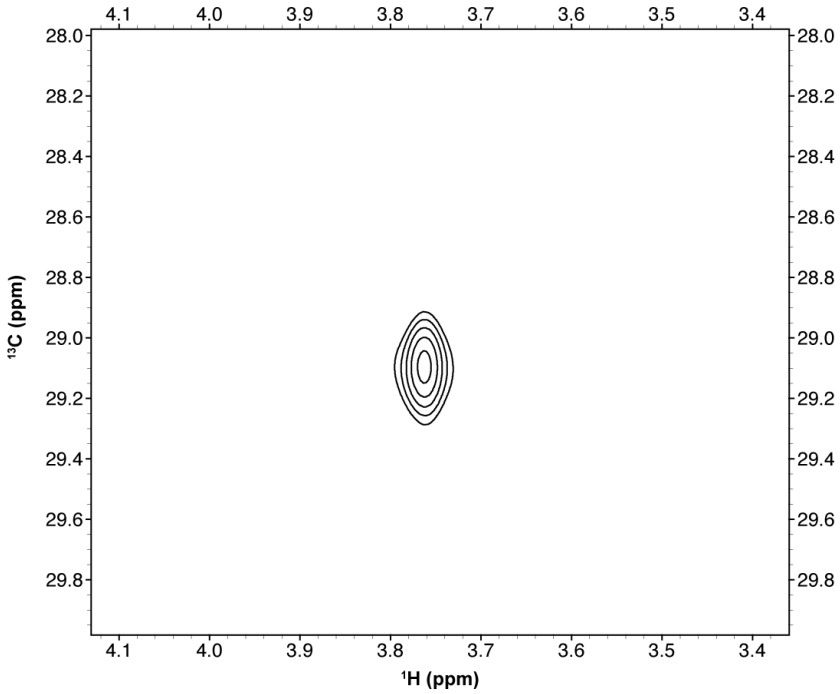

Figure 2.1: 2D-HSQC of SET7/9•^{13}C-methyl AdoMet complex

Previous studies of CH···O hydrogen bonding in small organic molecules show that the chemical shift calculations of hydrogen were usually accurate to 0.1 ppm of experiment (*134*, *166*, *167*). Error in biological molecules could arise from many sources, including but not limited to implicit solvation modeling and large or truncated molecules used in calculations. Therefore, to validate the accuracy of our calculations for the AdoMet methyl group, we first calculated the chemical shift of the AdoMet methyl group using the solution state NMR structure of free AdoMet in water. Methyl proton geometry was optimized, allowing for the appropriate number of hydrogen bonds formed in solution by the methyl group, followed by chemical shift calculation. By averaging all three values together to reproduce the single experimental methyl peak, the difference in chemical shift between the experimental (3.0 ppm) (*163*) and calculated values was less than 0.1 ppm. This level of accuracy prompted us to attempt to locate the methyl protons of AdoMet within the active site of SET7/9.

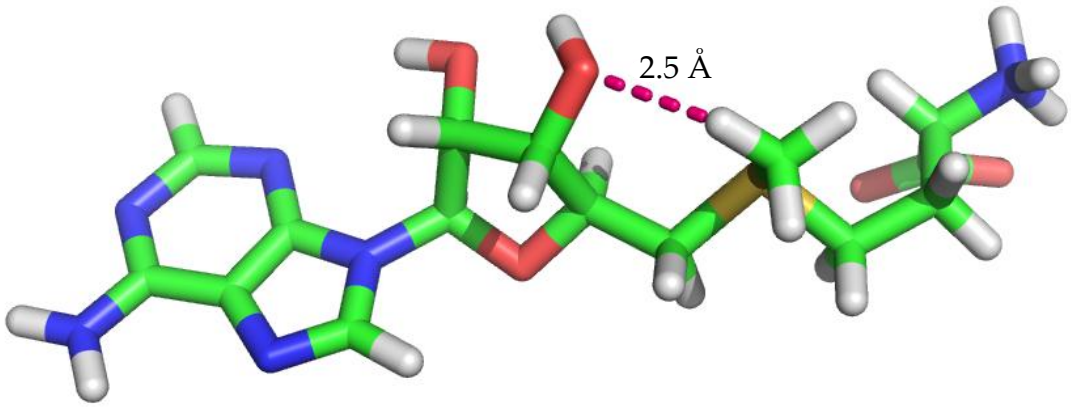

Figure 2.2: Free AdoMet in implicit solvent.

The CH···O hydrogen bond is shown in magenta.

To compare with our experimental data, we then modeled the active site of SET7/9 using its crystal structure bound to AdoMet at 1.7 Å resolution (*47*). Hydrogens were added to this model, and the geometry of the hydrogen positions was optimized (see Chapter 2.2 for details). Chemical shifts were then calculated for the methyl group protons and averaged to a single value, as performed for free AdoMet. The geometry-optimized structure of the AdoMet methyl group shows CH···O hydrogen bond formation from the AdoMet methyl group to the hydroxyl group of Tyr335 and main chain carbonyl oxygen of His-293, with H···O distances of 2.5 and 2.1 Å, respectively (Figure 2.3). Moreover, the C-H···O angles of 146° and 140°, respectively, were acceptable hydrogen bonding angles (*132, 133, 168*). As predicted, the calculations showed that protons participating in CH···O hydrogen bonds experienced significant downfield changes in chemical shift. The proton engaged in close hydrogen bonding with His-293 had a calculated chemical shift of 5.0 ppm, while the proton forming hydrogen bonds with Tyr335 had a calculated chemical shift of 3.3 ppm. The third proton's chemical shift was calculated to be 2.6 ppm. Averaging all three values

together yielded a calculated chemical shift of 3.7 ppm, which is within 0.1 ppm of the experimental value of 3.8 ppm.

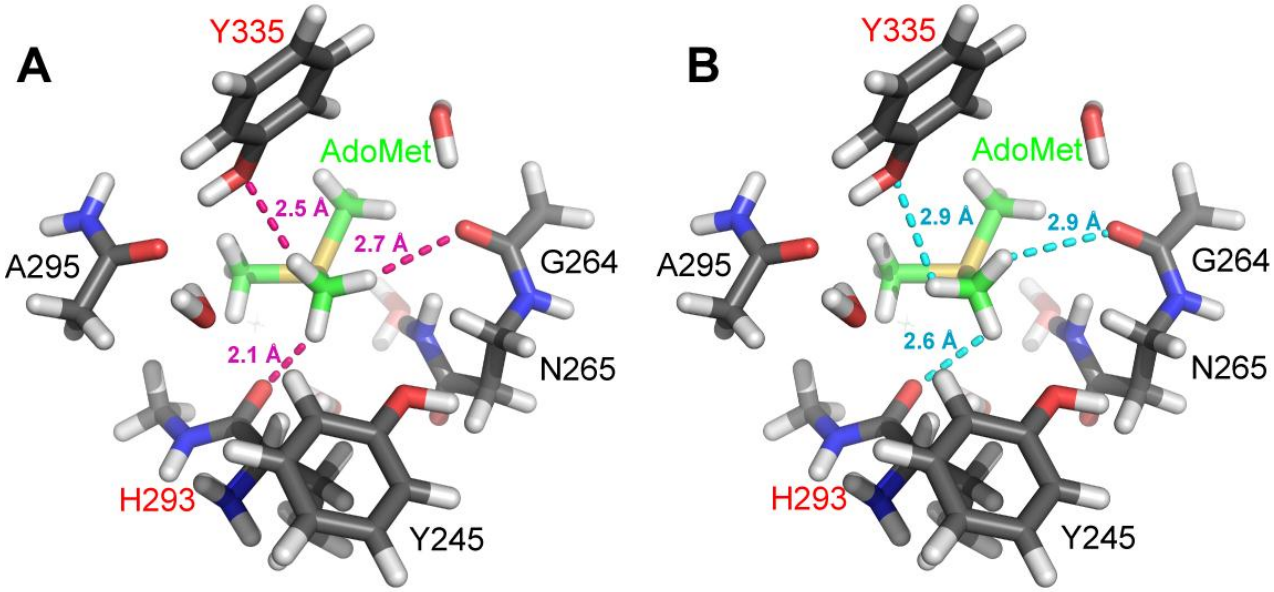

Figure 2.3: Optimized and broken CH···O hydrogen bonds in SET7/9.
(A) Optimized active site with bound AdoMet and (B) manually rotated geometry to eliminate CH···O hydrogen bonds. Truncated AdoMet and the protein are depicted with green and gray carbon atoms, respectively. Residues labeled in red designate CH···O acceptors

To further confirm the presence of CH···O hydrogen bonding, the methyl group was manually rotated through a range of 120° (example shown in Figure 2.3), and chemical shifts recalculated at intermediate geometries to explore whether alternative hydrogen positions could agree with the experimental data. The rotation of the AdoMet methyl group resulted in a smooth functional change in the calculated chemical shift, passing through a minimum at 36° (Figure 2.4). The minimum of this plot corresponded to the structure for which the model and the experimentally measured shift differ by 0.6 ppm. This difference represented: 1) the largest deviation between the experiment and model, 2) the most upfield of the calculated chemical shifts, and 3) the model in which the methyl rotamer formed minimal CH···O hydrogen bonds in the active site (Figure

2.3). These calculations indicated that AdoMet methyl rotamers precluding CH···O hydrogen bond formation were inconsistent with the experimental chemical shift data. All of the calculated chemical shifts that agreed closely with experimental chemical shifts represented structures that optimized CH···O hydrogen bonding for at least one AdoMet methyl hydrogen atom. Collectively, the experimental data and calculations yield direct evidence via NMR chemical shift for CH···O hydrogen bond formation between the AdoMet methyl group and oxygen atoms within the active site of SET7/9.

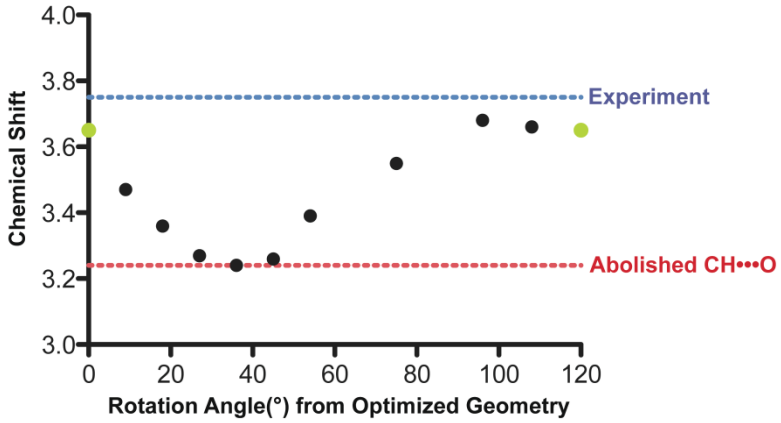

Figure 2.4: Chemical shift of the AdoMet methyl group as a function of rotation angle.
0° and 120°, rendered in green, are equivalent geometry optimized positions (displayed in Figure 2.3), whereas all other points were derived from manual rotation of the AdoMet methyl

These findings prompted us to examine the thermodynamic parameters for cofactor binding by SET7/9 using ITC. Binding affinities and enthalpies for SET7/9 were measured with AdoMet (Figure 2.6), its methyl transfer product AdoHcy, and Sinefungin, an AdoMet analog in which the methyl-sulfonium cation is substituted by an amine-methylene group that can participate in conventional NH···O hydrogen bonding. Crystal structures of SET7/9 and other SET domain enzymes illustrated that these cofactors share a structurally homologous binding mode (Figure 2.5) (*47, 48, 50*). The ITC data revealed that SET7/9 displayed nanomolar affinities for AdoMet and

Sinefungin, whereas its affinity for AdoHcy was ~1000-fold weaker (Table 1). These data were analogous to those obtained for the binding of these ligands to the SET domain protein Rubisco large subunit methyltransferase (LSMT), demonstrating consistency in rank order of binding affinities across different SET domain enzymes (*49*). In addition, SET7/9 and LSMT displayed comparable differences in binding enthalpy between AdoMet and Sinefungin (~3 kcal/mol). The high affinity that these enzymes displayed for AdoMet is presumably due to the ability of its methyl group to engage in CH···O hydrogen bonding, emphasizing the importance of these interactions in cofactor binding by SET domain methyltransferases.

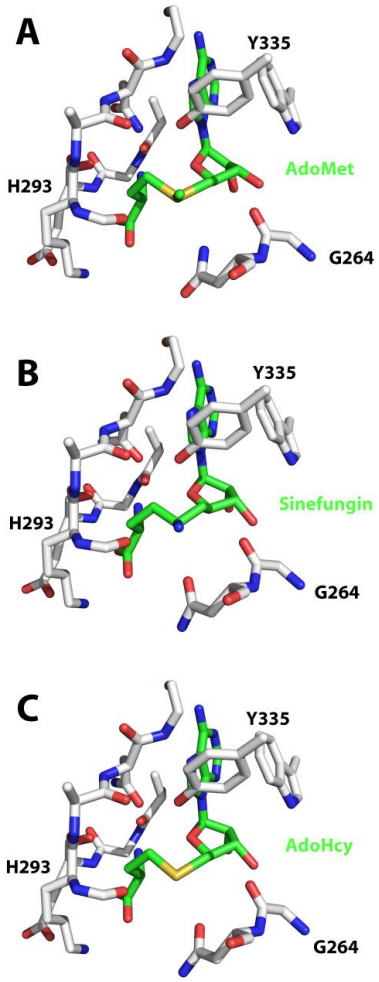

Figure 2.5: Binding modes of AdoMet (A), Sinefungin (B), and AdoHcy (C) to SET7/9 are nearly identical. Figures were rendered from PDB codes 1N6A (*47*), 3CBP (*50*) and 2F69 (*169*), respectively. Cofactor carbon atoms are rendered in green, and protein residues are rendered in gray. For clarity, some backbone atoms and side chains have been removed. The residues that can engage in CH···O hydrogen bonding with the AdoMet methyl group are labeled.

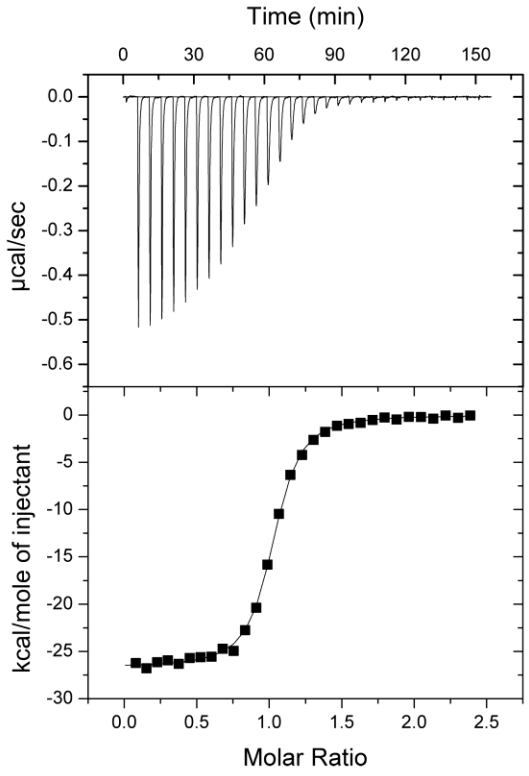

Figure 2.6: ITC analysis of AdoMet binding to SET7/9.
The top panel represents the titration of AdoMet into SET7/9, while the bottom panel represents the binding isotherm with fitted curve.

Table 2.1 Dissociation constants and binding enthalpies of AdoMet, Sinefungin, and AdoHcy to SET7/9

Ligand	K_D		ΔH	
	nM		*kcal/mol*	
AdoMet	53	± 3	-26.7	± 0.1
Sinefungin	330	± 24	-23.0	± 0.2
AdoHcy	60000	± 5000	-24.4	± 1.2

2.4 Discussion

The identification of CH···O hydrogen bonding between a methyl group and oxygen atoms within the SET domain active site has implications for these interactions in lysine methyl transfer reactions. As previously proposed, CH···O hydrogen bonds appear to play roles in binding AdoMet, positioning its methyl group in an appropriate

geometry for transfer, and stabilizing the S_N2 transition state (*49*). These data suggest that the CH···O hydrogen bonds confer a specific orientation for the methyl group to align it during catalysis and could potentially limit its motion within the active site.

The thermodynamic analyses of SET7/9 bound to AdoMet and its analogues are consistent with previous findings on LSMT (*49*). The consistency of rank order in binding affinity indicates that CH···O hydrogen bonds are important in cofactor binding by multiple methyltransferases due to the structural conservation of the SET domain active site. Moreover, these CH···O interactions may in part explain the importance of the evolutionarily invariant Tyr335 to enzyme function, given the propensity of its hydroxyl group to form CH···O hydrogen bonds with the AdoMet methyl group. It is also interesting to note that the change in binding affinity between AdoMet and its analogues is substantially more dramatic in SET7/9 than in LSMT. This effect could be a function of the plasticity of the SET7/9 active site (*52*), as compared to the preformed active site of LSMT (*170*). Future studies will need to address how conformational flexibility within the SET domain family influences CH···O hydrogen bonding to AdoMet.

In terms of methodology, our results have demonstrated that chemical shift can be used as a structural parameter for determining hydrogen positions and hydrogen bonding patterns within an enzyme active site. To our knowledge, these chemical shift data and computations provide the first direct, quantitative evidence of CH···O hydrogen bonding in an enzyme active site in solution. One advantage of chemical shift as a probe to examine CH···O hydrogen bonding in biomolecular structure is the relative ease of data acquisition. In the future, this methodology could be broadly applied to characterize CH···O hydrogen bonding in proteins, nucleic acids, and other

biological molecules, expanding our understanding of the functional importance of these interactions in macromolecular structure, ligand binding, and enzyme catalysis.

Chapter 3: **CH···O Hydrogen Bonds Mediate AdoMet-Dependent Methylation**

3.1 Objective and Approach

Similar to the difficulties in discovering and geometrically characterizing CH···O hydrogen bonds, little progress has been made in determining the importance or specific functions of CH···O hydrogen bonds in biological molecules. As discussed in Chapter 1, for SET domain methyltransferases, CH···O hydrogen bonds were proposed to aid in cofactor binding, stabilize the sp2 transition state of methyl transfer, and to aid NAC formation.

3.1.1 Tyrosine 335 Mutagenesis

To test the effect of CH···O hydrogen bonds upon both AdoMet binding, we continued our studies with the model system SET7/9. As discussed in the previous Chapter, one of the AdoMet methyl CH···O hydrogen bonding partners is an invariant tyrosine residue in the SET domain class. Unlike the backbone CH···O hydrogen bond acceptor in SET7/9, Tyr335 can be studied through mutagenesis. Thus, to probe the functional importance of the Tyr335 CH···O hydrogen bonds, we employed a biochemical approach to disrupt them through site-directed mutagenesis of Tyr335 using natural and unnatural amino acids (UAAs).

To disrupt the Tyr335 CH···O hydrogen bonds to AdoMet without disturbing the conventional hydrogen bond between the tyrosine hydroxyl group and the protein backbone, we employed an *amber* stop-codon suppression strategy to genetically substitute Tyr335 with the UAA *para*-aminophenylalanine (pAF) (*171*). This amino acid

is isosteric to tyrosine but replaces the hydroxyl group with an amine moiety. Site-specific incorporation of unnatural amino acids during protein translation can be accomplished by introducing a new tRNA and tRNA synthetase that are orthogonal to other tRNA-synthetase pairs in the bacterium. Such tRNA synthetases can be generated by directed evolution of current synthetases (*172*). The new tRNA is designed to recognize the *amber* stop codon (UAG), and thus permitting translation to proceed only by inserting the unnatural amino acid into the growing polypeptide chain.

We reasoned that the pAF mutation would abolish all CH···O hydrogen bonding by eliminating the Tyr335 hydrogen bond acceptor, but would not disrupt AdoMet methyl group positioning or the conventional hydrogen bond, enabling direct comparison of the structural, functional, and thermodynamic properties between WT and the Y335pAF mutant. In addition to the pAF substitution, we also generated a Y335F mutant to abolish all CH···O hydrogen bonding in addition to the conventional hydrogen bond formed by the tyrosine hydroxyl group. We reasoned that a comparison of the properties of the Y335F and Y335pAF mutants would offer insight into the specific functions of these CH···O hydrogen bonds.

3.1.2 NMR spin-relaxation

As previously mentioned, one possible role of CH···O hydrogen bonds is to promote NAC formation by aligning the AdoMet methyl group for nucleophilic attack. NAC formation is dependent on fixing both the nucleophile (lysine), and electrophile (AdoMet) into a reactive conformation. In the case of SET7/9, the lysine residue forms conventional NH···O hydrogen bonds that likely position it for transfer (*48*). To probe whether the AdoMet methyl group is fixed in space by the CH···O hydrogen bonds in the SET7/9 active site, we employed NMR spectroscopy. Biomolecular methyl group

dynamics are readily measured by NMR spin-relaxation and relaxation-dispersion experiments to probe the ps-ns and μs-ms timescales, respectively (*173, 174*). One major complication in implementing these experiments arises from the unusually large number of pathways for dipolar relaxation in methyl groups. In a $^{13}CH_3$ methyl group, there are sixteen separate single quantum coherences along which dipolar relaxation can occur, along with several double and zero quantum coherences (*175*). Relaxation by each of these pathways occurs at distinct rates, and thus the measured relaxation rates of a $^{13}CH_3$ methyl group cannot be fit to a mono-exponential curve and yield a single rate. To facilitate analysis, the relaxation rates are often split into fast-relaxation and slow-relaxation groups, and fit separately (*176*). Unfortunately, this method, and subsequent methods that measure these rates (*177, 178*), are only semi-quantitative, and thus not capable of directly determining the dynamics of a single, isolated methyl group without other methyl groups to serve as internal dynamic references.

The above-mentioned difficulties associated with methyl group relaxation measurements and fitting are easily circumvented by using deuterated methyl isotopomers. Carbon and deuterium relaxation can be accurately measured using $^{13}CHD_2$ and $^{13}CH_2D$ methyl groups respectively. Due to its spin = 1 state, deuterium does not cross-relax appreciably with ^{13}C-H groups, and contributes little in terms of direct dipolar relaxation (*174*). As such, relaxation of a $^{13}CHD_2$ group can be measured and analyzed in a very similar fashion as any isolated ^{13}CH group, with only minor corrections to account for ^{13}C-D dipolar relaxation (*174*). Although deuterium relaxation is also straightforward to analyze and can provide accurate methyl dynamics measurements, at the time of experimentation, only the $^{13}CHD_2$ methionine isotopomer was commercially available. Therefore, using $^{13}CHD_2$ methyl relaxation, in combination with natural and unnatural amino acid mutations, we sought to determine the functions

and importance of the Tyr335 CH···O hydrogen bonds to SET7/9 cofactor binding and catalysis.

3.2 **Materials and Methods:**

3.2.1 High-Resolution Crystal Structure PDB Survey

All co-crystal structures of methyltransferases bound to AdoMet with resolution better than 2.00 Å were downloaded and visually inspected before inclusion in the AdoMet and protein methyl and methylene CH···O hydrogen bond surveys. Duplicate structures were removed, keeping the highest resolution structure of each methyltransferase. Electron density from the EDS server was examined to verify model quality when available (*179*). Due to lack of $2F_o$-F_c density at the AdoMet methyl group position, several structures (2IGT, 1FPX, 1I9G, 2EGV, 2QE6, 2YQZ, 3H6L, 3EEO) were not used in the survey. For a few structures (3UJ7, 1N2X, 1ZQ9), analysis was restricted to a single copy only of the methyltransferase, due to lack of methyl density. All other structures with multiple copies within the PDB file had CH···O hydrogen bonding acceptor parameters averaged before analysis. 1N2X, 2ZUL and 1V2X were refined by real space parameters using Coot (*180*) to adjust for model building and refinement anomalies before inclusion in the survey. As hydrogen positions for these structures were unknown, methyl group hydrogen bonding was based on van der Waals radii (*75*) and non-hydrogen atom angles. In addition to AdoMet methyl groups, all protein side chain methyl C···O interactions were analyzed using identical hydrogen bonding cutoffs to provide a reference point for analyzing the properties of the AdoMet methyl and methylene CH···O hydrogen bonds.

Four separate criteria were used to determine if methyl groups participated in CH···O hydrogen bonding: C···O distance, X-C···O angle, and C···O-Y angle, where X

and Y are the non-hydrogen atoms bonded to carbon and oxygen, respectively, and sp2 oxygen-methyl elevation angle (*181*). Hydrogen bonding was determined using two scenarios 1) For C···O distances less than 3.25 Å, which is the sum of the van der Waals radii of carbon and oxygen; and 2) For distances greater or equal to 3.25 Å but less than 3.70 Å, which is the combined van der Waals radii of a C-H bond and oxygen arranged in linear geometry (*75*). In the former scenario, methyl groups were considered to participate in CH···O hydrogen bonding if the C···O distance was less than 3.25 Å, the X-C···O angle was between 60° and 160°, and the elevation angle was less than 75°. For the latter scenario, methyl groups were considered to undergo CH···O hydrogen bonding if the C···O distance was less than 3.70 Å, the X-C···O angle was between 85° and 140°, and if elevation angle was less than 50°. The C···O-Y angle was required to be greater than 90° or between 75° and 150° for sp2 or sp3 hybridized oxygen atoms, respectively. In-house software to measure methyl hydrogen bonding criteria was written using MATCH API to manipulate atoms (*182*). Based on this criterion, ~28% of the protein methyl groups surveyed exhibited CH···O hydrogen bonding. Protein methyl groups clearly favor relatively long CH···O interactions (mean of 3.8 Å), consistent with previous investigations of unpolarized methyl CH···O hydrogen bonding in small molecules (*183*). Distance and angular distributions were volume-corrected by multiplying the counts by $1/r^3$ and $1/\sin\theta$, respectively (*184*). Distributions of angles and distances were analyzed using Origin (OriginLab).

To calculate whether the AdoMet methyl CH···O hydrogen bonds were significantly different from that of protein methyl groups, we examined the average length of each of the two populations. The difference between the mean lengths of each population was determined to be 0.31 Å. To establish whether this difference is significant, we recombined both populations, and used the Fisher-Yates algorithm (*185*)

to randomly re-order the data set, followed by division into two subsets containing the same number of hydrogen bonding interactions as the two original populations (127 and 9881 members, respectively), and the difference between average lengths calculated. This process was repeated 100,000 times, generating a normal distribution with an average difference of 0.025 Å and standard deviation of 0.019 Å. These values yield a Z-score of 15.2σ, or a confidence value greater than 99.999%, permitting us to reject the null hypothesis.

In addition to methyl groups, all AdoMet and protein methylene CH···O hydrogen bonds were measured and volume corrected as described above for methyl groups, with the exception that hydrogens were modeled into the riding positions using Chimera (*156*), and used to define CH···O hydrogen bond angles. In addition to the longer distance defined above (<3.70 Å) and the same sp2 oxygen angles, the C-H···O angle was required to be between 110° and 250° for sp3 oxygens, and between 100° and 260° for sp2 oxygens, with a maximum hydrogen elevation angle with the sp2 oxygen plane of 60°.

3.2.2 Protein Expression and Purification

WT SET7/9 and SET7/9 Y335F were expressed as described previously in Chapter 2.2.1 . The gene encoding the catalytic domain of SET7/9 (110-366) was subcloned into the pBAD vector, followed by site-directed mutagenesis to convert Tyr335 to an amber stop codon for introduction of the UAA, pAF. SET7/9 Y335pAF was recombinantly expressed in DH10B *E.coli* containing pDule-pAF vector according to previously published protocol for genetic incorporation of UAAs (*186*). Addition of 1 mM D-pAF during induction yielded 100% labeled protein at half the yield of the WT enzyme. The remaining enzyme was truncated at residue 334. Induction without pAF yielded only

SET7/9 truncated at residue 334. Both SET7/9Y335F and SET7/9 Y335pAF were purified as previously described (*137*), with the exception that gel filtration chromatography was performed using a HiLoad Superdex 75 column (GE Healthcare) for pAF mutant. This chromatography effectively separated the SET7/9 Y335pAF from the truncated enzyme based on the difference in their elution profiles, permitting the latter to be eliminated as a contaminant.

3.2.3 X-ray Crystallography

WT, Y335pAF, and Y335F SET7/9 crystals were grown by hanging drop vapor diffusion at 20° C, as previously described (*130*) with the following exceptions: crystallization solution was mixed in a 1:1 ratio with 10-12 mg/mL SET7/9, 6 mM AdoMet or 3.0-4.5 mM AdoHcy, and 2.0 mM of a 10-residue peptide of TAF10 peptide or TAF10 K189A peptide in 20 mM TRIS pH 8.0, 100 mM NaCl and 2.0 mM Tris(2-carboxyethyl)phosphine. Crystallization solutions contained 0.86-1.07 M sodium citrate with 100 mM imidazole, pH 7.6-8.4 and 8.0-18 mM $NiCl_2$. Y335F and Y335pAF crystals were flash frozen in 1.5 M sodium citrate, with 1.0 mM AdoHcy, 10-13 mM $NiCl_2$ and 100 mM imidazole pH 8.2-8.7 after 2 days, or 3-5 weeks of crystallization time, respectively. For WT SET7/9, crystals were flash frozen in the same cryoprotectant, omitting AdoHcy, 16-28 hours after setup to obtain the AdoMet complex, and 3-5 weeks after setup for the AdoHcy complex. Data were collected at the Advanced Photon Source Synchrotron beamline 21-IDG (LS-CAT), at 100 K and wavelength 0.9786 nm, and were processed using HKL2000 (*187*). The structures were solved by molecular replacement in MOLREP (*188*), with the previously reported protein coordinates of the SET7/9 ternary complex (PDB Accession Code 3M53) and AdoMet-bound complex (PDB Accession Code 1N6A) (47, 130) as the search models. Model building and

refinement were carried out using Coot (180), MiFit (version 2010.10 http://code.google.com/p/mifit), and REFMAC (189), respectively. . The quality of the final model was verified by Molprobity (*190*). CNS was used to generate simulated annealing omit maps for each of the different structures (*191*). In the SET7/9 Y335pAF structure, K189 of the TAF10 peptide, pAF and AdoHcy were omitted. In the SET7/9 Y335F structure, K189, Phe335, and AdoHcy were omitted. In each of the WT SET7/9 structures, His293, Gly264, and cofactor were omitted. RMSD values were calculated and structural figure rendered in PyMol (Schrödinger, LLC). Exhaustive attempts to crystallize the SET7/9 Y335pAF•AdoMet and SET7/9Y335pAF•AdoMet•TAF10K189A complexes were unsuccessful.

3.2.4 Isothermal Titration Calorimetry (ITC)

SET7/9 Y335pAF•TAF10 and SET7/9 Y335F•TAF10 peptide ITC experiments were performed using a MicroCal VP-ITC calorimeter (GE Healthcare), as previously described for WT SET7/9 with identical buffer and temperature conditions (*130*). Stoichiometry constants ranged from 0.81-1.05. Kinetic ITC fits were performed on triplicate WT SET7/9•AdoMet binding experiments, as previously described (*192*).

3.2.5 Intrinsic Fluorescence Binding Assays

The weak cofactor affinities displayed by the SET7/9 Y335 mutants precluded analysis of their binding constants by ITC. To measure cofactor binding to the Y335 mutants, a fluorescence binding assay was developed that quantifies the change in intrinsic fluorescence of the active site tryptophan 352 that stacks adjacent to the cofactor adenine ring (*47*). Dissociation constants for AdoMet and the product AdoHcy were calculated using intrinsic fluorescence measurements in triplicate using a Safire II plate reader (Tecan Group Ltd.) and round-bottom, black Costar 96-well plates

(Corning). Buffer conditions were identical to previous SET7/9 ligand binding studies by ITC (20 mM sodium phosphate pH 7.0 and 100 mM NaCl), as described in Chapter 2.2.6 . Varying concentrations of ligand were added to 1.75 μM of SET7/9 Y335pAF or Y335F, and fluorescence monitored between 320 and 390 nm at 37°C after excitation at 280 nm. The measured fluorescence change was ~3-5 fold, depending on the enzyme and ligand used. Resulting fluorescence profiles were fit to a Gaussian function, and then calculated amplitudes were plotted as a function of ligand concentration, and dissociation constants determined by fitting in Prism (Graphpad). The fraction bound was determined by using the fitted saturation point of each individual experiment as = 1.0. For certain experiments, non-specific binding was additionally observed, as evidenced by a linear trend at high ligand concentration. This non-specific binding was subtracted out using a linear fit model. Resulting binding curves were verified to produce dissociation constants within error of those calculated using a fitting model with a linear non-specific binding term included in the fitting procedure. When measured separately, neither ligand nor SET7/9 showed discernible fluorescence change with increasing concentrations. To control for differences between ITC and fluorescence data, a binding experiment was first conducted for AdoHcy and WT SET7/9, yielding a dissociation constant (60 ± 5 μM and 134 ± 36 μM by ITC and fluorescence experiments, respectively) that differed by only two-fold compared to the K_D value measured by ITC.

3.2.6 Differential Scanning Calorimetry (DSC)

Melting enthalpy and temperatures of WT SET7/9, Y335F and Y335pAF were measured using a nanoDSC calorimeter (TA Instruments), with a scan rate of 1°C/min from 18° to 100°C at 3 atm constant pressure and a sample concentration of 0.9 mg/ml protein. Buffer conditions were identical to binding assay conditions. Data were

analyzed using nanoAnalyze software (TA Instruments), and the figure was prepared using Origin 8.5 (OriginLab).

3.2.7 NMR Spectroscopy

2D NMR and spin-relaxation experiments on WT SET7/9•AdoMet and SET7/9 Y335F•AdoMet were carried out on a 600-MHz Avance Bruker NMR spectrometer equipped with a triple resonance cryoprobe at 298 K. The SET7/9 Y335F sample used for chemical shift determination contained 0.1 mM $^{13}CH_3$-AdoMet (generated as previously described in Chapter 2.2.2) and 1.0 mM SET7/9 Y335F to enable saturation of the cofactor by enzyme. $^{13}CHD_2$-AdoMet experiments were performed in: 20 mM sodium phosphate, pH 7.5, 100 mM NaCl, and 0.5 mM TCEP, in 98% D_2O. $^{13}CHD_2$-AdoMet was enzymatically synthesized from $^{13}CHD_2$-methionine and ATP as previously described in Chapter 2.2.2. NMR data were processed using NMRPipe/NMRDraw(*159*) and iNMR (www.inmr.net), and analyzed using Sparky (*160*).

^{13}C relaxation rates were determined using $^{13}CHD_2$-AdoMet and a perdeuterated protein sample (see above), with a constant $R_1\rho$ spin-lock power of 3500 Hz to eliminate any R_{ex} contributions. ^{13}C relaxation rates were measured using pulse sequences (*193, 194*) that were modified to incorporate deuterium decoupling for the duration of the experiments, as previously described (*173*). WT SET7/9 correlation time (21 ns) was calculated from its structure (1N6A) (*47*) using HYDRONMR (*195, 196*), and D_2O viscosity of 1.100 cP (*197*). Relaxation rates were analyzed as previously described (*174*), with the exception that the methyl C-H distance was set to 1.08 Å, based on our quantum chemistry optimizations of the AdoMet methyl group in the presence of a CH···O hydrogen bonding partner, using in-house written software that incorporates spectral density analysis into an extended Model-Free formalism (*193*). Due to the

inability to generate stable perdeuterated SET7/9 Y335F and Y335pAF, we were unable to measure relaxation rates for these mutants. Relaxation data was fit using Prism 5 (Graphpad).

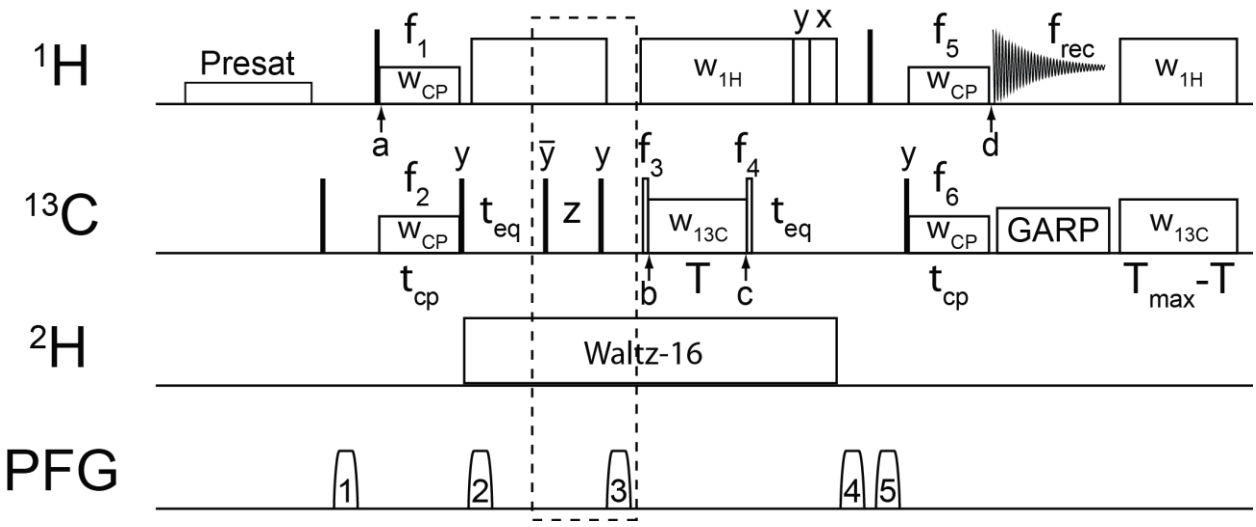

Figure 3.1: Selective ^{13}C $R_{1\rho}$ relaxation and relaxation-dispersion pulse sequence for $^{13}CHD_2$-methyl groups.
Adapted from previously published pulse sequence (*194*). Adapted previously published figure courtesy of Alex Hansen

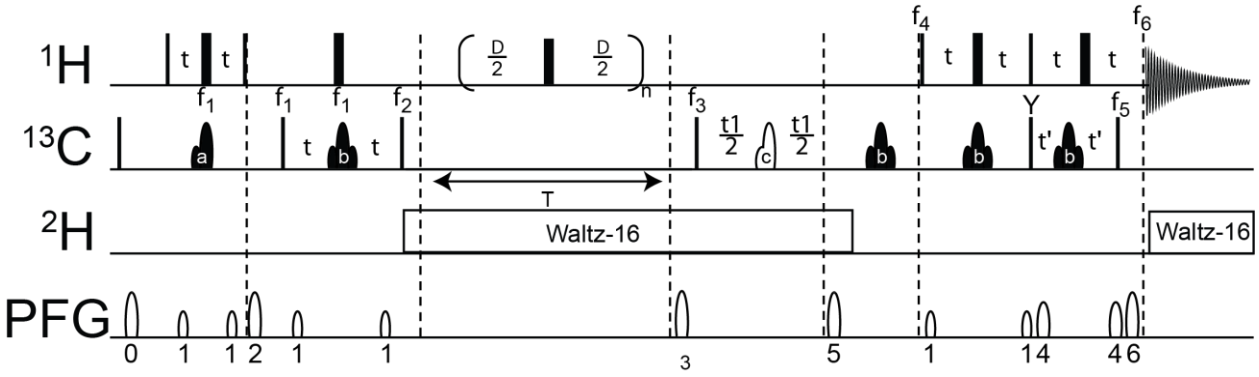

Figure 3.2: ^{13}C R_1 relaxation pulse sequence for $^{13}CHD_2$-methyl groups.
Adapted from previously published pulse sequence (*193*). Adapted previously published figure courtesy of Alex Hansen

Order parameters for AdoMet bound in the active site of methyltransferases were calculated for AdoMet molecules using the same pool of structures as the PDB

survey. AdoMet molecules were aligned by the sulfur and two adjacent methylene carbon atoms using the alignment protocol in PyMol (Schrödinger, LLC). Free AdoMet order parameters were calculated from a 10 ns MD simulation of AdoMet previously verified to correctly model the methyl group electronic environment in Chapter 3.2.9 . The AdoMet molecules in the MD simulation were aligned by the center of mass of the same three atoms as above, followed by an exhaustive rotation search. Order parameters were calculated from these ensembles as previously described (*198, 199*).

3.2.8 Radiometric AdoMet Assays

Purified WT SET7/9 (0.06-0.25 μM), and the mutants SET7/9 Y335pAF (10 μM) and SET7/9Y335F (0.3 μM), and were used in methyltransferase assays (*200*) with biotinylated TAF10 peptide (480 μM) (*130*). Assays contained 25 mM NaCl, 100 mM bicine pH 9.0, various concentrations of [^{3}H-*methyl*] AdoMet (diluted to specific activities from 0.18 to 1.8 μCi/nmole with purified AdoMet (*201*) in a final volume of 20 μL and incubated at 37°C. Assays were terminated at 1, 30, and 2 min for WT SET7/9, SET7/9 Y335pAF, and SET7/9Y335F respectively (rates measured were linear over these times), by addition of an equal volume of 200 mM MES pH 5 prior to addition of a two-fold molar excess of immobilized streptavidin resin (UltraLink; Pierce Biotechnology, Inc.). After 30 minutes at room temperature for binding of the peptide, the resin was collected by centrifugation, washed three times with 300 mM NaCl, and transferred to a vessel for measurement of tritium incorporated during the assay by liquid scintillation spectroscopy. For each activity determination, two separate control assays (without enzyme and without peptide substrate) were conducted in concert for background correction. Data were plotted and then fitted to the Michaelis-Menten equation using Prism (Graphpad).

Single turnover assays using WT SET7/9 were performed in analogous conditions to steady state experiments, except with increased concentrations of AdoMet (60 μM), TAF10 peptide (2.0 mM), and enzyme (70-140 μM), and that the quench was accomplished using an equal volume of 7 M acetic acid. Time points were taken between 0.1 and 13 seconds after reaction initiation. A chemistry rate constant was determined by fitting the data to a single mono-exponential curve (*202*).

3.2.9 Quantum Mechanical (QM) Modeling

Using the crystal structure of SET7/9 Y335pAF•TAF10•AdoHcy, a minimal active site model was generated for the both the SET7/9 Y335pAF•AdoHcy and SET7/9 Y335pAF•AdoMet complex to determine the position of the SET7/9 Y335pAF amine hydrogens and AdoMet methyl group. All hydrogens were added automatically using UCSF Chimera (*156*) followed by ONIOM/DFT methods (*203*) in Gaussian 09 (Gaussian Inc.) with solvation using the Polarizable Continuum Model (*153*). For the SET7/9 Y335pAF•AdoHcy complex, atoms closest to the pAF-amine group were modeled using B3LYP/6-311g+(2d,p) (*151, 204*), including all aniline atoms, the AdoHcy sulfur and adjacent methylene groups, the carbonyl of Ala295, and amide and Cα of Asn296, while remaining atoms were modeled using B3LYP/3-21G* (*157, 158*). All hydrogen atoms and the pAF amine nitrogen were allowed to optimize, while other atoms were held frozen. For the SET7/9 Y335pAF•AdoMet complex, the methyl group was also modeled using B3LYP/6-311g+(2d,p) and allowed to optimize its position.

For the CH···O geometry and energy calculations, all structures were geometry optimized using B3LYP/6-31g+(2d,p) in Gaussian 09 (Gaussian Inc.) with solvation using the Polarizable Continuum Model (*153*). The reactant state structures were calculated with no restraints. The methyl transfer transition state structure was

calculated using QST2 and QST3 algorithms. CH···O hydrogen bonders were geometry optimized with the transition state structure held fixed, or allowed to co-optimize. These two separate geometry optimizations yielded identical geometrical and energetic results. CH···O hydrogen bond gas-phase interaction energies of the geometry-optimized structures were calculated using counterpoise calculations (*205, 206*) with MP2/6-31g+(2d,p) (*207, 208*).

3.3 Results

3.3.1 Methyltransferase Structural Survey

Previous study of SET domain enzymes prompted us to examine whether CH···O hydrogen bonds are present in other classes of AdoMet-dependent methyltransferases. A survey using typical CH···O hydrogen bonding cutoffs (*75*) of 43 high-resolution crystal structures (< 2.0 Å resolution) of AdoMet-dependent methyltransferases, covering all of the major classes revealed that all but one crystal structure displays evidence of at least one CH···O hydrogen bond to the AdoMet methyl group (Figure 3.3, Appendix A.2). Notably, the single outstanding structure (Accession code 1VPT) has a C···O contact whose distance is well within the refinement coordinate error of the survey cut-off distance. Of these, half had multiple AdoMet methyl CH···O hydrogen bonds. Many of these interactions are closer than 3.25 Å, the van der Waals contact distance between a carbon and oxygen atom with no intervening hydrogen atom, similar to conventional hydrogen bonds (*75*). Additionally, we found that in every methyltransferase surveyed, at least one of the AdoMet methylene groups adjacent to the sulfonium cation engages in hydrogen bonding. The extensive structural conservation of AdoMet CH···O hydrogen bonding across diverse methyltransferase classes suggests that it is important to methylation.

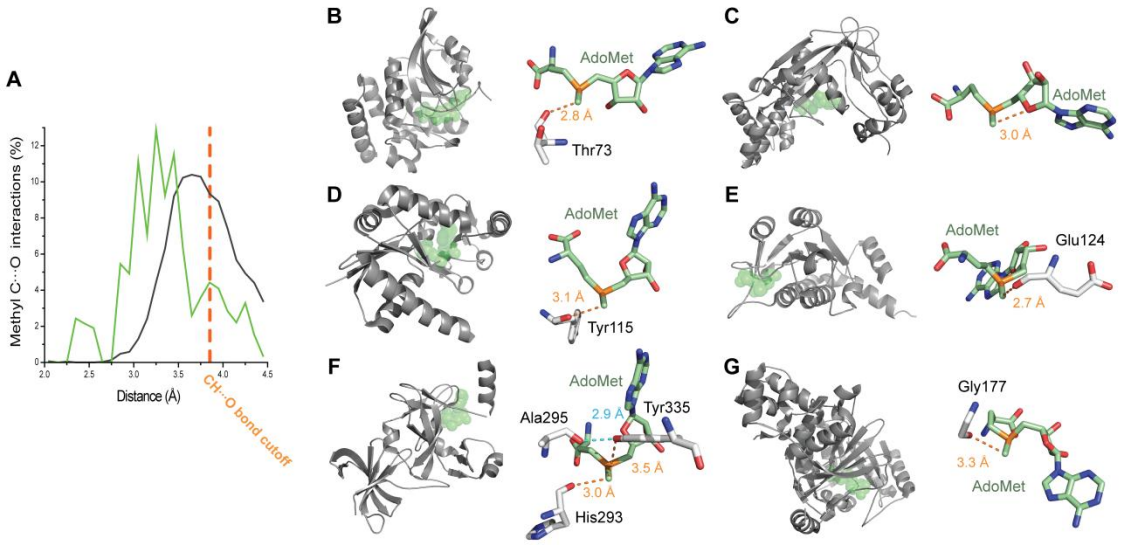

Figure 3.3: Methyl CH···O hydrogen bonding in AdoMet-dependent methyltransferases.
(A) Overlay of normalized histograms of methyl group-oxygen distances for AdoMet (green) and protein methyl groups (gray) that satisfy hydrogen bonding angular criteria (Chapter 3.2.1). Representative examples from the six available classes of AdoMet-dependent methyltransferases illustrating protein or cofactor CH···O hydrogen bonding (orange dashed lines) to the AdoMet methyl group. The hydrogen bond donor and methyl C···O interaction distances are labeled in each enzyme. (B) Class I**:** L-isoaspartyl O-methyltransferase (PDB accession code 1JG4) (*209*); (C) Class II: Methionine synthase reactivation domain (1MSK) (*36*); (D) Class III: CobF precorrin-6A synthase (2NPN); (E) Class IV: TrmH tRNA 2′-O methyltransferase (1V2X) (*210*); (F) Class V: SET7/9 KMT (1N6A) (*47*); and (G) Class VI: RlmN rRNA adenosine methyltransferase (3RFA) (*211*). Although not yet solved at high resolution with AdoMet bound, the other known classes also appear to have the potential for CH···O hydrogen bonding (*212-214*).

The surprisingly short distances of the AdoMet methyl CH···O hydrogen bonds potentially indicates a stronger interaction than typically observed for methyl groups. In addition to distance dependence, angular distributions can be indicative of hydrogen bond strength (*61*). It is well accepted that strong hydrogen bonding correlates with a tendency toward linearity between the donor, hydrogen and acceptor atoms, or in the case of an sp2 acceptor, positioning of the donor within the plane of the acceptor orbitals. As such, the angular dependence of AdoMet methyl hydrogen bonding versus that of other protein methyl groups can be used as a qualitative indicator of relative strength. For this comparison, the elevation angle of the methyl group to the plane of

the sp2 oxygen was used, as there is substantial uncertainty regarding the orbital and hydrogen positions in an sp3 acceptor as determined by X-ray crystallography. The planarity of the sp2 oxygen acceptor orbitals dictates that, in the ideal case, hydrogen bonding donor and acceptor atoms should be nearly coplanar.

The distribution of angles for all protein methyl groups displays a modest tendency towards linearity (Figure 3.4), in contrast to the AdoMet methyl group that exhibits stringent coplanarity (Figure 3.4). This observation implies that AdoMet methyl CH···O hydrogen bonding is stronger than analogous interactions formed by protein methyl groups. The combination of short C···O distances and a strong tendency towards co-planarity observed for AdoMet methyl hydrogen bonding is inconsistent with typical weak methyl CH···O hydrogen bonding (*57*, *61*). A similar analysis for the AdoMet methylene groups yielded a different trend. Unlike the AdoMet methyl group, the AdoMet methylene groups behave very similarly to protein methylenes, indicating that these groups do not form uniquely strong CH···O compared to protein methylene groups (Figure 3.5).

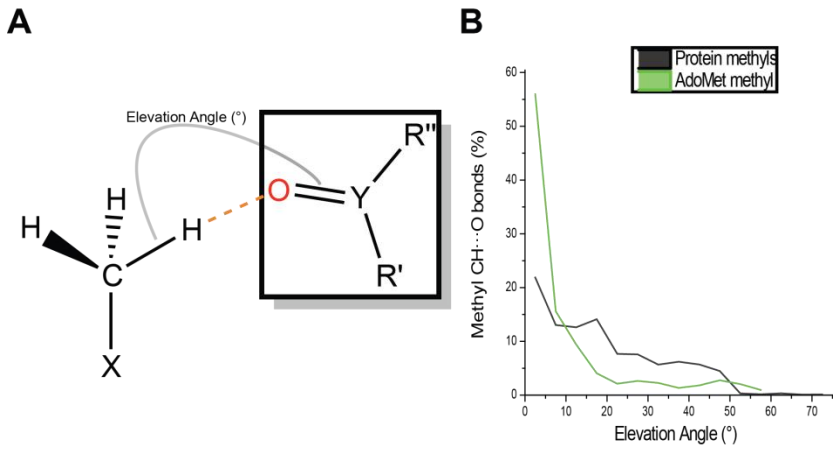

Figure 3.4: Angular distribution of methyl CH···O hydrogen bonds.
Diagram of (A) elevation angles measured in high-resolution crystal structures of methyltransferases, with C···O interaction distances < 3.7 Å that satisfy hydrogen bonding angular criteria (Supplementary Methods). Overlay of normalized distributions of methyl group hydrogen bond elevation angles (B) for AdoMet (green) and all protein methyl groups that form hydrogen bonds (black).

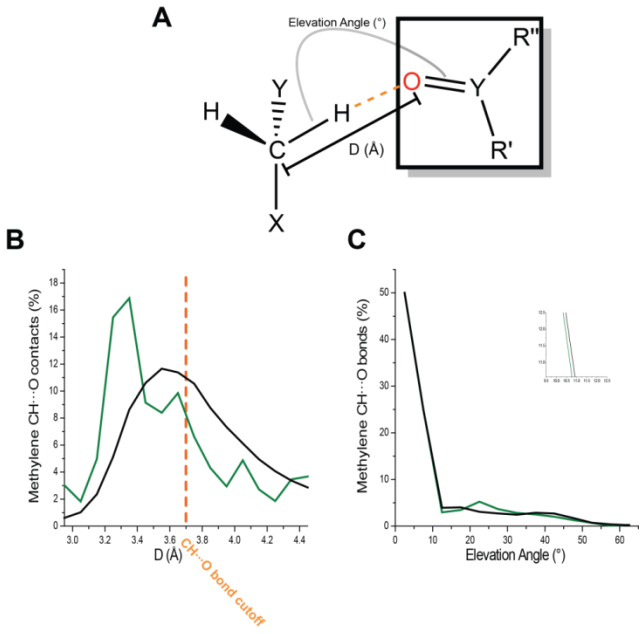

Figure 3.5: Distance and angular distribution of methylene CH···O hydrogen bonds.
Diagram of (A) elevation angles and distances measured in high-resolution crystal structures of methyltransferases, with C···O interaction distances < 3.7 Å that satisfy hydrogen bonding angular criteria (Supplementary Methods). Overlay of normalized distributions of methyl group hydrogen bond distances (B) and elevation angles (C) for AdoMet (green) and all protein methyl groups that form hydrogen bonds (black). Inset in (C) to distinguish overlapping section of elevation angle plot.

3.3.2 Design and testing of CH···O hydrogen bond disruption in SET7/9

We first performed a series of experiments to assess the feasibility of using the Y335F and Y335pAF mutations to probe methyl CH···O hydrogen bonding. Isothermal calorimetry experiments on wild type (WT) SET7/9 and the mutants demonstrated nearly identical protein substrate binding affinity (Figure 3.6). Thermal denaturation studies illustrated that the WT enzyme and Y335pAF mutant maintained equivalent stability to WT, whereas the Y335F mutant was mildly destabilized, consistent with the loss of the Tyr335-backbone hydrogen bond (Figure 3.6). To further examine the effects of these mutants on the enzyme's overall and active site structure, we determined their co-crystal structures of both mutants in complex with AdoHcy and TAF10 peptide (*130*) (Figure 3.7, Figure 3.8, Table 3.1). The coordinates of these complexes and the analogous WT complex superimpose with a Cα RMSD of 0.14 Å and 0.22 Å for the Y335pAF and Y335F mutants, respectively, illustrating that the WT enzyme and mutants are highly structurally similar. Further, the structural hydrogen bond between Tyr335 and the carbonyl group of Ala295 is preserved by the pAF substitution, with ideal hydrogen bond distance and geometry (Figure 3.7). In comparing both mutants to WT SET7/9, all other conventional hydrogen bonds between the enzyme and its ligands, AdoHcy and the TAF10 peptide, appear to be preserved, as well as the conformations and positions of the active site residues. Taken together, these results emphasize that the mutants retains native structure and protein substrate binding while abolishing the CH···O hydrogen bonds between Tyr335 and AdoMet, validating their use in detailed structure-function comparisons.

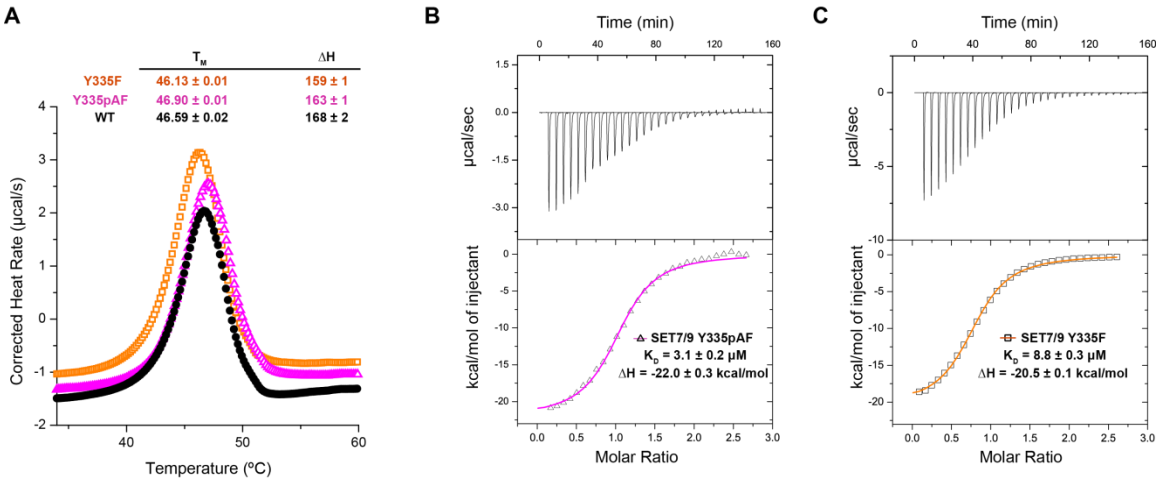

Figure 3.6: Calorimetric analysis of SET7/9 Y335pAF.
(A) Differential scanning calorimetry melting of SET7/9 Y335pAF (△), SET7/9 Y335F (□), and WT SET7/9 (●). For clarity, 10 points were omitted between each point shown. Isothermal titration calorimetry (ITC) analysis of TAF10 binding to (B) SET7/9 Y335pAF (△), and (C) Y335F (□). Top panel represents titration of peptide into SET7/9 Y335pAF, and bottom panel represents binding isotherm with fitted curve. Previous experiments found TAF10 binds WT SET7/9 with 4.9 ± 0.1 µM affinity (Table 1) (*130*).

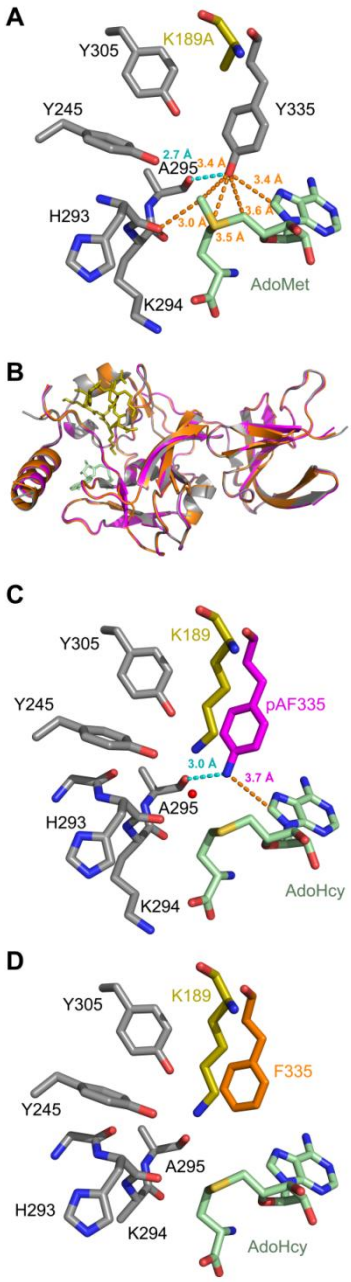

Figure 3.7: Structure of the SET7/9 Y335 mutants bound to AdoHcy and the TAF10 peptide.
(A) Overlay of WT SET7/9 (3M53, gray) (*130*) and SET7/9 Y335pAF (magenta), with Cα RMSD=0.14 Å. AdoHcy and TAF10 peptide are rendered as sticks. (B**)** F_o-F_c simulated annealing omit map (contoured at 2.0 σ) of SET7/9 Y335pAF (magenta carbon atoms), AdoHcy (green carbons), and TAF10-Lys189 (gold carbons). (C) Active site of SET7/9 Y335pAF. The NH···O hydrogen bond between Ala295 and pAF335 and van der Waals contact between pAF335 and the AdoHcy adenine are denoted by cyan and magenta dashes, respectively.

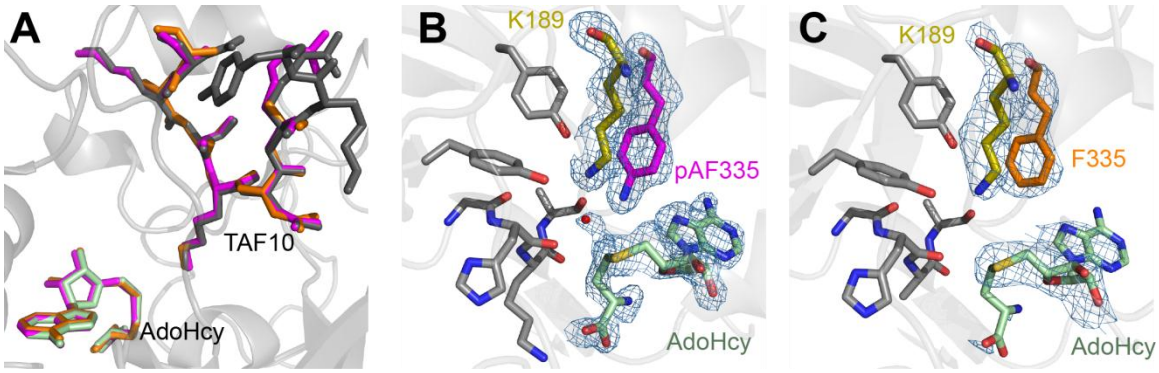

Figure 3.8 Ligand binding modes of WT and SET7/9 Y335pAF are conserved.
(A) Overlay of TAF10 peptide and AdoHcy in WT SET7/9 (3M53) (gray) and SET7/9 Y335pAF (magenta), and SET7/9Y335F (orange). Simulated annealing omit maps calculated at 2.0 σ for (B) Y335pAF and (C) Y335F crystals structures.

Table 3.1: Crystallographic statistics for structures of SET7/9 Y335pAF•TAF10•AdoHcy, SET7/9 Y335F•TAF10•AdoHcy, SET7/9•TAF10 K189A•AdoHcy, and SET7/9•TAF10 K189A•AdoMet.

	SET7/9 Y335pAF•TAF10•AdoHcy	SET7/9 Y335F•TAF10•AdoHcy	SET7/9•TAF10 K189A•AdoMet	SET7/9•TAF10 K189A•AdoHcy
Data collection				
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Cell dimensions				
a, *b*, *c* (Å)	83.0, 83.0, 95.5	82.3, 82.3, 95.3	83.5, 83.5, 95.8	87.2, 87.2, 94.2
α,*β*,*γ* (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	50.00-1.60(1.64-1.60)	50-2.56(2.62-2.56)	50.00-1.70(1.74-1.70)	50.00-1.63
R_{merge} (%)	7.7(44.2l)	10.5(51.6)	7.1(48.3)	6.3(56.2)
I/s*I*	26.7(3.1)	15.5(2.0)	27.4(7.1)	30.5(4.9)
Completeness (%)	99.4(94.9)	97.8(84.4)	99.9(100)	99.9(100)
Redundancy	10.7(4.8)	3.8(3.2)	11.1(11.1)	11.1(11.1)
Refinement				
Resolution (Å)	1.60	2.56	1.70	1.63
No. reflections	48193	12438	40793	48994
R_{work}/ R_{free}	0.20/0.22	0.20/0.26	0.19/0.21	0.19/0.23
No. atoms	2187	1868	2160	2228
Protein	1941	1838	1935	1920
Ligand/ion	75	25	66	76
Water	171	5	159	232
B-factors	**24.5**	**44.2**	**24.8**	**23.8**
Protein	23.4	46.9	24.6	23.2
Ligand/ion	34.9	71.1	31.5	27.9
Water	32	37.6	34.3	34.2
R.m.s deviations				
Bond lengths (Å)	0.014	0.014	0.015	0.015
Bond angles (°)	1.51	1.38	1.52	1.51
MolProbity Score				
Percentile	84%	94%	87%	92%
Resolution Range (Å)	1.59 ± 0.25	2.56± 0.25	1.70 ± 0.25	1.63 ± 0.25
Ramachandran				
Favored (%)	95.4	93.0	96.3	95.9
Allowed (%)	4.6	7.0	3.7	4.1
Outliers (%)	0	0	0	0

To verify that the Y335pAF mutation does not alter methyl group positioning, we examined the interactions between the AdoMet methyl group and the mutant by modeling AdoMet into the mutant's active site based on the coordinates of AdoHcy

using density functional theory calculations. The calorimetry experiments indicated that the Y335pAF substitution does not alter the charge of SET7/9 at this position in the active site, as expected due to the low pKa (~5.0) of the aniline group (*215*), such that its amine group could be confidently modeled as neutral.

After energy optimization, one of the pAF amine hydrogen atoms forms the expected hydrogen bond to the Ala295 carbonyl group (Figure 3.7), whereas the second amine hydrogen atom occupies the space adjacent to the methyl group, between the sulfur and the carbon atoms of the S-C bond (Figure 3.9). To avoid steric clashes with AdoMet methyl hydrogen atoms, the methyl group rotates about 5° relative to that in the WT enzyme, and the pAF amine group translates 0.2 Å away from the AdoMet molecule. Notably, the AdoMet methyl carbon position is virtually unchanged compared to the WT enzyme, indicating any decrease in the reaction rate is not likely due to an improper methyl carbon alignment in the SET7/9 Y335pAF mutant. Using MP2 QM calculations, the 5° rotation of the methyl group was calculated to have little impact on CH···O hydrogen bond or rotameric energy.

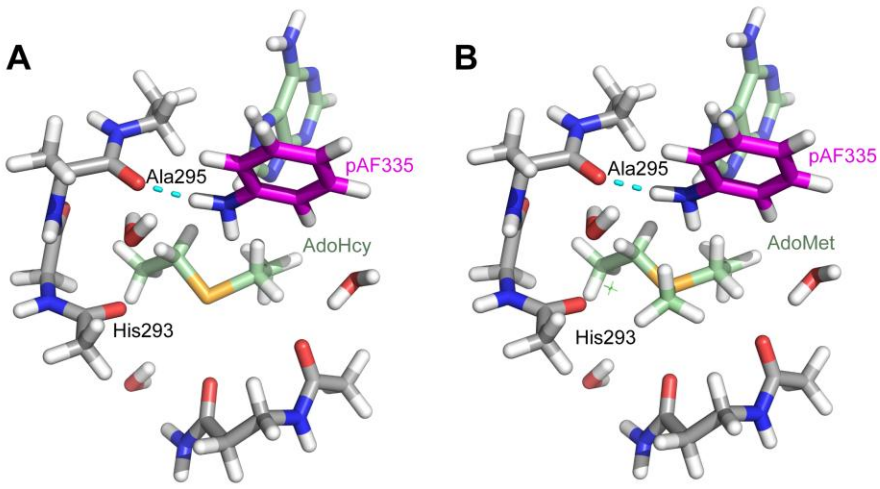

Figure 3.9: Model of SET7/9 Y335pAF hydrogen positions based on its crystal structure and density functional theory calculations.
Bound with AdoHcy (**A**) and AdoMet (**B**) pAF is rendered with magenta carbon atoms, and AdoHcy/AdoMet in green, while backbone carbons are rendered in gray. The structural hydrogen bond from the pAF amine to Ala295 is shown in cyan. The AdoMet methyl group is rotated 5° from its position in WT, as described in Chapter 2.3 , with negligible translational change in position.

As studies described in Chapter 2 demonstrated that the AdoMet methyl ^{1}H chemical shift is highly dependent on its hydrogen bonding state, we analyzed the AdoMet methyl ^{1}H chemical shift change when bound in the SET7/9 Y335F active site (Figure 3.10). Measurement of the AdoMet methyl ^{1}H chemical shift while bound to the Y335F mutant yielded a small upfield chemical shift change (0.1 ppm), indicative of the loss of the longer CH···O hydrogen bond between the AdoMet methyl and Tyr335 hydroxyl groups (Figure 3.10). Attempts to measure the AdoMet methyl chemical shift when bound to SET7/9 Y335pAF were unsuccessful due to the low binding affinity and intermediate exchange rates, as kinetic ITC analysis found that AdoMet binding is in the ms-μs regime Figure 3.11.

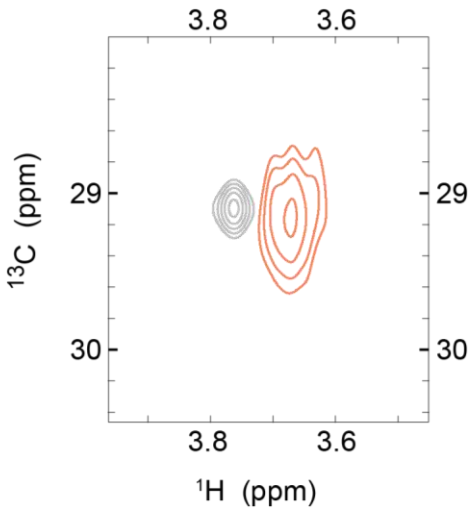

Figure 3.10: WT SET7/9•AdoMet and SET7/9 Y335F•AdoMet methyl ^{1}H chemical shift measurement.
WT (gray) and Y335F mutant (orange) ^{13}C 2D-HSQC spectra are overlaid. The Y335F mutation, which removes only the longer CH···O hydrogen bond, causes a small upfield ^{1}H chemical shift change relative to WT SET7/9. The increased linewidth and lower signal to noise of the AdoMet methyl resonance when bound to the Y335F mutant is likely due to lower stability of the Y335F mutant preventing long experiment time, and decreased AdoMet concentration to enable saturation by SET7/9 Y335F.

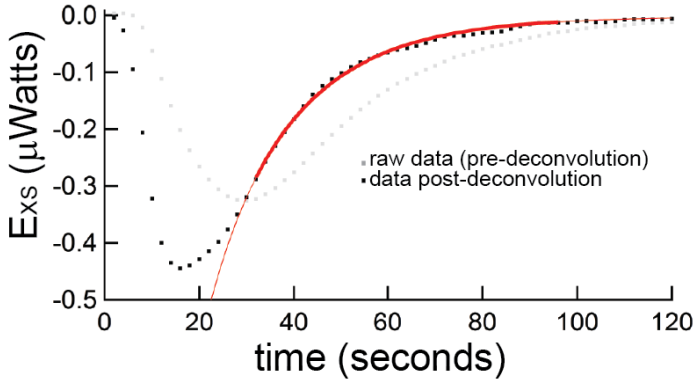

Figure 3.11: Representative injection peak and best-fit kinetic parameter trace from AdoMet•WT SET7/9 ITC experiments.
The red curves are nonlinear regression fits as previously described (*192*). Thick line represents the regression over the time period used in data fitting (31 seconds to 95 seconds). Faint red line is the extrapolation of the regression curve. AdoMet k_{on} was determined from triplicate experiments to be slower than diffusion limited = 23800 ± 2800 M^{-1} s^{-1}.

3.3.3 AdoMet Binding and Catalysis by SET7/9 Mutants

We next examined whether either the Y335F or Y335pAF mutations impaired cofactor binding or catalytic efficiency. The Y335F and Y335pAF mutations diminished AdoMet binding affinity by three and four orders of magnitude, respectively (Table 3.2, Figure 3.12), demonstrating that the loss of Tyr335-mediated CH···O hydrogen bonding substantially impaired cofactor binding. Despite the dramatic change in binding affinity, the change in catalytic rates of the mutants was comparatively small. The Y335F mutant exhibited a k_{cat} nearly identical to WT SET7/9, whereas the k_{cat} of Y335pAF was reduced about 60-fold relative to WT (Table 3.2, Figure 3.13). In terms of energetic stabilization of the transition state using the Eyring equation, this amounts to a 2.2 kcal/mol, or 10% increase, of the activation barrier. Single turnover experiments indicated that the methyl transfer is rate is nearly within error of the multiple-turnover k_{cat}, and well within 2-fold of each other, demonstrating that methyl transfer is limiting in SET7/9 , and thus the small changes can be attributed to changes in methyl transfer (Figure 3.14)

Table 3.2: WT SET7/9 and mutants Y335pAF, Y335F dissociation constants and catalytic parameters. AdoMet K_M was calculated using a saturating concentration of TAF10 peptide and varying AdoMet concentrations. [a,b]Isothermal titration calorimetry (ITC) values taken from Chapter 2.3 and Del Rizzo *et al.* (*130*), respectively. Measured by [c]intrinsic tryptophan fluorescence assay or [d]ITC. *High error for SET7/9 Y335F binding to AdoHcy due to large non-specific binding affects.

	K_D (μM)		
	WT	**Y335pAF**	**Y335F**
AdoMet	[a]0.053 ± 0.003	[c]976 ± 209	[c]52 ± 26
AdoHcy	[c]134 ± 36	[c]839 ± 166	[c]711 ± 500*
TAF10	[b]4.9 ± 0.1	[d]3.1 ± 0.2	[d]8.8 ± 0.3
	Catalytic Constants		
	WT	**Y335pAF**	**Y335F**
k_{cat} (min^{-1})	19.3 ± 1.5	0.34 ± 0.028	12.9 ± 0.8
k_{cat}/K_M ($\mu M^{-1}min^{-1}$)	3.0 ± 0.9	0.0018 ± 0.00044	0.18 ± 0.041

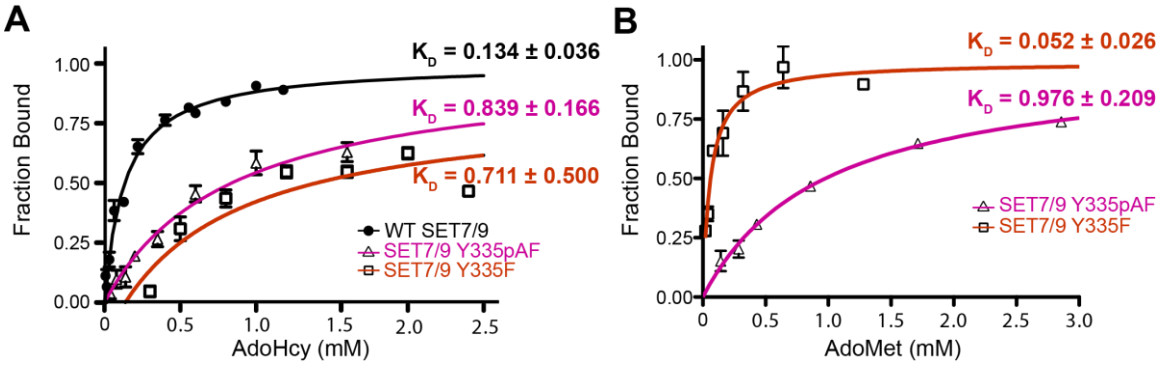

Figure 3.12: Measurement of cofactor binding affinity by intrinsic tryptophan fluorescence. (**A**) Binding of AdoHcy to WT SET7/9 (●), SET7/9 Y335pAF (△), and SET7/9 Y335F (□). (**B**) Binding of AdoMet to SET7/9 Y335pAF (△) or SET7/9 Y335F (□). All dissociation constants are reported in mM. WT SET7/9 and AdoMet dissociation constant previously measured as 0.053 ± 0.03 μM by ITC (Table 1) in Chapter 2.3 .

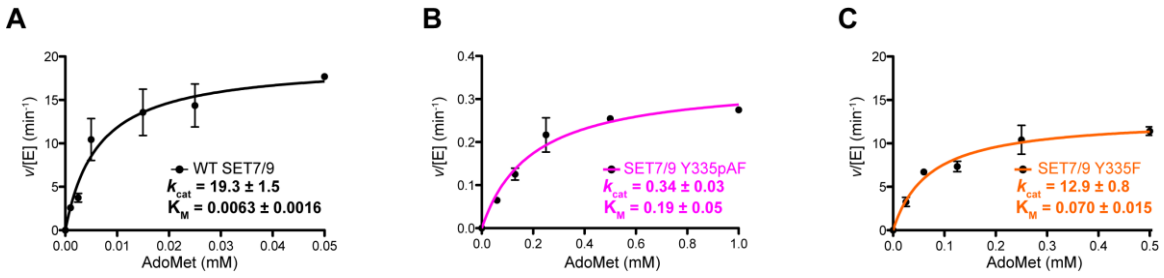

Figure 3.13: Radiometric methyltransferase assays with varying AdoMet concentration.
Data points and Michaelis-Menten fits are shown for (**A**) WT SET7/9 (black line) (**B**) SET7/9 Y335pAF (magenta line), and (**C**) SET7/9Y335F (orange line). All k_{cat} values are listed in min^{-1}, and AdoMet K_M values are reported in mM.

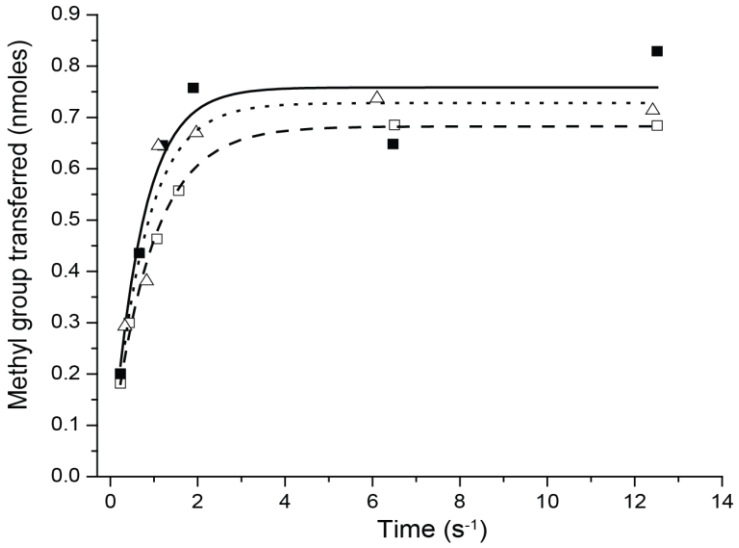

Figure 3.14: Triplicate single turnover experiments on SET7/9 at 37° C.
Average parameters from three separate fits: k_{chem} = 1.28 ± 0.19 s^{-1}, maximum product produced = 0.72 ± 0.04 nmoles.

3.3.4 Experimental CH···O Hydrogen Bond Strength

The SET7/9 Y335pAF mutant afforded an unexpected opportunity to experimentally assess the effective strength of a CH···O hydrogen bond in SET7/9•AdoHcy binding. The cofactor binding constants measured for WT SET7/9 and the Y335pAF mutant revealed a six-fold loss in AdoHcy affinity (Table 3.2). This diminished affinity was surprising given that AdoHcy lacks the ability to form methyl

CH···O hydrogen bonds, and as described above, the AdoMet methylene groups do not form strong hydrogen bonding interactions. Of importance, the pAF amine group cannot serve as an effective hydrogen bond acceptor, as its lone pair is in resonance with the ring To understand the defect in AdoHcy binding, we examined the superimposition of the SET7/9 Y335pAF•TAF10•AdoHcy peptide structure with the analogous WT complex. The aligned structures reveal a subtle shift in the position of the AdoHcy adenine group relative to the pAF335 side chain in the mutant ternary complex compared to the WT enzyme (Figure 3.7). Examining the interactions between the cofactor's adenine group and the invariant Tyr335 via a survey of 16 cofactor-bound structures of SET7/9 (resolution < 2.0 Å) illustrated a mean distance of 3.42 ± 0.06 Å between the Tyr335 hydroxyl group and the C8 atom in the cofactor's adenine group. These distances are indicative of an aromatic CH···O hydrogen bond between the Tyr335 hydroxyl and adenine C8 atom (denoted as the C^8H···O hydrogen bond). Conversely, in SET7/9 Y335pAF, the pAF amine group and the adenine C8 atom are separated by 3.71 Å with a refinement coordinate error of ± 0.05 Å (Figure 3.7). This distance is statistically different from those measured in the WT enzyme complexes (Table 3.3) and is indicative of a van der Waals contact between the pAF amine group and the adenine C8 atom in contrast to a hydrogen bond. Thus, these results reveal an unexpected role for the invariant tyrosine in adenine C^8H···O hydrogen bonding, in addition to its interactions with the AdoMet methyl and methylene groups.

Table 3.3: All AdoMet-C^8···Tyr335-OH distances in high-resolution (<2.0 Å) SET7/9 crystal structures.

PDB	C^8H···Tyr335 OH Distance (Å)
1N6A	3.28
2F69	3.35
3M53	3.46
3M54	3.45

3M55	3.45
3M56	3.45
3M57	3.47
3M58	3.37
3M59	3.38
3M5A	3.35
1XQH	3.41
1O9S	3.45
3OS5	3.48
3CBM	3.44
3CBO	3.48

As the shift in the cofactor adenine moiety was the only notable change in the SET7/9 Y335pAF structure, we attributed the six-fold decrease in AdoHcy affinity to the loss of the adenine $C^8H\cdots O$ hydrogen bond. Based on this decrease, we estimated the stabilization energy of this interaction at $\Delta\Delta G \approx -1$ kcal/mol, taking into account the difference in AdoHcy binding constants of WT SET7/9 and Y335pAF mutant. This free energy value is on the order of weak conventional hydrogen bond stabilization energy, demonstrating that CH···O hydrogen bonding can be relatively strong. Further, quantum calculations suggested that the adenine C8 forms CH···O hydrogen bonds with the tyrosine hydroxyl with an optimal distance of ≈ 3.5 Å, and gas-phase hydrogen bond energy of -1.4 kcal/mol, in good agreement with our measured $\Delta\Delta G$. The AdoHcy methylene groups do not form energetically favorable CH···O hydrogen bonds, and thus do not contribute to the loss of binding affinity. Finally, the $C^8H\cdots O$ hydrogen bond strength reported here is in good agreement with a previous experimentally determined value of a biological CH···O hydrogen bonds as discussed in Chapter 1.

3.3.5 Quantum Chemistry Calculations on Reactant and Transition States

To further evaluate our binding and kinetics data, we used quantum chemistry calculations to determine the energetic effect of CH···O bonding on AdoMet binding and transition state stabilization (Figure 3.15). After geometry optimization to either the

transition or reactant state, interaction energies were calculated between these states andoxygen acceptors. In both the reactant and transition states, CH···O hydrogen bonding was highly favorable, and of a similar strength to conventional hydrogen bonds, between -5-10 kcal/mol (*216*). Interestingly, the CH···O hydrogen bonds energetically stabilize both the transition state and reactant state similarly (Figure 3.15). Approximately similar stabilization of both the reactant and transition state is analgous to a substantial change in binding affinity, but a relatively minor effect on the turnover number (*217*), consistent with our measured binding and kinetics table for WT SET7/9 and the Tyr335 mutants (Table 3.2). Accordingly, theTyr335-mediated CH···O hydrogen bonds promote tight binding to AdoMet as well as the transition state, but do not specifically recognize and stabilize the transition state in preference to the reactant state.

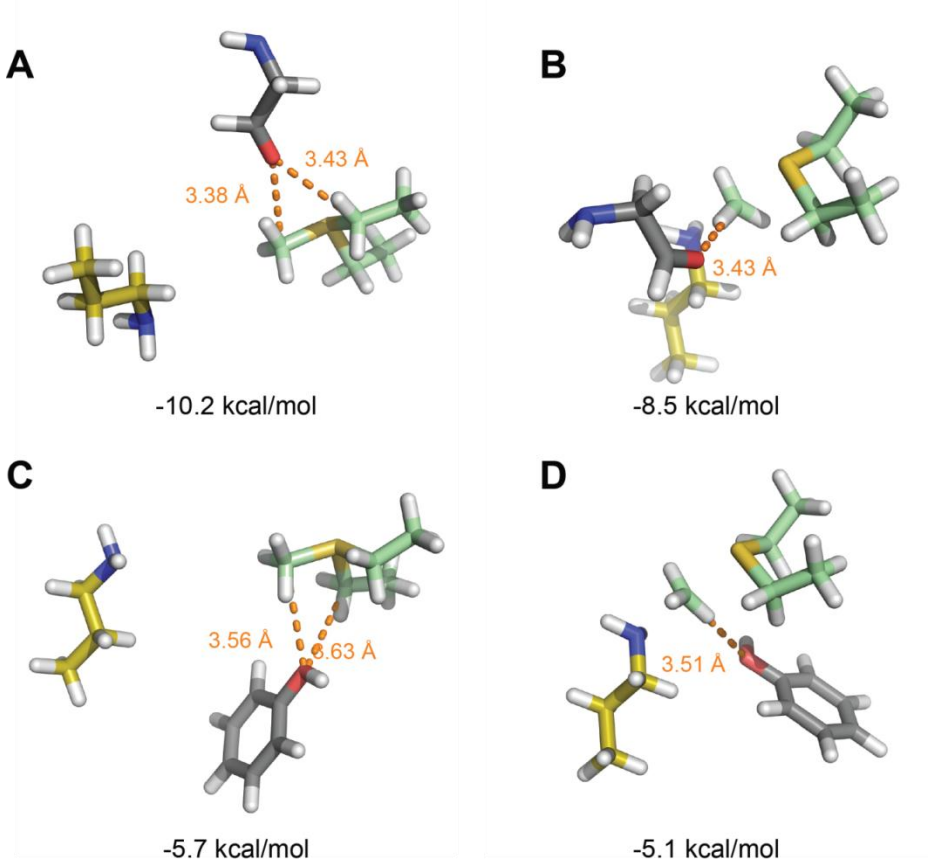

Figure 3.15: Quantum chemistry geometry optimizations and interaction energies of AdoMet and methyl transfer transition state CH···O hydrogen bonds.
(A) Reactant and (B) transition state CH···O hydrogen bonds to backbone carbonyl group, respectively. (C) Reactant and (D) transition state CH···O hydrogen bonds to tyrosine side chain, respectively. C···O distances are depicted in orange.

3.3.6 Methyl Group Dynamics and Pore Constriction

As previously mentioned, one mechanism by which methyl transfer can be catalyzed is through NAC formation. It is conceivable that Tyr335, in combination with the backbone CH···O hydrogen bond between AdoMet and the His293 carbonyl group, may contribute specifically to catalysis through methyl group positioning To evaluate whether CH···O hydrogen bonds could contribute to NAC formation by restricting methyl group motion while bound to the active site, we undertook experiments to

examine how the aperture of the oxygen-lined methyl transfer pore may change during catalysis. To accomplish this, we solved two additional structures of WT SET7/9 ternary complexes bound to a non-reactive peptide and either AdoMet or AdoHcy to mimic the movements of the methyl transfer pore throughout the methyl transfer reaction.

Analysis of these structures in combination with other high-resolution liganded structures of SET7/9 indicates that the methyl transfer pore contracts as the methyl group translates from the reactant to the product state (Figure 3.16). The SET7/9•AdoMet binary complex exhibits the largest methyl transfer pore area, due to the methyl group's steric footprint, with no counteracting electrostatic effects or cooperative protein motion from TAF10 peptide binding to offset methyl transfer pore widening (Figure 3.16). Similarly, the SET7/9•TAF10 K189A•AdoMet complex also displays a relatively wide methyl transfer pore due to the presence of the methyl group, but cooperative TAF10 peptide binding in the ternary complex induces further active site closure, thus decreasing methyl transfer pore size. The SET7/9•TAF10 K189A•AdoHcy complex is virtually identical to that of the AdoMet complex, but lacks the cofactor methyl group, resulting in a further contraction of the methyl transfer pore. By comparison, the SET7/9•TAF10 K189me1•AdoHcy product complex contains a methyl group in the active site, but its position is adjacent to the methyl transfer pore and thus does not occupy the full steric space of the methyl group observed in the AdoMet complexes. Additionally, the presumptive positive charge of the TAF10 Lys189 ε-ammonium cation in this structure (*54, 130, 218*) electrostatically attracts the partially negatively charged oxygen atoms that form the methyl transfer pore, promoting methyl transfer pore contraction. Consequently, the SET7/9•TAF10•AdoHcy complex, which contains the positively charged TAF10 K189 side chain but lacks a methyl group, yields the smallest methyl transfer pore size. In summary, the steric footprint of the methyl

group forces the methyl transfer pore to be larger in the reactant state than in the transition or product states. As such, the methyl transfer pore could then sterically restrict the motion of the methyl group, and favor NAC formation.

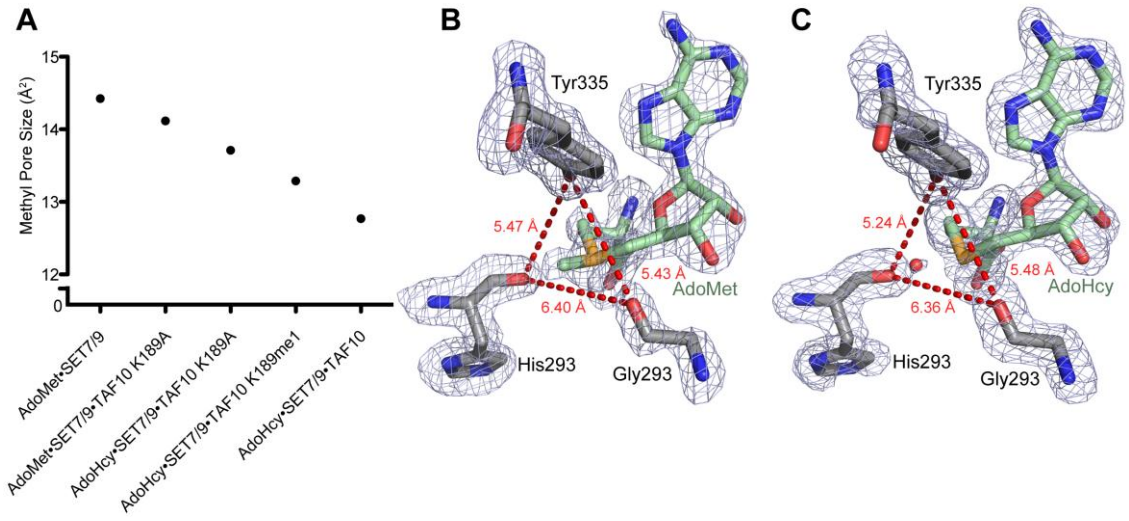

Figure 3.16: Methyl transfer pore contraction in SET7/9.
(A) Methyl transfer pore size decreases throughout model methyl transfer reaction. Methyl transfer pore is defined as the triangle formed by the carbonyl oxygen atoms of His293 and Gly264, and the hydroxyl group of Tyr335. PDB accession codes for the other SET7/9•TAF10 complexes are, in order of appearance: 1N6A (*47*), 2F69 (*169*), and 3M53 (*130*). (B, C) Simulated annealing omit maps and methyl transfer pore sizes for crystal structures of SET7/9•TAF10 K189A•AdoMet and SET7/9•TAF10 K189A•AdoHcy contoured at 2.5 σ, respectively, with methyl transfer pore triangle shown in red dashes.

To further probe this question, we analyzed the motion of the AdoMet S-C_{methyl} axis using the generalized order parameter (S^2_{axis}), which measures the amplitude of motion of the methyl group on a scale of 0 to 1. Importantly, it can both be back-calculated from structural ensembles as well as derived using Model-Free analysis and NMR spin-relaxation experiments. . For a point of reference for our analysis, the average alanine methyl S^2_{axis} in proteins, which shares some stereochemical similarity to AdoMet due to its relatively fixed sp3 Cα atom, was previously shown to be 0.8 (*219*). To evaluate the extent of motion of the AdoMet methyl group when not bound to the

methyltransferase active site, we calcaluted the S^2_{axis} for the free AdoMet using an MD simulation previously verified to accurately reflect the level of AdoMet methyl CH···O hydrogen bonding in water (*220*), and found the S^2_{axis} = 0.67. Then, to compare the methyltransferase-bound S^2_{axis} with the free AdoMet, we calculated S^2_{axis} from the methyltransferase crystal structures used in the PDB survey described above, and found that in the crystal structures, the AdoMet methyl group was fixed in position to a greater degree than in solution or for the analogous alanine methyl groups, with the S^2_{axis} = 0.96. Finally, to confirm this observation, we directly measured by NMR ^{13}C spin-relaxation experiments the methyl motion of AdoMet while bound to SET7/9 (Figure 3.17). Using the Model-Free formalism (*174*), the AdoMet methyl group was observed to have an order parameter (S^2_{axis}) of 1.2 ± 0.2, indicating that the S-C axis of the methyl group experiences little to no motion within the methyltransferase active site, and substantially less motion than when free in solution. Thus, the oxygen-lined pore restricts the motion of the AdoMet methyl group in the active site, optimizing catalytic alignment and favoring NAC formation.

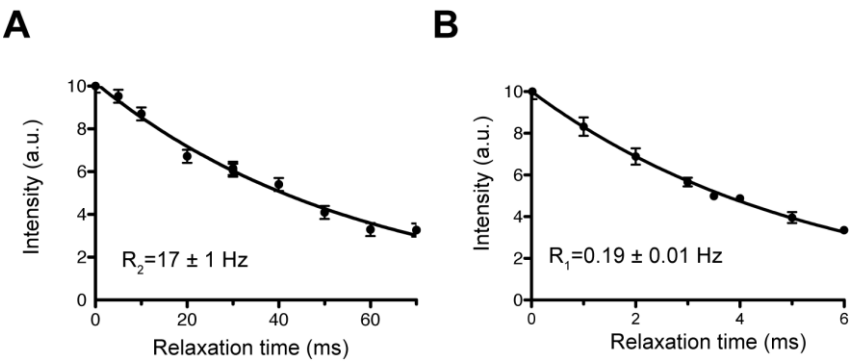

Figure 3.17: (A) R_2 and (B) R_1 ^{13}C spin relaxation fits of $^{13}CHD_2$-methyl AdoMet bound to SET7/9.

3.3.7 Discussion

Collectively, the data presented here illustrate that methyl CH···O hydrogen bonds are a highly conserved and important contributor to AdoMet-dependent methylation (Figure 3.18). In SET7/9, the Y335 CH···O hydrogen bonds are important for cofactor binding and stabilizing the transition state. Further, the methyl transfer pore formed by CH···O hydrogen bond acceptors is responsible for limiting the motion of the AdoMet methyl group, and may contribute to NAC formation. These findings point toward a convergent evolutionary model that is contingent upon methyl CH···O hydrogen bonding.

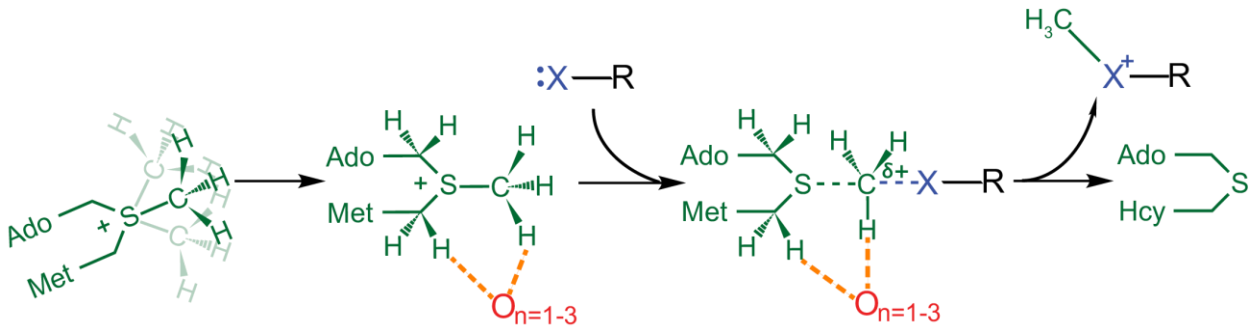

Figure 3.18: Model of AdoMet-dependent methylation.
CH···O hydrogen bond formation restricts methyl movement and stabilizes both the reactant sp2 transition states.

As described in Chapter 3.3.2, neither the Y335F nor pAF mutants disrupted the CH···O hydrogen bond between the AdoMet methyl group and the backbone carbonyl moiety that is strictly conserved in the SET domain class. It is possible that the backbone CH···O hydrogen bond has a different role than the Y335 hydrogen bond in catalysis. Although it is difficult to analyze the effects of this CH···O hydrogen bond due to its immutable nature, future studies may be able to differentiate the role of this hydrogen bond by analyzing the properties of the Y335 mutants in varying conditions that probe the strength and purpose of the remaining backbone CH···O hydrogen bond.

The Y335pAF mutation was employed to specifically perturb the SET7/9 active site while permitting formation of the backbone-Y335 OH···O hydrogen bond. Still, a small defect (about 10% of the barrier) in methyl transfer was detected that was not present in the Y335F mutant. There are multiple possibilities that could account for this small change. One explanation is that methyl dynamics is greater in this mutant due to the loss of CH···O hydrogen bonding. That the Y335F mutant does not display a similar loss in catalysis could be an artifact, as it may be masked by a change in the contribution of the lost OH···O hydrogen bond between the bound and unbound SET7/9 states. Another possibility is that, as has been suggested in other methyltransferases and discussed in Chapter 1, SET7/9 undergoes active site compression to facilitate methyl transfer. The position of the extra hydrogen atom between the sulfur and methyl group of AdoMet could inhibit AdoMet motions parallel to the S-C_{methyl} axis that might be vital for active site compression. Further, the NMR relaxation methods used here are insensitive to these motions. Future studies using 2D infrared spectroscopy may be able to detect these motions, and test for them specifically in the pAF mutant.

Chapter 4: **Conclusions and Future Directions**

With respect to AdoMet-dependent methylation, the experiments presented here demonstrate: 1) the existence of CH···O hydrogen bonding in the active site of SET domain methyltransferases, 2) the importance of these interactions in methylation by SET7/9, 3) that CH···O hydrogen bonding stabilizes both the reactant and transition states of methyl transfer and aids in NAC formation in SET7/9, and 4) the strong possibility that these interactions and roles are conserved in all classes of AdoMet-dependent methyltransferases. Together, these results suggest that CH···O hydrogen bonding is a universal element of methyl transfer, and potentially a driving force of methyltransferase convergent evolution. Despite these discoveries, we are still unable to provide a full description of the importance of the importance or roles of CH···O hydrogen bonds in SET7/9. Until a method is developed to successfully disrupt all CH···O hydrogen bonding in the active site of SET7/9 simultaneously, as well as one at a time, we will not be able to appreciate the full affect CH···O hydrogen bonding has on both cofactor binding and catalysis in this model methyltransferase.

In addition to experiments that may be able to disrupt more, or different, CH···O hydrogen bonds, other methyltransferase experiments may still be most appropriate to attempt using SET7/9. As an example, 2D infrared spectroscopy experiments in collaboration with the Kubarych group could be used to evaluate the dynamics of the bound AdoMet methyl group at ultrafast timescales, and potentially to follow the dynamics of the methyl group as methyl transfer proceeds. These experiments could define the reaction coordinate of the methyl transfer reaction at a level of detail

unprecedented in enzymology. SET7/9 is an ideal choice for such experiments, due to its high solubility, tight AdoMet binding, and good yield from bacterial expression.

Despite the advantages of using SET7/9 as a model system, the experiments described in this dissertation represent only a single methyltransferase case study. To elaborate, we found using preliminary single turnover assays at lower temperature that in WT SET7/9 the methyl transfer reaction is relatively temperature insensitive, suggesting that the activation barrier to methylation is surprisingly small. Temperature dependent experiments on other methyltransferases, however, have shown a wide variation in activation barriers of enzymatic methylation, from 7 kcal/mol to 19 kcal/mol (*221, 222*). This variability suggests that the various forces that facilitate methylation are not present in the same magnitude for all methyltransferases. Correlatively, it is possible that CH···O hydrogen bonding plays different roles either within or between different methyltransferase classes. Although the SET domain is well-conserved within Class V enzymes, there are variations that lead to differing hydrogen bond distances and angles. For example, within Class I, the large divergence in sequence in the AdoMet-binding pocket has led to a wide variety of CH···O hydrogen bond acceptors in unique positions. Seemingly the most common acceptor in Class I methyltransferases is the backbone carbonyl group, similar to Class V enzymes. In Class I, a backbone carbonyl acceptor is usually found in a position analogous to that of the invariant tyrosine hydroxyl group in Class V (Appendix A.2). As carbonyl and tyrosine acceptors have differing properties in terms of preferred angles, distances, and hydrogen bond strength, this switch could change how CH···O hydrogen bonding affects cofactor binding and methyl transfer. As a result, similar investigations to those presented here should be carried out in all methyltransferase classes to fully understand how CH···O hydrogen bonding contributes to AdoMet-dependent methylation.

As such, we have begun to investigate other methyltransferase classes. As described in Appendix A.1, we have used chemical shift measurements to determine that the Class I enzyme TylM1 forms CH···O hydrogen bonds, similar to the SET domain enzymes. Efforts are currently underway in the Trievel group to prepare members of each of the other methyltransferase classes for evaluation by chemical shift in a similar manner, or through infrared spectroscopy. We envision these tests as a medium-throughput diagnostic to explore CH···O hydrogen bond formation within multiple methyltransferases.

Although the chemical shift experiments described in Chapter 2 provide direct evidence for the existence of CH···O hydrogen bonds in the SET domain active site, these experiments are not sufficient to obtain high-resolution details of the geometric properties of CH···O hydrogen bonds. Neutron crystallography, as discussed in Chapter 1, is a fast-evolving technique capable of providing a detailed snapshot of methyl group positioning (*87*). As presented in Appendix A.3, we now have the ability to crystallize perdeuterated SET7/9 in complex with methyl-deuterated AdoMet, which facilitates neutron crystallography experiments. Further, in collaboration with the Holden group at the University of Wisconsin-Madison, we have evidence that TylM1 crystals can be grown to the dimensions required for neutron crystallography. In collaboration with Dean Myles at Oak Ridge National Laboratory, we are preparing to utilize the next-generation neutron crystallography beamline, MaNDI, which can be used to collect high resolution neutron diffraction data on biomolecular crystals that are as small as a tenth the volume as needed at traditional neutron beamlines. Using this new resource, we anticipate being able to solve high-resolution neutron structures of multiple methyltransferases to directly observe CH···O hydrogen bond formation and properties within methyltransferase active sites.

In addition to chemical shift and neutron crystallography experiments on TylM1, we also are planning to biochemically characterize the CH···O hydrogen bonds in this enzyme using a similar strategy to that used in Chapter 3. TylM1 contains two tyrosine residues that appear to be forming CH···O hydrogen bonds with the AdoMet methyl group based on the chemical shift experiments and calculations described in Appendix A.1. These two tyrosines are prime candidates for substitution by both phenylalanine and pAF to examine the effects of cofactor binding and catalysis by removing CH···O hydrogen bonds. Our working hypothesis is formulated from the data described in Chapter 3, in that the CH···O hydrogen bonds will be vital for cofactor binding and stabilizing the transition state. One advantage of working with TylM1 is that we may be able to also examine the effects of removing both tyrosine CH···O hydrogen bonds simultaneously using combinations of pAF and phenylalanine substitutions to replace each of the tyrosine residues. In this manner, we may be able to achieve a more in depth view of the functions of individual CH···O hydrogen bonds both separately and in concert. Thus, the investigations of TylM1 serve not only to validate the possibilities raised by the SET7/9 experiments, but also to gain an in-depth understanding of the functions of CH···O hydrogen bonds in methylation.

As discussed in Chapter 1, methyltransferases are involved in many disease states, and are now being targeted to treat several disorders. Multiple strategies that utilize CH···O hydrogen bonding in inhibitor design have very recently been discovered. For example, Tolcapone, an inhibitor of the Class I methyltransferase Catechol-O-methyltransferase used to treat Parkinson's disease, was recently shown to feature a hydrogen bond acceptor pointing towards the AdoMet methyl group when in complex with the enzyme, likely enhancing its binding affinity by forming a CH···O hydrogen bond (*223*). Using a different approach, Luo et al. introduced a new class of

transition-state analogue inhibitors based on the Trievel group's work on CH···O hydrogen bonding in the Class V methyltransferases. These inhibitors are derivatives of the natural product inhibitor sinefungin, in which the methyl sulfonium cation is substituted by an amine-methylene group that forms conventional NH···O hydrogen bonds in place of the methyl CH···O hydrogen bonds, as described in Chapter 2. Luo et al. designed a strategy to synthesize selective transition-state analogue inhibitors by derivitizing sinefungin with various alkyl groups, and demonstrated the strategy's feasibility using the Class V methyltransferase SETD2 (*224*). In the future, the Luo laboratory plans to use this strategy to develop specific inhibitors for many Class V enzymes. In a similar example, the Class I enzyme PRMT5 was shown to be powerfully inhibited by a sinefungin derivative that hydrogen bonds to the methyltransferase's putative methyl CH···O hydrogen bond acceptors. In summary, inhibitors that mimic either the AdoMet methyl CH···O hydrogen bond donors or acceptors are a promising avenue for future drug development that necessitates understanding these interactions within the context of methyltransferase structure and mechanism.

As examined in Appendix A.3, methyl CH···O hydrogen bonding is not limited to just polarized methyl groups such as AdoMet and methyllysine, but is surprisingly prevalent in nonpolar methyl groups in proteins. The question still remains: how important are these interactions in non-polar methyl groups to protein structure, folding, and dynamics? We have begun molecular dynamics simulations in collaboration with the Brooks group to answer this question by removing methyl CH···O hydrogen bonds from proteins by deleting the charge on methyl hydrogen atoms. Preliminary analysis of these simulations suggests that methyl CH···O hydrogen bonds are important to protein folding and dynamics by tempering the hydrophobic effect. As a result, the protein is able to tolerate methyl groups that are exposed to

water, and avoid kinetic traps in the folding pathway. This model corroborates the finding discussed in Chapter 1 that large sections of the surface of lysozyme interact with water solely through CH···O hydrogen bonds (*110*). We believe that this technique could be used to analyze the importance of hydrogen bonds in many benchmark biological systems, including DNA and RNA, and could provide crucial information on the importance and functions of CH···O hydrogen bonds in many contexts.

Although there are still many questions involving the functions of CH···O hydrogen bonding in proteins, as discussed in Chapter 1.2.3, there have been few studies on the importance of CH···O hydrogen bonding in nucleic acids. Analogously to the use of site-specific unnatural amino acid mutations in proteins as described here, it may be possible to evaluate the effects of CH···O hydrogen bonding in nucleic acids by utilizing modified nucleotides. For example, the combined use of 8-aza-2'deoxyadenosine(*225*) and 2-aza-2'deoxyadenosine (*226*) could remove the putative CH···O hydrogen bonds from Hoogsteen and Watson-Crick base-pairing, respectively. Quantum chemistry calculations predict that these modified bases would impact base-pair stability, and previous study indicates that the 2-aza base destabilizes duplex formation (*227*). These bases could also be used to analyze the importance of CH···O hydrogen bonding in RNA folding, and in protein-nucleic acid recognition.

Another important outcome of the work described in this dissertation is the facile and inexpensive means to create specifically labeled AdoMet by enzymatic synthesis. Through the use of methyltransferases, methyl-labeling could be employed to site-specifically label many different biological systems. This property could be especially useful for NMR spectroscopy, as a means to examine the structure and dynamics, or as an assignment strategy, for especially large biomolecules. For example, rRNA

methyltransferases could be used to site-label the ribosome with ^{13}C methyl groups, and these methyl groups could then be used to probe ribosome dynamics. Similarly, methyl groups placed at the ends of very long DNA molecules could be used to examine the effects of length on DNA dynamics, as currently NMR spectroscopy is limited to short nucleotide sequences. Histone or DNA methyltransferases could be employed to label nucleosomes, or nucleosomal arrays, and could then probe the relationship of nucleosome dynamics and genetic regulation. These are only a few of the many possible systems that current NMR techniques could examine in great detail through the use of site-specific methyl labeling.

Similarly, the genetic incorporation of unnatural amino acids into proteins is another new technique that holds great promise for many scientific fields. As demonstrated here, unnatural amino acids can be a useful tool for enzymologists, as it expands the biochemical toolkit beyond the standard 20 amino acids. Given the ever-increasing number of unnatural amino acids that can be genetically incorporated, and the recent opening of the first center for unnatural amino acids research at Oregon State University, the number, diversity, and availability of unnatural amino acids to the scientific community should rapidly increase over the next few years. Engineering enzymes using unnatural amino acids could produce new substrate specificity, or even new chemical reactions previously not achievable through standard mutagenesis (*228*). Soon, it may be possible to engineer nearly any functional moiety desired into enzymes both to study the fundamentals of enzyme mechanism, and to produce novel enzymes for industrial purposes.

We are only beginning to appreciate the breadth of CH···O hydrogen bonding in protein structure and function, and many fundamental questions concerning these

interactions in biological molecules and processes remain unresolved, representing fruitful avenues for future research. It is important to note that CH···O bonds represent a large category of underappreciated interactions in biomolecular structure, including CH···N, CH···S (*136*) and π hydrogen bonds as well as n··· π^* interactions that are typified in proline residues (*229*). Similar to CH···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.

Appendix A

A.1 Direct Evidence of CH···O Hydrogen Bonding to AdoMet in Methyltransferase TylM1

As discussed in Chapter 4, the experimental results presented in this dissertation are not necessarily applicable to all methyltransferases, and may only reflect SET domain enzymes. As a result, we have begun to repeat our analyses of SET7/9 on methyltransferases from other structural classes. As Class I is the largest and most studied class of methyltransferases, we chose to analyze a Class I enzyme for our second case study. TylM1 is a nitrogen dimethyltransferase responsible for completing the synthetic pathway of D-mycaminose, a natural product from *Streptomyces fradiae* (*230*), and the high resolution crystal structure of TylM1 bound to AdoMet revealed several potential CH···O hydrogen bond interactions (*230*).

Using the protocol for NMR experiments and chemical shift calculations outlined in Chapter 2.2, with the expression and purification scheme as previously published (*230*), we sought to verify the position of the hydrogen atoms, and determine the extent of CH···O hydrogen bonding in the active site. The calculated chemical shift of the methyl group, from the optimized structure shown in Figure A.1, was 3.3 ppm. From the NMR experiments, using saturating concentrations of TylM1 and dTDP inhibitor, the chemical shift was measured to be 3.2 ppm.

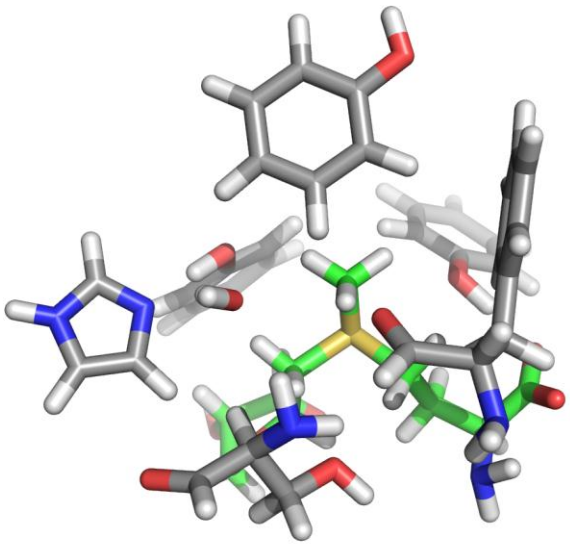

Figure A.1: QM optimization of AdoMet methyl group bound to TylM1

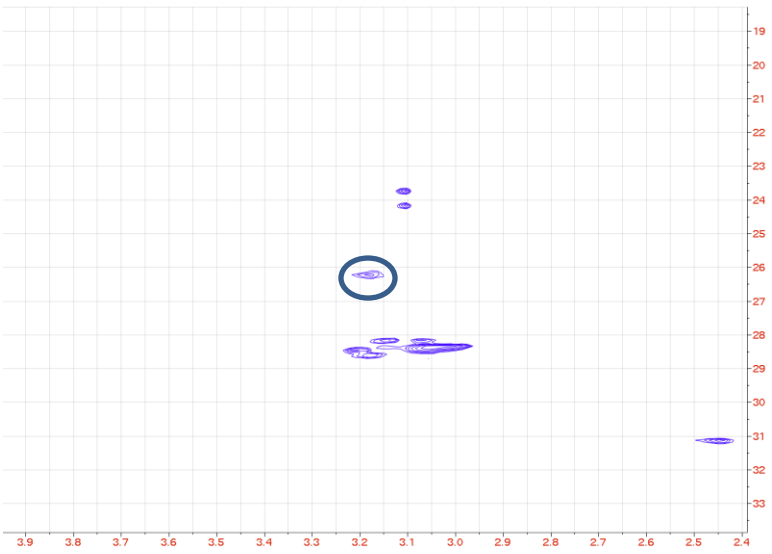

Figure A.2: NMR spectrum of AdoMet bound to TylM1.
The circle indicates the assigned bound AdoMet methyl group, as it is the only resonance not to appear in the control spectra.

Like SET7/9, the calculated and measured chemical shifts of the AdoMet methyl group were in agreement to 0.1 ppm difference. Also like SET7/9, the optimized structure demonstrated that the AdoMet methyl group positions itself to optimize CH···O hydrogen bonding within the active site. The primary difference between the

two enzymes is the magnitude change of the chemical shift from the unbound (3.0 ppm) to the bound state (3.2 for TylM1 and 3.7 for SET7/9). There are two contributing reasons for this difference. First, the conformation of the AdoMet is different in Class I and Class V enzymes, both of which are different from the solution state conformation. Each of these three conformations was calculated to have a distinct basal chemical shift, with that of the Class I methyltransferases being the furthest upfield of the three. Secondly, all of the hydrogen bonding interactions in TylM1 are substantially longer than the shortest interaction in SET7/9, which dominates its chemical shift change. These differences may have consequences in terms of hydrogen bond strength, as well as catalysis, which will be explored in future studies.

A.2 All AdoMet CH···O Hydrogen Bonds

PDB	Acceptor Residue	Acceptor Atom	Acceptor Residue Number	C···O distance (Å)
1EIZ	ASP	O	124	3.38
1EIZ	HOH	O	421	3.59
1EJ0	HOH	O	421	3.49
1EJ0	ASP	O	124	3.56
1G60	HOH	O	561/570	3.00
1G60	HOH	O	583	3.00
1JG4	THR	O	73	2.81
1MSK	SAM	O4'	1301	3.02
1MSK	HOH	O	2174	3.50
1N2X	ASP	OD1	103	2.43
1N2X	ASP	O	103	3.16

1N6A	HSD	O	293	3.00
1N6A	TYR	OH	335	3.48
1RJD	HOH	O	905/845	3.49
1RJD	CYS	O	202	3.54
1V2X	GLU	O	124	2.59
1V2X	HOH	O	442	2.96
1ZQ9	HOH	O	4166	3.09
2AVD	SER	O	84	3.27
2AVD	ASP	OD1	185	3.51
2B9E	ASP	O	305	3.44
2CL5	BIE	O3	1218	3.24
2CL5	ASP	O	141	3.37
2CL5	MET	O	40	3.38
2CL5	HOH	O	2205/2186	3.67
2DPM	ASP	O	194	3.25
2NPN	TYR	O	115	3.08
2NPN	HOH	O	4781	3.08
2NXE	HOH	O	336	2.82
2NYU	ASP	O	104	3.42
2PLW	HOH	O	216	3.22
2WA2	HOH	O	2149	3.46
2ZUL	ASN	O	305	3.15

3AIA	HSD	O	158	3.06
3AIA	ASP	O	157	3.21
3CKK	PHE	O	131	2.90
3DLC	GLY	O	118	3.29
3DMF	ASN	O	305	3.21
3DOU	HOH	O	255	3.61
3DOU	ASP	O	111	3.68
3DXY	PHE	O	142	2.92
3ELW	ASP	O	146	3.40
3G89	ALA	O	159	3.12
3G89	HOH	O	425	3.22
3G89	HOH	O	948	3.25
3G89	HOH	O	955	3.38
3G89	HOH	O	686	3.47
3GCZ	ASP	O	146	3.17
3IHT	GLY	O	114	3.53
3KKZ	GLN	O	26	2.83
3KKZ	HOH	O	777/701	3.28
3M6W	HOH	O	985	3.32
3M6W	HOH	O	984	3.41
3MB5	HOH	O	267	3.28
3MB5	ASP	OD2	169	3.40

3MB5	GLN	O	75	3.53
3MTE	PHE	O	105	3.02
3NDI	HOH	O	481	3.56
3NDI	HOH	O	482	3.56
3NDI	HOH	O	480	3.66
3OOI	HOH	O	247	3.11
3OOI	TYR	OH	207	3.49
3P97	HOH	O	343	2.30
3P97	HOH	O	635	2.94
3P97	HOH	O	462	3.34
3P97	ASP	O	146	3.40
3PFG	HOH	O	485	3.39
3PFG	PHE	O	118	3.49
3PFG	TYR	OH	14	3.49
3QWP	SER	O	202	3.08
3QWP	ASN	OD1	181	3.25
3QWP	TYR	OH	239	3.34
3RFA	GLY	O	177	3.26
3RQ4	PHE	O	160	3.22
3RQ4	HOH	O	267	3.34
3RQ4	TYR	OH	217	3.34
3S8P	PHE	O	250	3.10

3S8P	ALA	O	269	3.18
3S8P	TYR	OH	307	3.40
3UJ7	ASP	O	128	2.85

A.3 Perdeuterated Diffraction of SET7/9 Crystals

As discussed in Chapter 1.2, it is currently difficult to examine the existence and geometrical properties of CH···O hydrogen bonds in biomolecules, and one technique that holds great promise is neutron crystallography. To achieve sufficient signal-to-noise ratios to determine high resolution structures often requires large perdeuterated crystals. As such, we prepared perdeuterated SET7/9 as described in Chapter 3.2.2, and crystallized it using similar conditions to those described in Chapter 3.2.3. A perdeuterated crystal, and diffraction to better than 1.6 Å resolution, are shown in Figure A.3. With the upcoming commissioning of the MaNDi beamline at Oak Ridge National Laboratories, these crystals may be sufficient to determine high-resolution neutron structures of SET7/9 to study the active site CH···O hydrogen bonding in great depth.

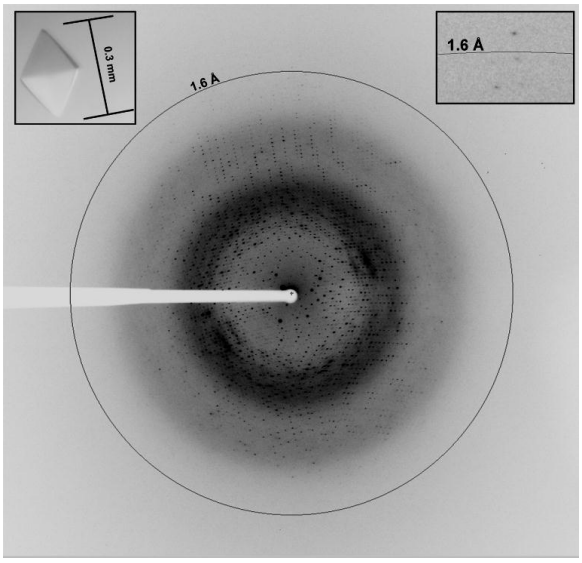

Figure A.3: Preliminary diffraction analysis of crystals of deuterated SET7/9 bound to AdoHcy and a CD3-monomethyl-TAF10K189 peptide.
The circle denotes the 1.6 Å diffraction boundary. The left inset panel depicts a perdeuterated SET7/9 crystal. The right inset panel illustrates reflections beyond the 1.6 Å diffraction ring.

A.4 Side Chain Methyl CH···O Hydrogen Bonds

The previous chapters have demonstrated the ubiquitous nature and importance of CH···O hydrogen bonding to the AdoMet methyl group in methyltransferases. One of the more intriguing results to come from these studies is that although side chain methyl groups form CH···O hydrogen bonds less preferentially than the AdoMet methyl group, they still form at a surprisingly high rate. In total, 28% of all methyl groups in proteins were considered to form CH···O hydrogen bonds based on the criteria used in our methyltransferase survey in Chapter 3. Further, as was shown in Figure 3.3, a substantial percentage of all side chain methyl CH···O contacts are closer than 3.7 Å. These findings prompted us to more closely examine the extent of side chain methyl CH···O hydrogen bonding in proteins.

We analyzed methyl CH···O hydrogen bonding in all neutron structures in which methyl hydrogen atoms should be discernible (*231*), to directly evaluate the level of CH···O hydrogen bonding in proteins. Angular and distance cutoffs were similar to that used in Chapter 3. Unexpectedly, 36% of all CH···O contacts in proteins fell within the angular and distance criteria of being hydrogen bonds. For comparison, the analogous percentage of all CH···O bonds, consisting primarily of backbone CH···O bonds in β-sheets, was discovered by Derewenda et al. to be substantially less than that of methyl groups, at 13% (*75*). This juxtaposition is surprising for multiple reasons: 1) Unlike methyl groups, the backbone CH···O hydrogen bonds are predisposed to form by secondary structure. 2) Methyl groups typically form weaker CH···O hydrogen bonds than the Cα group, due to the unique surrounding sp2 character of the backbone. 3) Most methyl groups are thought to reside in the hydrophobic core of the protein, sequestered from hydrophilic oxygen atoms. By comparing the distribution of CH···O contacts to CH···C van der Waals contacts as a control, the CH···O distribution clearly favors shorter interactions than the van der Waals contact, similar to other CH···O hydrogen bonds (Figure A.4) (*75*).

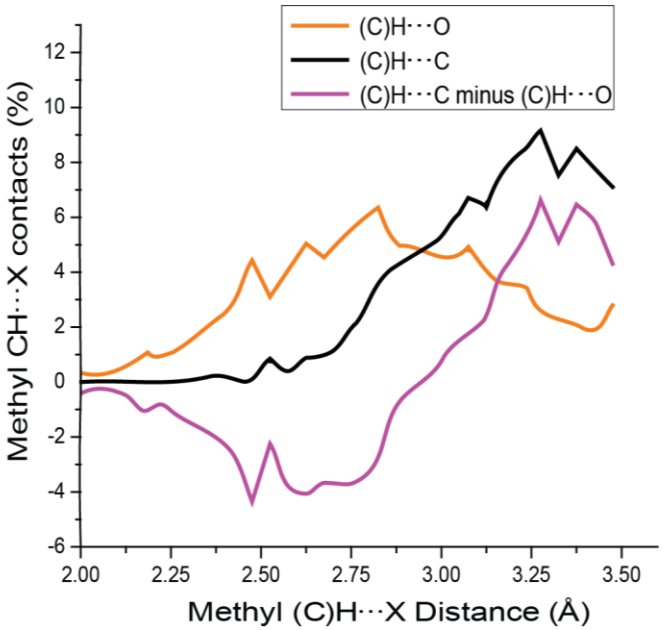

Figure A.4: Methyl (C)H···O contact distances in proteins.
Magenta line is the difference between the CH···C (black) and CH···O distributions (orange).

Importantly, the angular distribution of the CH···O interactions is also consistent with hydrogen bond formation. The C-H···O angle distribution displays a greater tendency towards linearity than that of C-H···C angles. Similarly, the CH···O elevation angle displays a strong trend towards co-planarity. Combined with the distance distribution shown in Figure A.4, the angular distributions clearly demonstrate that methyl groups in proteins are able to form CH···O hydrogen bonds (*57*).

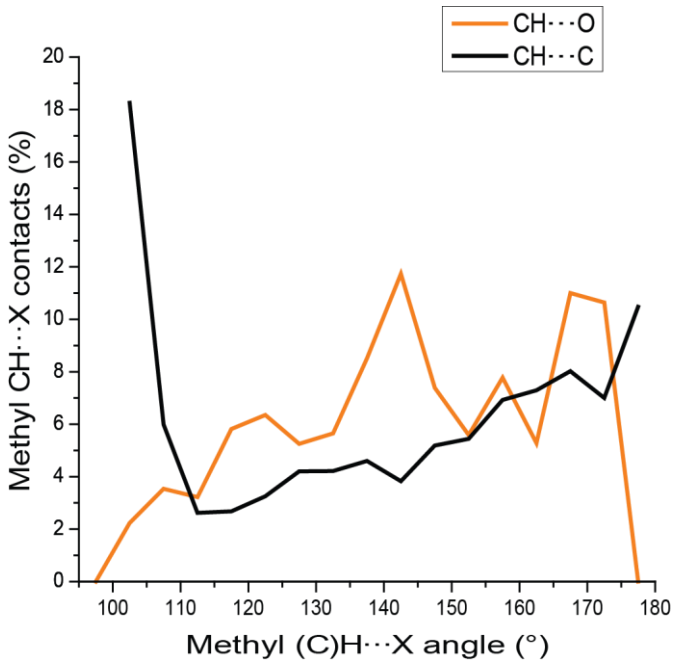

Figure A.5: Methyl C-H···O (orange) and C-H···C (black) angles.

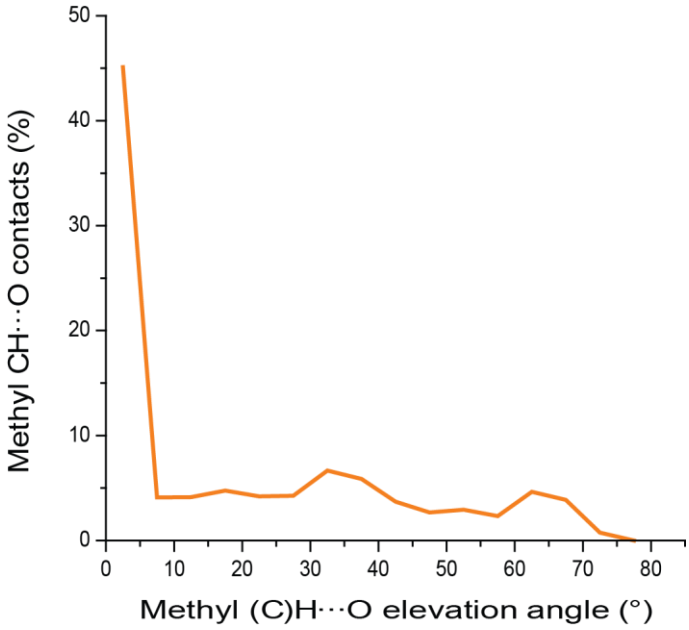

Figure A.6 Methyl CH···O elevation angles

References

1. Cantoni, G. L. (1975) Biological Methylation - Selected Aspects, *Annu Rev Biochem 44*, 435-451.
2. Schubert, H. L., Blumenthal, R. M., and Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence, *Trends Biochem Sci 28*, 329-335.
3. Blanco, S., Kurowski, A., Nichols, J., Watt, F. M., Benitah, S. A., and Frye, M. (2011) The RNA-Methyltransferase Misu (NSun2) Poises Epidermal Stem Cells to Differentiate, *Plos Genet 7*.
4. Cheng, X. D., and Blumenthal, R. M. (2010) Coordinated Chromatin Control: Structural and Functional Linkage of DNA and Histone Methylation, *Biochemistry 49*, 2999-3008.
5. Albert, M., and Helin, K. (2010) Histone methyltransferases in cancer, *Semin Cell Dev Biol 21*, 209-220.
6. Greer, E. L., and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance, *Nat Rev Genet 13*, 343-357.
7. Copeland, R. A., Solomon, M. E., and Richon, V. M. (2009) Protein methyltransferases as a target class for drug discovery, *Nat Rev Drug Discov 8*, 724-732.
8. Cha, B., and Jho, E. H. (2012) Protein arginine methyltransferases (PRMTs) as therapeutic targets, *Expert Opin Ther Tar 16*, 651-664.
9. Poewe, W. (2004) The role of COMT inhibition in the treatment of Parkinson's disease, *Neurology 62*, S31-S38.
10. Mannisto, P. T., and Kaakkola, S. (1999) Catechol-O-methyltransferase (COMT): Biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors, *Pharmacol Rev 51*, 593-628.
11. Serretti, A., and Olgiati, P. (2012) Catechol-O-Methyltransferase and Alzheimer's Disease: A Review of Biological and Genetic Findings, *Cns Neurol Disord-Dr 11*, 299-305.

12. Borroni, B., Agosti, C., Archetti, S., Costanzi, C., Bonomi, S., Ghianda, D., Lenzi, G. L., Caimi, L., Di Luca, M., and Padovani, A. (2004) Catechol-O-methyltransferase gene polymorphism is associated with risk of psychosis in Alzheimer Disease, *Neurosci Lett 370*, 127-129.
13. Pagans, S., Kauder, S. E., Kaehlcke, K., Sakane, N., Schroeder, S., Dormeyer, W., Trievel, R. C., Verdin, E., Schnolzer, M., and Ott, M. (2010) The Cellular Lysine Methyltransferase Set7/9-KMT7 Binds HIV-1 TAR RNA, Monomethylates the Viral Transactivator Tat, and Enhances HIV Transcription, *Cell Host Microbe 7*, 234-244.
14. Grosjean, H. (2009) *DNA and RNA modification enzymes : structure, mechanism, function, and evolution*, Landes Bioscience, Austin, Tex.
15. Kojic, M., Topisirovic, L., and Vasiljevic, B. (1992) Cloning and Characterization of an Aminoglycoside Resistance Determinant from Micromonospora-Zionensis, *J Bacteriol 174*, 7868-7872.
16. Cubrilo, S., Babic, F., Douthwaite, S., and Vlahovicek, G. M. (2009) The aminoglycoside resistance methyltransferase Sgm impedes RsmF methylation at an adjacent rRNA nucleotide in the ribosomal A site, *Rna 15*, 1492-1497.
17. Coward, J. K., and Sweet, W. D. (1971) Kinetics and Mechanism of Methyl Transfer from Sulfonium Compounds to Various Nucleophiles, *J Org Chem 36*, 2337-&.
18. Coward, J. K. (1977) Chemical Mechanisms of Methyl Transfer Reactions: Comparison of Methylases with Nonenzymic "Model Reactions", In *The Biochemistry of adenosylmethionine : [proceedings of an international symposium on the biochemistry of adenosylmethionine, sponsored by the Accademia nazionale dei Lincei, held in Rome, Italy May 21-26, 1974]* (Salvatore, F., and Accademia nazionale dei Lincei., Eds.), pp xv, 588 p., Columbia University Press, New York.
19. Gray, C. H., Coward, J. K., Schowen, K. B., and Schowen, R. L. (1979) Alpha-Deuterium and C-13 Isotope Effects for a Simple, Inter-Molecular Sulfur-to-Oxygen Methyl-Transfer Reaction - Transition-State Structures and Isotope Effects in Transmethylation and Transalkylation, *J Am Chem Soc 101*, 4351-4358.
20. Hegazi, M. F., Borchardt, R. T., and Schowen, R. L. (1979) Alpha-Deuterium and C-13 Isotope Effects for Methyl Transfer Catalyzed by Catechol O-Methyltransferase - Sn2-Like Transition-State, *J Am Chem Soc 101*, 4359-4365.
21. Iwig, D. F., Grippe, A. T., McIntyre, T. A., and Booker, S. J. (2004) Isotope and elemental effects indicate a rate-limiting methyl transfer as the initial step in the reaction catalyzed by Escherichia coli cyclopropane fatty acid synthase, *Biochemistry 43*, 13510-13524.

22. Mihel, I., Knipe, J. O., Coward, J. K., and Schowen, R. L. (1979) Alpha-Deuterium Isotope Effects and Transition-State Structure in an Intra-Molecular Model System for Methyl-Transfer Enzymes, *J Am Chem Soc 101*, 4349-4351.
23. Rodgers, J., Femec, D. A., and Schowen, R. L. (1982) Isotopic Mapping of Transition-State Structural Features Associated with Enzymic Catalysis of Methyl Transfer, *J Am Chem Soc 104*, 3263-3268.
24. Williams, I. H. (1984) Theoretical Modeling of Compression Effects in Enzymic Methyl Transfer, *J Am Chem Soc 106*, 7206-7212.
25. Boyd, R. J., Kim, C. K., Shi, Z., Weinberg, N., and Wolfe, S. (1993) Secondary H/D Isotope Effects and Transition-State Looseness in Nonidentity Methyl Transfer-Reactions - Implications for the Concept of Enzymatic Catalysis Via Transition-State Compression, *J Am Chem Soc 115*, 10147-10152.
26. Soriano, A., Castillo, R., Christov, C., Andres, J., Moliner, V., and Tunon, I. (2006) Catalysis in glycine N-methyltransferase: Testing the electrostatic stabilization and compression hypothesis, *Biochemistry 45*, 14917-14925.
27. Ruggiero, G. D., Williams, I. H., Roca, M., Moliner, V., and Tunon, I. (2004) QM/MM determination of kinetic isotope effects for COMT-catalyzed methyl transfer does not support compression hypothesis, *J Am Chem Soc 126*, 8634-8635.
28. Zhang, J. Y., and Klinman, J. P. (2011) Enzymatic Methyl Transfer: Role of an Active Site Residue in Generating Active Site Compaction That Correlates with Catalytic Efficiency, *J Am Chem Soc 133*, 17134-17137.
29. Zheng, Y. J., and Bruice, T. C. (1997) A theoretical examination of the factors controlling the catalytic efficiency of a transmethylation enzyme: Catechol O-methyltransferase, *J Am Chem Soc 119*, 8137-8145.
30. Lau, E. Y., and Bruice, T. C. (1998) Importance of correlated motions in forming highly reactive near attack conformations in catechol O-methyltransferase, *J Am Chem Soc 120*, 12387-12394.
31. Olsson, M. H. M., and Warshel, A. (2004) Solute solvent dynamics and energetics in enzyme catalysis: The S(N)2 reaction of dehalogenase as a general benchmark, *J Am Chem Soc 126*, 15167-15179.
32. Kanaan, N., Pernia, J. J. R., and Williams, I. H. (2008) QM/MM simulations for methyl transfer in solution and catalysed by COMT: ensemble-averaging of kinetic isotope effects, *Chem Commun*, 6114-6116.
33. Cheng, X. D., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. (1993) Crystal-Structure of the Hhal DNA Methyltransferase Complexed with S-Adenosyl-L-Methionine, *Cell 74*, 299-307.

34. Vidgren, J., Svensson, L. A., and Liljas, A. (1994) Crystal-Structure of Catechol O-Methyltransferase, *Nature 368*, 354-358.
35. Martin, J. L., and McMillan, F. M. (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold, *Curr Opin Struc Biol 12*, 783-793.
36. Dixon, M. M., Huang, S., Matthews, R. G., and Ludwig, M. (1996) The structure of the C-terminal domain of methionine synthase: Presenting S-adenosylmethionine for reductive methylation of B-12, *Structure 4*, 1263-1275.
37. Petrossian, T. C., and Clarke, S. G. (2011) Uncovering the Human Methyltransferasome, *Mol Cell Proteomics 10*.
38. Schubert, H. L., Blumenthal, R. M., and Cheng, X. D. (2003) Many paths to methyltransfer: a chronicle of convergence, *Trends Biochem Sci 28*, 329-335.
39. Grove, T. L., Radle, M. I., Krebs, C., and Booker, S. J. (2011) Cfr and RlmN Contain a Single [4Fe-4S] Cluster, which Directs Two Distinct Reactivities for S-Adenosylmethionine: Methyl Transfer by S(N)2 Displacement and Radical Generation, *J Am Chem Soc 133*, 19586-19589.
40. Guo, H. B., and Guo, H. (2007) Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity, *P Natl Acad Sci USA 104*, 8797-8802.
41. Zhang, X. D., and Bruice, T. C. (2008) Product specificity and mechanism of protein lysine methyltransferases: Insights from the histone lysine methyltransferase SET8, *Biochemistry 47*, 6671-6677.
42. Krishnan, S., Horowitz, S., and Trievel, R. C. (2011) Structure and Function of Histone H3 Lysine 9 Methyltransferases and Demethylases, *Chembiochem 12*, 254-263.
43. Kim, J. K., Samaranayake, M., and Pradhan, S. (2009) Epigenetic mechanisms in mammals, *Cell Mol Life Sci 66*, 596-612.
44. Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. D. (2005) The SET-domain protein superfamily: protein lysine methyltransferases, *Genome Biol 6*, 227.
45. Taylor, W. R., Xiao, B., Gamblin, S. J., and Lin, K. (2003) A knot or not a knot? SETting the record 'straight' on proteins, *Comput Biol Chem 27*, 11-15.
46. Trievel, R. C., Beach, B. M., Dirk, L. M., Houtz, R. L., and Hurley, J. H. (2002) Structure and catalytic mechanism of a SET domain protein methyltransferase, *Cell 111*, 91-103.
47. Kwon, T., Chang, J. H., Kwak, E., Lee, C. W., Joachimiak, A., Kim, Y. C., Lee, J. W., and Cho, Y. J. (2003) Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet, *Embo J 22*, 292-303.

48. Xiao, B., Jing, C., Wilson, J. R., Walker, P. A., Vasisht, N., Kelly, G., Howell, S., Taylor, I. A., Blackburn, G. M., and Gamblin, S. J. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9, *Nature 421*, 652-656.
49. 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.
50. Subramanian, K., Jia, D., Kapoor-Vazirani, P., Powell, D. R., Collins, R. E., Sharma, D., Peng, J. M., Cheng, X. D., and Vertino, P. M. (2008) Regulation of estrogen receptor alpha by the SET7 lysine methyltransferase, *Mol Cell 30*, 336-347.
51. Jacobs, S. A., Harp, J. M., Devarakonda, S., Kim, Y., Rastinejad, F., and Khorasanizadeh, S. (2002) The active site of the SET domain is constructed on a knot, *Nat Struct Biol 9*, 833-838.
52. Xiao, B., Jing, C., Kelly, G., Walker, P. A., Muskett, F. W., Frenkiel, T. A., Martin, S. R., Sarma, K., Reinberg, D., Gamblin, S. J., and Wilson, J. R. (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7, *Gene Dev 19*, 1444-1454.
53. Hu, P., and Zhang, Y. K. (2006) Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: An ab initio QM/MM-FE study with multiple initial structures, *J Am Chem Soc 128*, 1272-1278.
54. Zhang, X., and Bruice, T. C. (2007) Histone lysine methyltransferase SET7/9: formation of a water channel precedes each methyl transfer, *Biochemistry 46*, 14838-14844.
55. Hu, P., Wang, S., and Zhang, Y. (2008) How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by ab initio QM/MM molecular dynamics simulations, *J Am Chem Soc 130*, 3806-3813.
56. Trievel, R. C., Flynn, E. M., Houtz, R. L., and Hurley, J. H. (2003) Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT, *Nat Struct Biol 10*, 545-552.
57. Steiner, T., and Desiraju, G. R. (1998) Distinction between the weak hydrogen bond and the van der Waals interaction, *Chem Commun*, 891-892.
58. Steiner, T. (2002) The hydrogen bond in the solid state, *Angew Chem Int Edit 41*, 48-76.
59. Scheiner, S. (2011) Weak H-bonds. Comparisons of CH center dot center dot center dot O to NH center dot center dot center dot O in proteins and PH center dot center dot center dot N to direct P center dot center dot center dot N interactions, *Phys Chem Chem Phys 13*, 13860-13872.

60. Cannizzaro, C. E., and Houk, K. N. (2002) Magnitudes and chemical consequences of (RN+)-N-3-C-H center dot center dot center dot O=C hydrogen bonding, *J Am Chem Soc 124*, 7163-7169.
61. Gilli, G., and Gilli, P. (2009) *The Nature of the Hydrogen Bond*, Oxford University Press, New York.
62. Desiraju, G. R. (2011) A bond by any other name, *Angew Chem Int Ed Engl 50*, 52-59.
63. Desiraju, G. R., and Steiner, T. (1999) *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, New York.
64. Arunan, E., Desiraju, G. R., Klein, R. A., Sadlej, J., Scheiner, S., Alkorta, I., Clary, D. C., Crabtree, R. H., Dannenberg, J. J., and Hobza, P. (2011) Definition of the hydrogen bond (IUPAC Recommendations 2011), *Pure and Applied Chemistry 83*, 1637.
65. Scheiner, S., Gu, Y., and Kar, T. (2000) Evaluation of the H-bonding properties of CH center dot center dot center dot O interactions based upon NMR spectra, *J Mol Struc-Theochem 500*, 441-452.
66. Krimm, S., and Kuroiwa, K. (1968) Low Temperature Infrared Spectra of Polyglycines and C-H...O=C Hydrogen Bonding in Polyglycine 2, *Biopolymers 6*, 401-&.
67. Scheiner, S. (2009) Identification of Spectroscopic Patterns of CH center dot center dot center dot O H-Bonds in Proteins, *J Phys Chem B 113*, 10421-10427.
68. Gu, Y. L., Kar, T., and Scheiner, S. (1999) Fundamental properties of the CH center dot center dot center dot O interaction: Is it a true hydrogen bond?, *J Am Chem Soc 121*, 9411-9422.
69. Desiraju, G. R. (1996) The C-H center dot center dot center dot O hydrogen bond: Structural implications and supramolecular design, *Accounts Chem Res 29*, 441-449.
70. Kumler, W. D. (1935) The effect of the hydrogen bond on the dielectric constants and boiling points of organic liquids, *J Am Chem Soc 57*, 600-605.
71. Glasstone, S. (1937) The structure of some molecular complexes in the liquid phase., *T Faraday Soc 33*, 0200-0207.
72. Ramachan.Gn, and Chandras.R. (1968) Interchain Hydrogen Bonds Via Bound Water Molecules in Collagen Triple Helix, *Biopolymers 6*, 1649-&.
73. Ramachan.Gn, Sasisekh.V, and Ramakris.C. (1966) Molecular Structure of Polyglycine 2, *Biochim Biophys Acta 112*, 168-&.
74. Krimm, S. (1967) Hydrogen Bonding of C-H...O=C in Proteins, *Science 158*, 530-&.

75. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) The Occurrence of C-H-Center-Dot-Center-Dot-Center-Dot-O Hydrogen-Bonds in Proteins, *J Mol Biol 252*, 248-262.
76. Taylor, R., and Kennard, O. (1982) Crystallographic Evidence for the Existence of C-H...O, C-H...N, and C-H...C1 Hydrogen-Bonds, *J Am Chem Soc 104*, 5063-5070.
77. Wahl, M. C., and Sundaralingam, M. (1997) C-H center dot center dot center dot O hydrogen bonding in biology, *Trends Biochem Sci 22*, 97-102.
78. Schwalbe, C. H. (2012) June Sutor and the C–H··· O hydrogen bonding controversy, *Crystallography Reviews 18*, 189-204.
79. Arbely, E., and Arkin, I. T. (2004) Experimental measurement of the strength of a C alpha-H center dot center dot center dot O bond in a lipid bilayer, *J Am Chem Soc 126*, 5362-5363.
80. 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 sterochemical analysis, *J Mol Biol 297*, 713-732.
81. Sandalova, T., Schneider, G., Kack, H., and Lindqvist, Y. (1999) Structure of dethiobiotin synthetase at 0.97 angstrom resolution, *Acta Crystallogr D 55*, 610-624.
82. 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.
83. Cordier, F., Barfield, M., and Grzesiek, S. (2003) Direct observation of Calpha-Halpha...O=C hydrogen bonds in proteins by interresidue h3JCalphaC' scalar couplings, *J Am Chem Soc 125*, 15750-15751.
84. Sheppard, D., Li, D. W., Godoy-Ruiz, R., Bruschweiler, R., and Tugarinov, V. (2010) Variation in quadrupole couplings of alpha deuterons in ubiquitin suggests the presence of C(alpha)-H(alpha)...O=C hydrogen bonds, *J Am Chem Soc 132*, 7709-7719.
85. 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 H-1 NMR chemical shifts support (His) C-epsilon 1-H center dot center dot center dot O = C H-bond: Proposal for reaction-driven ring flip mechanism in serine protease catalysis, *P Natl Acad Sci USA 97*, 10371-10376.
86. 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.

87. Blakeley, M. P., Langan, P., Niimura, N., and Podjarny, A. (2008) Neutron crystallography: opportunities, challenges, and limitations, *Curr Opin Struc Biol 18*, 593-600.
88. 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, *P Natl Acad Sci USA 107*, 6817-6822.
89. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The C alpha-H center dot center dot center dot O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions, *P Natl Acad Sci USA 98*, 9056-9061.
90. Shefter, E., Barlow, M., Sparks, R., and Trueblood, K. (1964) Crystal + Molecular Structure of Beta-Adenosine-2]-Beta-Uridine-5]-Phosphoric Acid, *J Am Chem Soc 86*, 1873-&.
91. 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-&.
92. 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.
93. Brandl, M., Lindauer, K., Meyer, M., and Suhnel, J. (1999) C-H...O and C-H...N interactions in RNA structures, *Theor Chem Acc 101*, 103-113.
94. 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.
95. 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.
96. 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 center dot center dot center dot O hydrogen bonds, *Chem-Eur J 3*, 1400-1404.
97. 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.

98. Deng, J. P., Xiong, Y., and Sundaralingam, M. (2001) X-ray analysis of an RNA tetraplex (UGGGGU)(4) with divalent Sr2+ ions at subatomic resolution (0.61 angstrom), *P Natl Acad Sci USA 98*, 13665-13670.
99. 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.
100. Castellano, R. K., Gramlich, V., and Diederich, F. (2002) Rebek imides and their adenine complexes: Preferences for Hoogsteen binding in the solid state and in solution, *Chem-Eur J 8*, 118-129.
101. Quinn, J. R., Zimmerman, S. C., Del Bene, J. E., and Shavitt, I. (2007) Does the A center dot T or G center dot C base-pair possess enhanced stability? Quantifying the effects of CH center dot center dot center dot O interactions and secondary interactions on base-pair stability using a phenomenological analysis and ab initio calculations, *J Am Chem Soc 129*, 934-941.
102. 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.
103. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1994) Crystal-Structure of Intercalated 4-Stranded D(C3t) at 1.4 Angstrom Resolution, *P Natl Acad Sci USA 91*, 11636-11640.
104. 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. T., 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-U741.
105. Yokoyama, T., Mizuguchi, M., Nabeshima, Y., Kusaka, K., Yamada, T., Hosoya, T., Ohhara, T., Kurihara, K., Tomoyori, K., Tanaka, I., and Niimura, N. (2012) Hydrogen-bond network and pH sensitivity in transthyretin: Neutron crystal structure of human transthyretin, *J Struct Biol 177*, 283-290.
106. 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 C-H to O hydrogen bond in an engineered protein cavity, *J Am Chem Soc 119*, 9083-9084.
107. 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.
108. Klaholz, B. P., and Moras, D. (2002) C-H center dot center dot center dot O hydrogen bonds in the nuclear receptor RAR gamma - a potential tool for drug selectivity, *Structure 10*, 1197-1204.

109. 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.
110. Bon, C., Lehmann, M. S., and Wilkinson, C. (1999) Quasi-Laue neu iron-diffraction study of the water arrangement in crystals of triclinic hen egg-white lysozyme, *Acta Crystallogr D 55*, 978-987.
111. Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L., and Zhurkin, V. B. (1998) A role for CH center dot center dot center dot O interactions in protein-DNA recognition, *J Mol Biol 277*, 1129-1140.
112. Castellano, R. K. (2004) Progress toward understanding the nature and function of C-H center dot center dot center dot O interactions, *Curr Org Chem 8*, 845-865.
113. Jiang, L., and Lai, L. H. (2002) CH center dot center dot center dot O hydrogen bonds at protein-protein interfaces, *J Biol Chem 277*, 37732-37740.
114. Lee, K. M., Chang, H. C., Jiang, J. C., Chen, J. C. C., Kao, H. E., Lin, S. H., and Lin, I. J. B. (2003) C-H---Ohydrogen bonds in beta-sheetlike networks: Combined X-ray crystallography and high-pressure infrared study, *J Am Chem Soc 125*, 12358-12364.
115. Yohannan, S., Faham, S., Yang, D., Grosfeld, D., Chamberlain, A. K., and Bowie, J. U. (2004) A C-alpha-H center dot center dot center dot O hydrogen bond in a membrane protein is not stabilizing, *J Am Chem Soc 126*, 2284-2285.
116. 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-U1273.
117. Shi, Z. S., Olson, C. A., Bell, A. J., and Kallenbach, N. R. (2002) Non-classical helix-stabilizing interactions: C-H...O H-bonding between Phe and Glu side chains in alpha-helical peptides, *Biophys Chem 101*, 267-279.
118. Strop, P., and Mayo, S. L. (2000) Contribution of surface salt bridges to protein stability, *Biochemistry 39*, 1251-1255.
119. Gehring, K., Leroy, J. L., and Gueron, M. (1993) A Tetrameric DNA-Structure with Protonated Cytosine.Cytosine Base-Pairs, *Nature 363*, 561-565.
120. Berger, I., Egli, M., and Rich, A. (1996) Inter-strand C-H center dot center dot center dot O hydrogen bonds stabilizing four-stranded intercalated molecules: Stereoelectronic effects of 04' in cytosine-rich DNA, *P Natl Acad Sci USA 93*, 12116-12121.

121. Leroy, J. L., Snoussi, K., and Gueron, M. (2001) Investigation of the energetics of C-H center dot center dot center dot O hydrogen bonds in the DNA i-motif via the equilibrium between alternative intercalation topologies, *Magn Reson Chem 39*, S171-S176.
122. Das, R., Karanicolas, J., and Baker, D. (2010) Atomic accuracy in predicting and designing noncanonical RNA structure, *Nat Methods 7*, 291-294.
123. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) (His)C-Epsilon-H...O=C Hydrogen-Bond in the Active-Sites of Serine Hydrolases, *J Mol Biol 241*, 83-93.
124. 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.
125. Bach, R. D., Thorpe, C., and Dmitrenko, O. (2002) C-H center dot center dot center dot carboxylate oxygen hydrogen bonding in substrate activation by acyl-CoA dehydrogenases: Synergy between the H-bonds, *J Phys Chem B 106*, 4325-4335.
126. Sengoku, T., and Yokoyama, S. (2011) Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A, *Gene Dev 25*, 2266-2277.
127. Horton, J. R., Upadhyay, A. K., Qi, H. H., Zhang, X., Shi, Y., and Cheng, X. D. (2010) Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases, *Nat Struct Mol Biol 17*, 38-U52.
128. 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.
129. 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.
130. Del Rizzo, P. A., Couture, J. F., Dirk, L. M. A., 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.

131. Ramachan.Gn, and Sasisekh.V. (1965) Refinement of Structure of Collagen, *Biochim Biophys Acta 109*, 314-&.
132. Desiraju, G. R. (1991) The C-H...O Hydrogen-Bond in Crystals - What Is It, *Accounts Chem Res 24*, 290-296.
133. Derewenda, Z., Lee, L., Kobos, P., and Derewenda, U. (1995) The Occurrence of C-H-Center-Dot-Center-Dot-Center-Dot-O Hydrogen-Bonds in Proteins, *Faseb J 9*, A1246-A1246.
134. Yates, J. R., Pham, T. N., Pickard, C. J., Mauri, F., Amado, A. M., Gil, A. M., and Brown, S. P. (2005) An investigation of weak CH center dot center dot center dot O hydrogen bonds in maltose anomers by a combination of calculation and experimental solid-state NMR spectroscopy, *J Am Chem Soc 127*, 10216-10220.
135. Peralta, J. E., de Azua, M. C. R., and Contreras, R. H. (1999) Natural bond orbitals analysis of C-H center dot center dot center dot O interactions in NCH/H2O and NCH/OCH2, and their effect on nuclear magnetic shielding constants, *J Mol Struc-Theochem 491*, 23-31.
136. Westler, W. M., Lin, I. J., Perczel, A., Weinhold, F., and Markley, J. L. (2011) Hyperfine-Shifted C-13 Resonance Assignments in an Iron-Sulfur Protein with Quantum Chemical Verification: Aliphatic C-H center dot center dot center dot S 3-Center-4-Electron Interactions, *J Am Chem Soc 133*, 1310-1316.
137. Trievel, R. C., Beach, B. M., Dirk, L. M. A., Houtz, R. L., and Hurley, J. H. (2002) Structure and catalytic mechanism of a SET domain protein methyltransferase, *Cell 111*, 91-103.
138. Zhang, X., Yang, Z., Khan, S. I., Horton, J. R., Tamaru, H., Selker, E. U., and Cheng, X. D. (2003) Structural basis for the product specificity of histone lysine methyltransferases, *Mol Cell 12*, 177-185.
139. Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. D. (2005) The SET-domain protein superfamily: protein lysine methyltransferases, *Genome Biol 6*, -.
140. 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*.
141. Iwig, D. F., and Booker, S. J. (2004) Insight into the polar reactivity of the onium chalcogen analogues of S-adenosyl-L-methionine, *Biochemistry 43*, 13496-13509.
142. Schalk-Hihi, C., and Markham, G. D. (1999) The conformations of a substrate and a product bound to the active site of S-adenosylmethionine synthetase, *Biochemistry 38*, 2542-2550.
143. Chirlian, L. E., and Francl, M. M. (1987) Atomic Charges Derived from Electrostatic Potentials - a Detailed Study, *J Comput Chem 8*, 894-905.

144. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; . (2004) Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT.

145. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: the biomolecular simulation program, *J Comput Chem 30*, 1545-1614.

146. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water, *J Chem Phys 79*, 926-935.

147. Nose, S., and Klein, M. L. (1986) Constant-temperature-constant-pressure molecular-dynamics calculations for molecular solids: Application to solid nitrogen at high pressure, *Phys Rev B Condens Matter 33*, 339-342.

148. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A Smooth Particle Mesh Ewald Method, *J Chem Phys 103*, 8577-8593.

149. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell, A. D., Jr. (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J Comput Chem 31*, 671-690.

150. Markham, G. D., Norrby, P. O., and Bock, C. W. (2002) S-Adenosylmethionine conformations in solution and in protein complexes: Conformational influences of the sulfonium group, *Biochemistry 41*, 7636-7646.
151. Becke, A. D. (1993) Density-Functional Thermochemistry .3. The Role of Exact Exchange, *J Chem Phys 98*, 5648-5652.
152. Ditchfie.R, Hehre, W. J., and Pople, J. A. (1971) Self-Consistent Molecular-Orbital Methods .9. Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules, *J Chem Phys 54*, 724-&.
153. Miertus, S., Scrocco, E., and Tomasi, J. (1981) Electrostatic Interaction of a Solute with a Continuum - a Direct Utilization of Abinitio Molecular Potentials for the Prevision of Solvent Effects, *Chem Phys 55*, 117-129.
154. Wolinski, K., Hinton, J. F., and Pulay, P. (1990) Efficient Implementation of the Gauge-Independent Atomic Orbital Method for Nmr Chemical-Shift Calculations, *J Am Chem Soc 112*, 8251-8260.
155. Scheiner, S. (2010) Effect of CH center dot center dot center dot O Hydrogen Bond Length on the Geometric and Spectroscopic Features of the Peptide Unit of Proteins, *Int J Quantum Chem 110*, 2775-2783.
156. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF chimera - A visualization system for exploratory research and analysis, *J Comput Chem 25*, 1605-1612.
157. Binkley, J. S., Pople, J. A., and Hehre, W. J. (1980) Self-Consistent Molecular-Orbital Methods .21. Small Split-Valence Basis-Sets for 1st-Row Elements, *J Am Chem Soc 102*, 939-947.
158. Clark, T., Chandrasekhar, J., Spitznagel, G. W., and Schleyer, P. V. (1983) Efficient Diffuse Function-Augmented Basis-Sets for Anion Calculations .3. The 3-21+G Basis Set for 1st-Row Elements, Li-F, *J Comput Chem 4*, 294-301.
159. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes, *J Biomol Nmr 6*, 277-293.
160. Goddard, T. G., Kneller D.G. SPARKY 3, University of California, San Francisco.
161. Markham, G. D., Hafner, E. W., Tabor, C. W., and Tabor, H. (1980) S-Adenosylmethionine Synthetase from Escherichia-Coli, *J Biol Chem 255*, 9082-9092.
162. Park, J., Tai, J. Z., Roessner, C. A., and Scott, A. I. (1995) Overcoming Product Inhibition of S-Adenosyl-L-Methionine (Sam) Synthetase - Preparation of Sam on the 30 Mm Scale, *Bioorg Med Chem Lett 5*, 2203-2206.

163. Seeger, K., Lein, S., Reuter, G., and Berger, S. (2005) Saturation transfer difference measurements with SU(VAR)3-9 and S-adenosyl-L-methionine, *Biochemistry 44*, 6208-6213.
164. Harris, R. K. (1983) *Nuclear Magnetic Resonance Spectroscopy*, Pitman books limited, London.
165. Afonin, A. V., Ushakov, I. A., Kuznetsova, S. Y., Petrova, O. V., Schmidt, E. Y., and Mikhaleva, A. I. (2002) C-H center dot center dot center dot X (X = N, O, S) intramolecular interaction in 1-vinyl-2-(2 '-heteroaryl)pyrroles as monitored by H-1 and C-13 NMR spectroscopy, *Magn Reson Chem 40*, 114-122.
166. Sigalov, M., Vashchenko, A., and Khodorkovsky, V. (2005) Aromatic C-H center dot center dot center dot O interactions in a series of bindone analogues. NMR and quantum mechanical study, *J Org Chem 70*, 92-100.
167. Afonin, A. V., Ushakov, I. A., Vashchenko, A. V., Simonenko, D. E., Ivanov, A. V., Vasil'tsov, A. M., Mikhaleva, A. I., and Trofimov, B. A. (2009) C-H center dot center dot center dot N and C-H center dot center dot center dot O intramolecular hydrogen bonding effects in the H-1, C-13 and N-15 NMR spectra of the configurational isomers of 1-vinylpyrrole-2-carbaldehyde oxime substantiated by DFT calculations, *Magn Reson Chem 47*, 105-112.
168. Sarkhel, S., and Desiraju, G. R. (2004) N - H...O, O - H...O, and C - H...O hydrogen bonds in protein-ligand complexes: Strong and weak interactions in molecular recognition, *Proteins 54*, 247-259.
169. Couture, J. F., Collazo, E., Hauk, G., and Trievel, R. C. (2006) Structural basis for the methylation site specificity of SET7/9, *Nat Struct Mol Biol 13*, 140-146.
170. Trievel, R. C., Flynn, E. M., Houtz, R. L., and Hurley, J. H. (2003) Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT, *Nat Struct Biol 10*, 545-552.
171. Mehl, R. A., Anderson, J. C., Santoro, S. W., Wang, L., Martin, A. B., King, D. S., Horn, D. M., and Schultz, P. G. (2003) Generation of a bacterium with a 21 amino acid genetic code, *J Am Chem Soc 125*, 935-939.
172. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli, *Science 292*, 498-500.
173. Ishima, R., Louis, J. M., and Torchia, D. A. (1999) Transverse C-13 relaxation of CHD2 methyl isotopmers to detect slow conformational changes of protein side chains, *J Am Chem Soc 121*, 11589-11590.
174. Ishima, R., Petkova, A. P., Louis, J. M., and Torchia, D. A. (2001) Comparison of methyl rotation axis order parameters derived from model-free analyses of H-2

and C-13 longitudinal and transverse relaxation rates measured in the same protein sample, *J Am Chem Soc 123*, 6164-6171.
175. Ollerenshaw, J. E., Tugarinov, V., and Kay, L. E. (2003) Methyl TROSY: explanation and experimental verification, *Magn Reson Chem 41*, 843-852.
176. Nicholson, L. K., Kay, L. E., Baldisseri, D. M., Arango, J., Young, P. E., Bax, A., and Torchia, D. A. (1992) Dynamics of Methyl-Groups in Proteins as Studied by Proton-Detected C-13 Nmr-Spectroscopy - Application to the Leucine Residues of Staphylococcal Nuclease, *Biochemistry 31*, 5253-5263.
177. Zhang, X., Sui, X. G., and Yang, D. W. (2006) Probing methyl dynamics from C-13 autocorrelated and cross-correlated relaxation, *J Am Chem Soc 128*, 5073-5081.
178. Tugarinov, V., and Kay, L. E. (2006) Relaxation rates of degenerate H-1 transitions in methyl groups of proteins as reporters of side-chain dynamics, *J Am Chem Soc 128*, 7299-7308.
179. Kleywegt, G. J., Harris, M. R., Zou, J. Y., Taylor, T. C., Wahlby, A., and Jones, T. A. (2004) The Uppsala Electron-Density Server, *Acta Crystallogr D 60*, 2240-2249.
180. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D 66*, 486-501.
181. Vargas, R., Garza, J., Dixon, D. A., and Hay, B. P. (2000) How strong is the C(alpha)-H center dot center dot center dot O=C hydrogen bond?, *J Am Chem Soc 122*, 4750-4755.
182. Yesselman, J. D., Price, P. J., Knight, J. L., and Brooks, C. L., 3rd. (2011) MATCH: An Atom-Typing Toolset for Molecular Mechanics Force Fields, *J Comput Chem*.
183. Steiner, T. (2000) Influence of C-H center dot center dot center dot O interactions on the conformation of methyl groups quantified from neutron diffraction data, *J Phys Chem A 104*, 433-435.
184. Pierce, A. C., Sandretto, K. L., and Bemis, G. W. (2002) Kinase inhibitors and the case for CH center dot center dot center dot O hydrogen bonds in protein-ligand binding, *Proteins 49*, 567-576.
185. Fisher, R. a., and Yates, F. (1948) *Statistical Tables for Biological, Agricultural, and Medical Research*, III ed., Oliver & Boyd, London.
186. Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C., and Mehl, R. A. (2007) Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization, *Nat Protoc 2*, 2601-2607.
187. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Method Enzymol 276*, 307-326.
188. Vagin, A., and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement, *Acta Crystallogr D 56*, 1622-1624.

189. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use, *Acta Crystallogr D 60*, 2184-2195.
190. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D 66*, 12-21.
191. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr D 54*, 905-921.
192. Meulen, K. A. V., and Butcher, S. E. (2012) Characterization of the kinetic and thermodynamic landscape of RNA folding using a novel application of isothermal titration calorimetry, *Nucleic Acids Res 40*, 2140-2151.
193. Hansen, A. L., and Al-Hashimi, H. M. (2007) Dynamics of large elongated RNA by NMR carbon relaxation, *J Am Chem Soc 129*, 16072-16082.
194. Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A., and Al-Hashimi, H. M. (2009) Extending the Range of Microsecond-to-Millisecond Chemical Exchange Detected in Labeled and Unlabeled Nucleic Acids by Selective Carbon R(1 rho) NMR Spectroscopy, *J Am Chem Soc 131*, 3818-+.
195. Zhang, Q., Sun, X. Y., Watt, E. D., and Al-Hashimi, H. M. (2006) Resolving the motional modes that code for RNA adaptation, *Science 311*, 653-656.
196. de la Torre, J. G., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations, *J Magn Reson 147*, 138-146.
197. Cho, C. H., Urquidi, J., Singh, S., and Robinson, G. W. (1999) Thermal offset viscosities of liquid H2O, D2O, and T2O, *J Phys Chem B 103*, 1991-1994.
198. Markwick, P. R. L., Bouvignies, G., and Blackledge, M. (2007) Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy, *J Am Chem Soc 129*, 4724-4730.
199. Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992) A 500-Ps Molecular-Dynamics Simulation Study of Interleukin-1-Beta in Water - Correlation with Nuclear-Magnetic-Resonance Spectroscopy and Crystallography, *J Mol Biol 226*, 239-250.

200. Couture, J. F., Dirk, L. M. A., Brunzelle, J. S., Houtz, R. L., and Trievel, R. C. (2008) Structural origins for the product specificity of SET domain protein methyltransferases, *P Natl Acad Sci USA 105*, 20659-20664.
201. Chirpich, T. P., Zappia, V., Costilow, R. N., and Barker, H. A. (1970) Lysine 2,3-Aminomutase - Purification and Properties of a Pyridoxal Phosphate and S-Adenosylmethionine-Activated Enzyme, *J Biol Chem 245*, 1778-&.
202. Jones, B. N., Quang-Dang, D. U., Oku, Y., and Gross, J. D. (2008) A kinetic assay to monitor RNA decapping under single- turnover conditions, *Methods in enzymology 448*, 23-40.
203. Dapprich, S., Komaromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *J Mol Struc-Theochem 461*, 1-21.
204. Ditchfield, R., Hehre, W. J., and Pople, J. A. (1971) Self-Consistent Molecular-Orbital Methods .9. Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules, *J Chem Phys 54*, 724-&.
205. Boys, S. F., and Bernardi, F. (1970) Calculation of Small Molecular Interactions by Differences of Separate Total Energies - Some Procedures with Reduced Errors, *Mol Phys 19*, 553-&.
206. Simon, S., Duran, M., and Dannenberg, J. J. (1996) How does basis set superposition error change the potential surfaces for hydrogen bonded dimers?, *J Chem Phys 105*, 11024-11031.
207. Frisch, M. J., Headgordon, M., and Pople, J. A. (1990) A Direct Mp2 Gradient-Method, *Chem Phys Lett 166*, 275-280.
208. Headgordon, M., Pople, J. A., and Frisch, M. J. (1988) Mp2 Energy Evaluation by Direct Methods, *Chem Phys Lett 153*, 503-506.
209. Griffith, S. C., Sawaya, M. R., Boutz, D. R., Thapar, N., Katz, J. E., Clarke, S., and Yeates, T. O. (2001) Crystal structure of a protein repair methyltransferase from Pyrococcus furiosus with its L-isoaspartyl peptide substrate, *J Mol Biol 313*, 1103-1116.
210. Nureki, O., Watanabe, K., Fukai, S., Ishii, R., Endo, Y., Hori, H., and Yokoyama, S. (2004) Deep knot structure for construction of active site and cofactor binding site of tRNA modification enzyme, *Structure 12*, 593-602.
211. Boal, A. K., Grove, T. L., McLaughlin, M. I., Yennawar, N. H., Booker, S. J., and Rosenzweig, A. C. (2011) Structural Basis for Methyl Transfer by a Radical SAM Enzyme, *Science 332*, 1089-1092.

212. Yang, J., Kulkarni, K., Manolaridis, I., Zhang, Z., Dodd, R. B., Mas-Droux, C., and Barford, D. (2011) Mechanism of Isoprenylcysteine Carboxyl Methylation from the Crystal Structure of the Integral Membrane Methyltransferase ICMT, *Mol Cell 44*, 997-1004.
213. Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C., and Warren, M. J. (1998) The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase, *Nat Struct Biol 5*, 585-592.
214. Suzuki, Y., Noma, A., Suzuki, T., Ishitani, R., and Nureki, O. (2009) Structural basis of tRNA modification with CO2 fixation and methylation by wybutosine synthesizing enzyme TYW4, *Nucleic Acids Res 37*, 2910-2925.
215. Brown, H. C., McDaniel, D. H., and Hafliger, O. (1955) Dissociation Constants, In *Determination of Organic Structures by Physical Methods* (Braude, E. A., and Nachod, F. C., Eds.), pp 567-662, Academic Press Inc., New York, N.Y.
216. Scheiner, S. (2005) Relative strengths of NH center dot center dot O and CH center dot center dot O hydrogen bonds between polypeptide chain segments, *J Phys Chem B 109*, 16132-16141.
217. Fersht, A. (1999) *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, W.H. Freeman, New York.
218. Dirk, L. M. A., Flynn, E. M., Dietzel, K., Couture, J. F., Trievel, R. C., and Houtz, R. L. (2007) Kinetic manifestation of processivity during multiple methylations catalyzed by SET domain protein methyltransferases, *Biochemistry 46*, 3905-3915.
219. Igumenova, T. I., Frederick, K. K., and Wand, A. J. (2006) Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution, *Chem Rev 106*, 1672-1699.
220. 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.
221. Merriam, G. R., Maclusky, N. J., Picard, M. K., and Naftolin, F. (1980) Comparative Properties of the Catechol Estrogens, .1. Methylation by Catechol-O-Methyltransferase and Binding to Cytosol Estrogen-Receptors, *Steroids 36*, 1-11.
222. Thithapa.A. (1972) Substrate-Specificity and Heterogeneity of N-Methyltransferases, *Biochem Bioph Res Co 47*, 301-&.
223. Ellermann, M., Lerner, C., Burgy, G., Ehler, A., Bissantz, C., Jakob-Roetne, R., Paulini, R., Allemann, O., Tissot, H., Grunstein, D., Stihle, M., Diederich, F., and Rudolph, M. G. (2012) Catechol-O-methyltransferase in complex with substituted

3'-deoxyribose bisubstrate inhibitors, *Acta crystallographica. Section D, Biological crystallography 68*, 253-260.

224. Zheng, W., Ibanez, G., Wu, H., Blum, G., Zeng, H., Dong, A., Li, F., Hajian, T., Allali-Hassani, A., Amaya, M. F., Siarheyeva, A., Yu, W., Brown, P. J., Schapira, M., Vedadi, M., Min, J., and Luo, M. (2012) Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2, *J Am Chem Soc 134*, 18004-18014.
225. Seela, F., Münster, I., Lüchner, U., and Rosemeyer, H. (2004) 8-Azaadenosine and Its 2′-Deoxyribonucleoside: Synthesis and oligonucleotide base-pair stability, *Helvetica chimica acta 81*, 1139-1155.
226. Sugiyama, T., Schweinberger, E., Kazimierczuk, Z., Ramzaeva, N., Rosemeyer, H., and Seela, F. (2000) 2-Aza-2′-deoxyadenosine: Synthesis, Base-Pairing Selectivity, and Stacking Properties of Oligonucleotides, *Chem-Eur J 6*, 369-378.
227. Sugiyama, T., Schweinberger, E., Kazimierczuk, Z., Ramzaeva, N., Rosemeyer, H., and Seela, F. (2000) 2-aza-2 '-deoxyadenosine: Synthesis, base-pairing selectivity, and stacking properties of oligonucleotides, *Chem-Eur J 6*, 369-378.
228. Hoesl, M. G., and Budisa, N. (2012) Recent advances in genetic code engineering in Escherichia coli, *Current opinion in biotechnology 23*, 751-757.
229. Bartlett, G. J., Choudhary, A., Raines, R. T., and Woolfson, D. N. (2010) n ->pi* interactions in proteins, *Nat Chem Biol 6*, 615-620.
230. Carney, A. E., and Holden, H. M. (2011) Molecular architecture of TylM1 from Streptomyces fradiae: an N,N-dimethyltransferase involved in the production of dTDP-D-mycaminose, *Biochemistry 50*, 780-787.
231. 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.