UNIVERSITY OF CALIFORNIA
RIVERSIDE

Investigation of Proton-Bridged Amines

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Hou Uy Ung

December 2015

Dissertation Committee:
Dr. Thomas H. Morton, Chairperson
Dr. Catharine Larsen
Dr. Michael Marsella

The Dissertation of Hou Uy Ung is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgement

The last five years of graduate school has been the most stressful and rewarding experience of my life. Without the guidance of Prof. Thomas Morton, I would not be the person I am today. He didn't just teach me to become a better chemist, but also a better person. The most memorable lesson I can recall is if something doesn't work, use a bigger hammer.

I want to thank Dr. Leonard Mueller for taking the time to hood me at the commencement ceremony, Dr. Richard Kondrat, and Dr. Dan Borchardt for teaching me mass spectrometry and NMR. IR multiple photon dissociation experiments would not be possible if not for the help of Dr. Jos Oomens, Dr. Giel Berden, along with the staffs and students formally at the FOM Institute for Plasma Physics.

I want to also take this time to thank the current graduate students and those who graduated. When I first join the Morton Lab, Dr. Aaron Moehlig was the only person here to mentor me. He was easy to talk to and always ready to help. After Dr. Aaron Moehlig graduated, I turned to Dr. Michael Young for advice. Whenever I was stumped by a problem, my first instinct was to find Mike. I would also like to thank all of the graduate and undergraduate students: Jay-Ar Bendo, Eric Commendatore, Dr. Omar Hamdy, Dr. Yoo-Jin Ghang, Erik Romero, Taylor Le, Carroll Hy, Evelyn Lopez, Ranier Rivera, the Hooley Group, and Larsen Group for making the last five years memorable and tolerable.

I would like to thank my family. To my parents and my brother, thank you for always supported what I did. Finally, I would like to thank my loving wife, Quyen Nguyen. She is always supportive of my life choices and has sacrificed so much to make graduate school possible for me. None of this would have been possible without the support of my family.

ABSTRACT OF THE DISSERTATION

Investigation of Proton-Bridged Amines

by

Hou Uy Ung

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, December 2015
Dr. Thomas H. Morton, Chairperson

Hydrogen bonds range from those observed between neutral partners (such as water molecules) to ionic hydrogen bonds (where a charge is associated with either the hydrogen bond donor or hydrogen bond acceptor). This dissertation explores the formation and stabilization of ionic hydrogen bonding species in: mono- and diprotonated polyamines, cytosine and isocytosine proton-bound dimers (PBDs), and protonated nucleosides.

Vibrational spectra of two polyamines, capable of forming N-$H^{+\cdots}$O and N-$H^{+\cdots}$N intramolecular hydrogen bonds, are recorded using IR Multiple Photon Dissociation (IRMPD) in the fingerprint and CH/NH/OH stretching domain. Density functional theory (DFT) calculations (B3LYP/6-311++G** level) were utilized to predict dissociation pathways of each charged species in the gas phase. Experimental IRMPD spectra of the two polyamines are plotted against

their deuterated analogs. Bands that are present in the protonated spectrum but absent in the deuterated spectrum were assigned to be motions associated with the bridging proton.

^{13}C and ^{15}N solid state NMR spectra of 1-methylcytosine PBD salts suggest the bridging proton is shared unequally between the two nucleobases, resulting in an asymmetrical unit. X-ray diffraction shows two different crystal structure dimensions and morphology of 1-methylcytosine PBD iodide salt with one being identical to 1-methylcytosine PBD iodide salt-d_5. Inelastic neutron scattering (INS) spectra compared completely protonated PBDs with deuterated analogs of 1-methylcytosine PBD iodide salt and revealed bands potentially corresponding to methyl torsions and N-H bends.

Proton affinities of nucleosides along with the IRMPD spectra of their protonated analogs in the CH/NH/OH domain are presented. DFT calculations (at B3LYP/6-311++G**) were utilized to predict the favored protonation sites. Partial deuteration experiments suppress combination bands and overtones of fundamental bands that are otherwise present in the CH/NH/OH domain. Conformation of the furanose ring prefers the South orientation for all three protonated nucleosides

Three target molecules are presented and their different synthetic pathways are outlined to achieve each target molecule. The three target molecules are modified cytosine and isocytosine derivatives and upon

protonation, both molecules are capable of forming proton-bound dimer with either neutral or protonated cytosine and itself.

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Charts

List of Equations

Chapter I:

Polyamines

Introduction

Inter- and intramolecular hydrogen bonds are essential, for without these forces life on earth could not be possible. This type of bonding ranges from ordinary hydrogen bonds (as seen between two water molecules) to strong (also known as ionic) hydrogen bonds between an ion and a neutral molecule. The difference between these two types of hydrogen bonds is that, in the latter, a positive charge is associated with the hydrogen bond donor, or a negative charge is associated with the hydrogen bond acceptor. In this context, the term proton bridge refers to the intramolecular ionic hydrogen bond between two basic sites, regardless of whether the proton resides between two nitrogens, between two oxygens, or between a nitrogen and an oxygen.

Hydrogen bonds tend to be much stronger in the gas phase than in solution. The enthalpy for dissociating the strongest ionic hydrogen bond, the bifluoride ion (FHF^-) is about 40 kcal/mol[1,2]. Because the $T\Delta S$ term at room temperature and atmospheric pressure is on the order of 9 kcal/mole, single hydrogen bonds between neutral molecules are infrequently seen in the gas phase unless the temperature is lower than ambient. However, ionic hydrogen bonds are quite common and are easily seen in the mass spectrometer.

Water is primarily thought of when discussing hydrogen bonds. In 2003, Asmis *et al.* obtained the first vibrational spectra of the proton-bound dimer of water ($H_5O_2^+$), between 600 – 1900 cm^{-1}, in the gas phase, by plotting the dissociation of H_3O^+ and D_3O^+ as a function of wavenumber[3]. As reproduced in

Figure 1.1, five band maxima are reported at 788, 921, 1043, 1317, and 1741 cm^{-1}, and all five bands shift to lower frequencies upon deuteration by a factor of approximately $1/\sqrt{2}$. The band originally at 788 cm^{-1} shifted out of the range investigated. Both oxygens are somewhat pyramidal and the O-H-O bond, while symmetrically placed, deviates slightly from linearity.

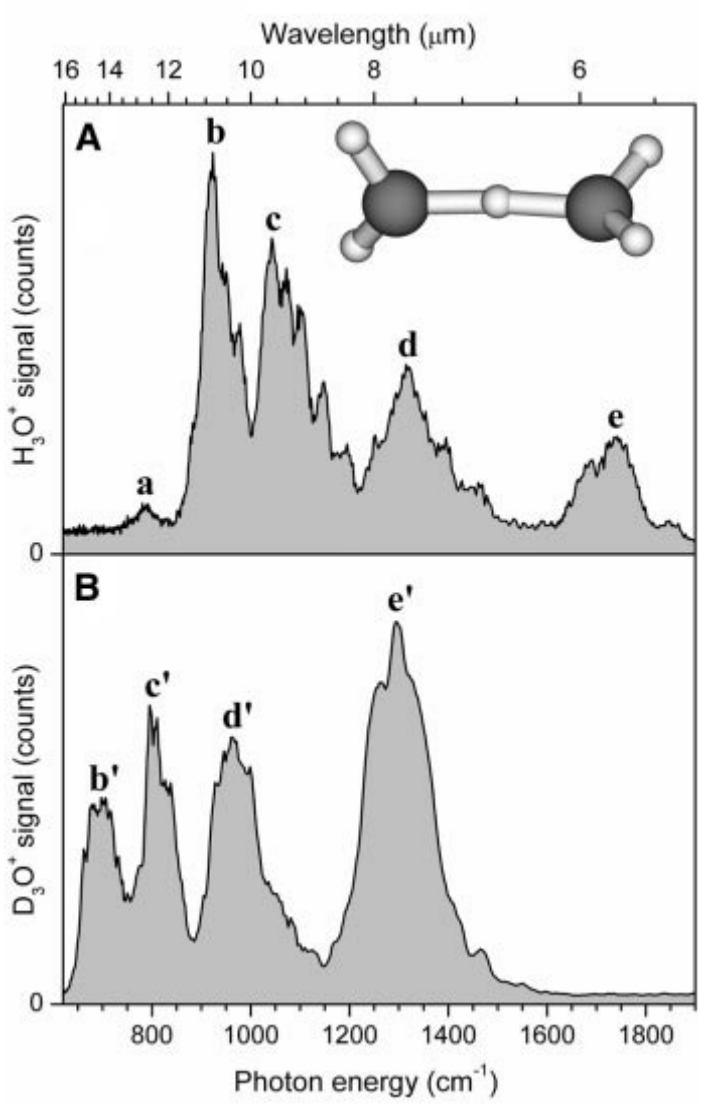


Figure 1.1[3]: **A**) Experimental vibrational spectrum of protonated water dimer, showing five maxima, in the domain from 600 – 1900 cm^{-1}. **B**) Experimental vibrational spectrum of deuterated deuterium oxide dimer in the domain from 600 – 1900 cm^{-1}. All peaks in panel **A** are shifted to lower frequencies upon deuteration, as shown in panel **B**. Peak a from **Panel A** shifts out of the investigated domain.

Water is not the only molecule capable of forming $O\text{-}H^{+}\cdots O$ ionic hydrogen bond. Protonated ethers also have the capability to form inter- and intramolecular ionic hydrogen bonds. Eyler *et al.* reported the IRMPD spectra of PBDs of

protonated dimethyl (**Figure 1.2A**) and diethyl (**Figure 1.2B**) ethers with their corresponding neutral partner[4]. The intramolecular ionic hydrogen bond of protonated diglyme was also investigated. Experimental IRMPD spectra matched well to that predicted by Density Functional Theory (DFT), except in the case of protonated diglyme. The lowest energy conformation of protonated diglyme (**Figure 1.2C**) along with the protonated symmetric analog (**Figure 1.2D**) were both calculated and compared to the experimental spectrum, but neither predicted spectra match, except for the $O\text{-}H^{+\cdots}O$ asymmetric stretch predicted around 750 cm^{-1}. The $O\text{-}H^{+\cdots}O$ asymmetric stretch is mixed with the symmetric CO stretch and HCO bending modes of the monomer units, which are observed in the domain from 750 – 820 cm^{-1}. The $O\text{-}H^{+\cdots}O$ bending modes are coupled with the HCH bending and are observed in the domain from 1500 – 1600 cm^{-1} [4].

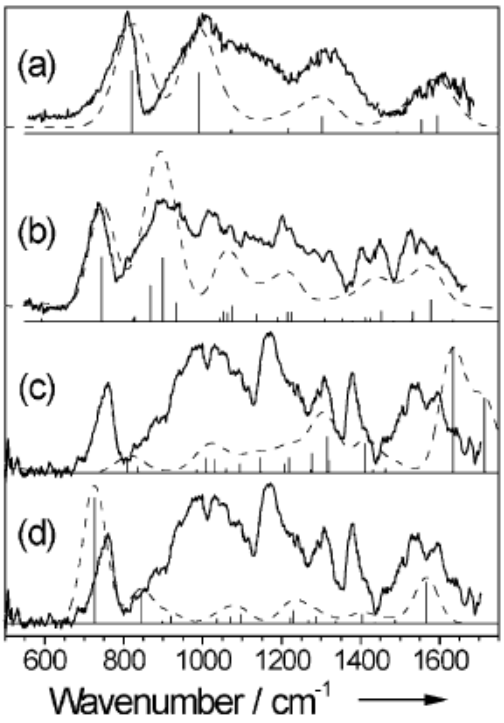


Figure 1.2[4]: Solid black traces represent the experimental IR spectra, while the dashed black lines depict the predicted IR spectra obtained at the MP2/cc-PVDZ level. **A)** Overlap of predicted and experimental IR spectra of the PBD of protonated dimethylether with its corresponding neutral component. **B)** Overlap of predicted and experimental IR spectra of the PBD of protonated diethylether with its neutral component. **C)** Overlap of predicted and experimental IR spectra of the PBD of protonated diglyme in its lowest calculated energy state. **D)** Overlap of predicted and experimental IR spectra of the PBD of protonated diglyme with symmetry applied to the molecule.

Kelly Theel and Prof. Gregory Beran attempted to map out a 2-dimensional potential energy surface (PES) diagram for the bridging proton in protonated diglyme. The calculation was performed by fixing the $O\text{-}H^{+}\cdots O$ bond at B3LYP/6-31+G(d,p). The PES diagram shows a double-well potential, but extremely flat.

Eyler *et. al.* recently published another article regarding the IRMPD comparison between the experimental IRMPD spectrum and prediction, at MP2/cc-pVTZ, of diglyme in the fingerprint domain from 930 – 1090 cm^{-1} (**Figure 1.3**)[5]. The $O\text{-}H^{+\cdots}O$ bond angle is calculated to be 177° and the distance between the bridging proton to the two methoxy oxygens are 1.12 and 1.31 Å. Unlike the previously IRMPD spectrum of diglyme, this spectrum was obtained using a tunable CO_2 laser. There is more detail in this IRMPD spectrum than that previously obtained. A large band is observed expermentally around 1040 cm^{-1} and what appear to be two smaller bands between 960 – 990 cm^{-1}. The domain below 960 cm^{-1} shows at least one band whereas calculations only predict one. Overall, the experimental spectrum agrees well with prediction except for a band that is predicted to be present at 1010 cm^{-1} but absent in the experimental spectrum.

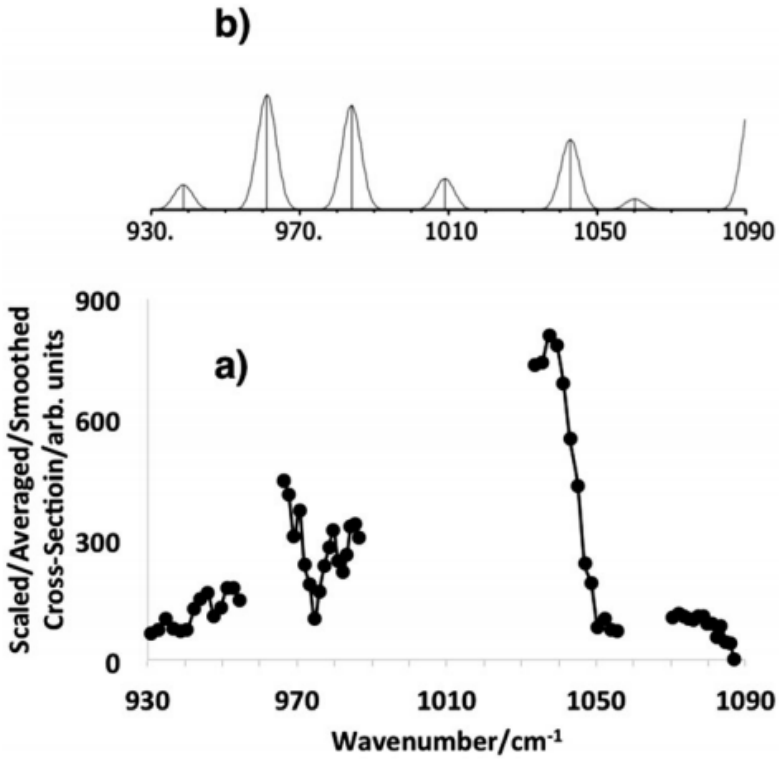


Figure 1.3[5]: **A)** Experimental IRMPD spectrum of protonated diglyme in the fingerprint domain from 930 – 1090 cm^{-1} using a tunable CO_2 laser. **B)** Scaled predicted IR spectrum of protonated diglyme calculated at MP2/cc-pVTZ.

The motion associated with the bridging proton across two basic sites can be represented by a double-well potential. If the two basic sites are of equal basicity, then the double-well should be symmetric. **Figure 1.4** depicts the potential energy diagram of the motion of a bridging proton, on protonated 3,3-N,N,N',N'-hexamethyl-1,5-pentanediamine, between two equally basic sites. The colored curves depict the wave functions within the double-well potential when the barrier lies below the zero point energy. As shown by the red curve, the wave function for the lowest level of the asymmetric stretch of the bridging proton has a

maximum near the top of the barrier for proton transit. This phenomenon represents a close to ideal case of low-barrier hydrogen bonding (LBHB).

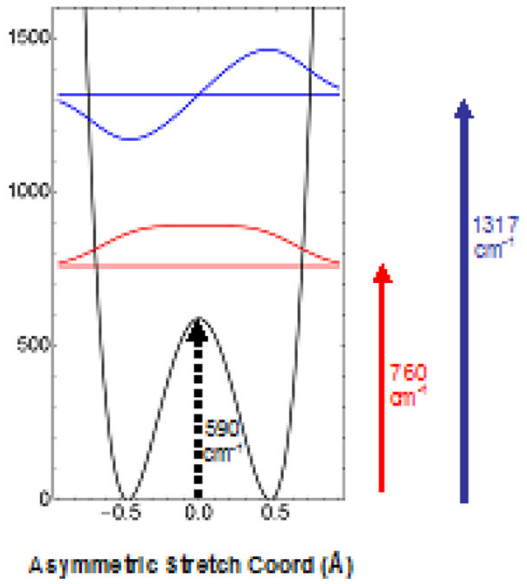
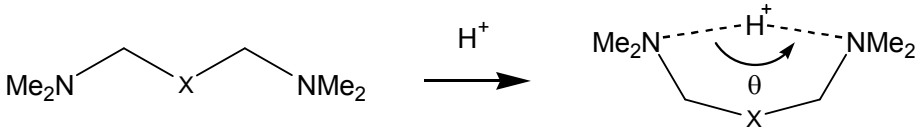


Figure 1.4[6]: 2-dimensional B2-P3LYP/cc-pVTZ/ / B3LYP/6-31G** potential for protonated 3,3-N,N,N',N'-hexamethyl-1,5-pentanediamine showing the positions and vibrational wave functions for the zero point (v = 0) and the v = 1 level for the asymmetric N-H^{+}···N stretch.

Background

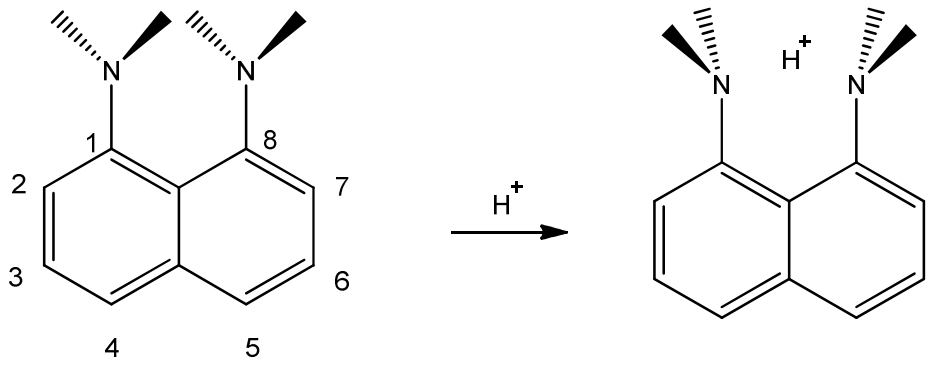
In 2011, Beran *et al*. put forth several criteria for modeling an intramolecular proton bridge as a linear triatomic having a LBHB[6]: (1) the proton must have a low barrier for transfer (2) the two basic groups must have identical proton affinities and (3) the geometry must be close to linear (θ > 170°) when the molecule cyclizes upon itself as depicted by **Eq 1**. It was determined experimentally that ions that give the lowest barrier for proton transit from one basic site to another also have N-H···N bond angles close to 180^{o}. Solution phase studies by Pluth *et al.* suggest that chain length also affects the strength of the

proton bridge[7]. Having the amino groups methylated avoids any unwanted hydrogen bonding interaction the molecule might have with the surrounding solvent[8].



(1)

Protonated 1,8-bis(dimethylamino)naphthalene (DMAN) is a cation that satisfies two of the aforementioned criteria, but has a N-$H^{+\cdots}$N bond angle < 170°, (around 160° depending upon the counterion[8-12]) upon protonation. DMAN is commonly referred to as a "proton sponge" due to its high proton affinity, as shown in **Eq 2**. DMAN was first discovered by Alder *et al*.[9] and has been the subject of many studies ever since. The positioning of the dimethylamino groups at the 1- and the 8-position of naphthalene produces a strained molecule if the nitrogens are coplanar with the rings. The steric bulk from the methyl groups force the two amines out of conjugation with the naphthalene ring and also pucker the naphthalene ring slightly out of planarity[8,9]. Protonation alleviates the steric stress because the methyl groups on each amine can orient away from each other.



(2)

The conjugate acid of DMAN creates an intramolecular proton bridge that should be shared by both dimethylamino groups, but is the bridging proton really shared equally? Many experiments have been performed in attempt to answer this question, but different experiments have produced different results[8-15]. Results from solution phase NMR experiments of protonated DMAN[8] do not agree with solid state NMR (ssNMR) or X-ray experiments. X-ray diffraction analysis places the proton in the center of the two dimethylamino groups, but solution phase NMR experiments has the bridging proton on either one amino group or the other. The use of different counter ions and different conditions also produced varying results[8-15]. To date, many different derivatives of DMAN have been synthesized and investigated in attempt to answer this question.

Although neutral DMAN may look symmetrical, solid state NMR (ssNMR) studies have shown the molecule to be unymmetrical[10-14]. The distance between the two amino groups was determined to be 2.789 Å. Results from ^{13}C ssNMR experiments show four distinct individual peaks, each peak corresponding to a methyl group, while the ^{1}H ssNMR experiments show two separate methyl

resonances. Steric interactions between methyl groups force the molecule to be unymmetrical.

Different conjugate acid DMAN salts were examined using ssNMR, X-ray crystallography, infrared spectroscopy, and theoretical DFT calculations[13,14]. Using ^{1}H ssNMR, a peak is observed at 19.7 ppm, corresponding to the bridging proton, in both protonated DMAN thiocyanate and tetrafluoroborate salts[13]. X-ray diffraction of the protonated DMAN bromide salt and the deuterated DMAN bromide salt suggest both sides of the $N\text{-}H^{+\cdots}N$ bonds to be symmetrical. Calculations at the MP2/6-31G** level places the $N\text{-}H^{+\cdots}N$ distance at 2.567 Å. The experimental $N\text{-}H^{+\cdots}N$ distance was determined to be 2.567 Å, which agrees with calculation. N-H and N-D distances were determined to be exactly midway between the two nitrogens, resulting in a symmetric molecule upon protonation. There was no change in distance when the bridging proton was replaced with a bridging deuterium. The infrared spectrum in the fingerprint domain of the protonated DMAN bromide salt shows a band 488 cm^{-1}, which is assigned to be the $N\text{-}H^{+\cdots}N$ asymmetric stretch, and the $N\text{-}D^{+\cdots}N$ asymmetric stretch occurs at 235 cm^{-1} [14].

A derivative of DMAN that has an extra methylene group between naphthalene and the dimethylamino group was studied by Sobczyk *et al.*[15]. 1,8-dimethylaminomethylnaphthalene (DMAMN) isn't as strained as DMAN and is allowed more flexibility. The protonated perchlorate salt of DMAMN was investigated using X-Ray diffraction. Unlike the salts of $DMANH^{+}$ Br^{-} and

$DMAND^+$ Br^-, the protonated perchlorate salt of DMAMN was determined to be asymmetrical due to the flexibility allowed by the extra carbon. The distance between the two dimethylamino groups was determined to be 2.675 Å with a bridging $N\text{-}H^{+\cdots}N$ bond angle of 167^o, 7^o closer to linearity than protonated DMAN. The N-H distance of the bound proton is 1.18 Å, while the other N-H distance is 1.51 Å, indicating the bridging proton prefers to reside close to one amino group rather than in the center.

What happens if one of the amino groups is exchanged a different functional group capable of hydrogen bonding? Johnson *et al.* investigated the gas phase vibrational spectra of DMAN derivatives where the dimethylamino group at the 8-position is exchanged for a hydroxy, fluorine, and benzoate group[16]. Due to the difference in basicity, the bridging proton doesn't reside in the middle, but closer to one side of the molecule than the other. The protonated and deuterated IR spectra were obtained by monitoring the photodissociation of a H_2 adduct that is weakly bound to the mass-selected ion. Comparison between the protonated and deuterated IR spectra are reproduced in **Figure 1.5**. For all spectra in **Figure 1.5**, the upward traces correspond to the experimental spectra while downward traces correspond to the DFT calculated harmonic spectra. A comparison between protonated and deuterated 1-dimethylamino-8-hydroxy-naphthalene (**Figure1.5C** and **Figure1.5D**) shows two bands that shift upon deuteration from approximately 3050 and 2800 to 2250 cm^{-1}.

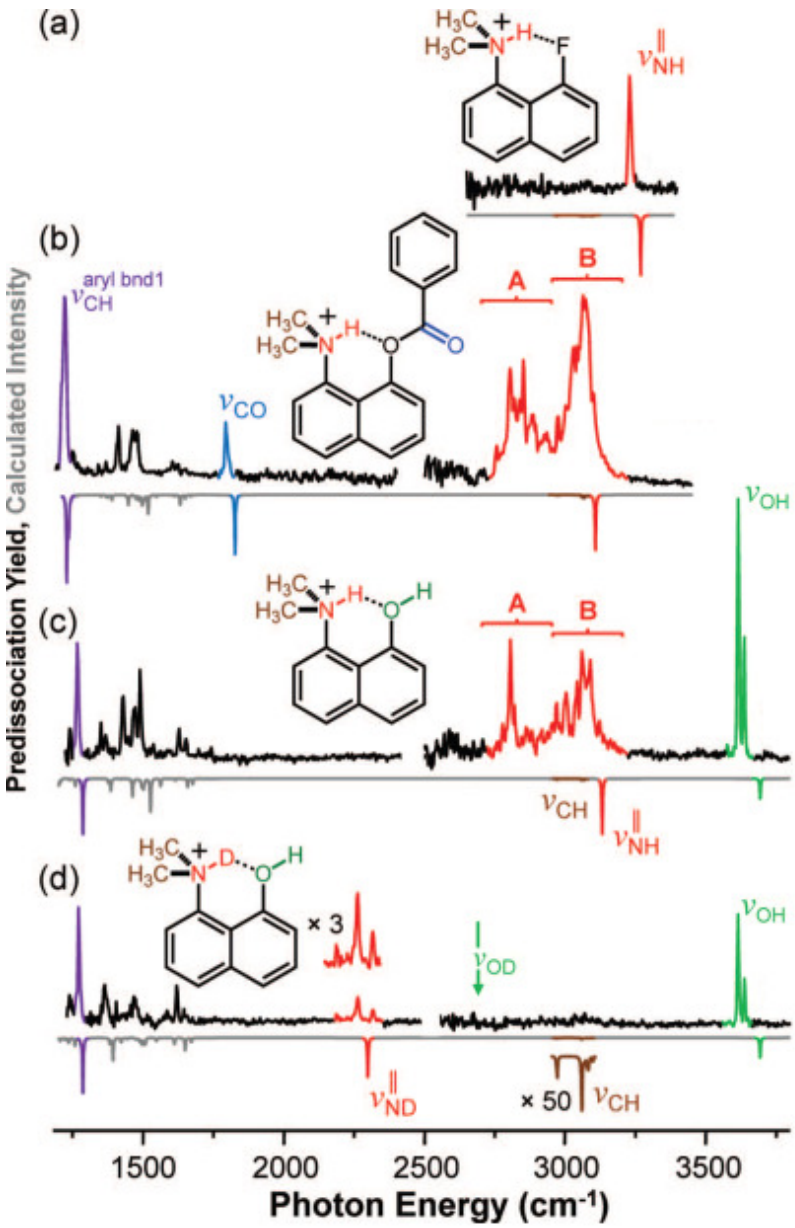


Figure 1.5[16]: Traces flowing upward correspond to the experimental gas phase spectra while downward traces represent the predicted harmonic IR spectra calculated at the B3LYP/6-311++G** level. All of the groups at the 8-position form an intramolecular hydrogen bond with the protonated dimethylamino group **A)** A fluorine is at the 8-position of naphthalene. **B)** A benzoate group is at the 8-position of naphthalene. **C)** A hydroxyl group resides at the 8-position of naphthalene. **D)** The same hydroxyl group is at the 8-position as in **C** but with all of the exchangeable hydrogens are replaced with deuterium.

Although solid state experiments place the bridging proton in the middle between the two amino groups of DMAN, a solution phase experiment does not provide the same picture. Protonated DMAN bromide, chloride, perchlorate and tetraphenylborate salts were investigated using solution phase NMR with D_2O and deuterated acetronitrile as solvents. Results indicate that different counter ions and different solvents affect the N-H distances. In terms of N-H bond distances, the conjugate acid of DMAN tetraphenylborate salt has the longest N-H bond distance at 1.115 Å. The shortest N-H bond distance belonged to the conjugate acid of DMAN chloride salt at 1.075 Å[10]. The conjugate acid DMAN bromide salt was investigated under different solvents. In water, the N-H distance was determined to be 1.089 Å. When the same experiment was conducted in acetonitrile, the N-H distance was determined to be slightly longer at 1.095 Å[10]. The interaction with different solvent environments affected the proton transfer potentials, resulting in different N-H bond distances.

The HBr salt of 1,8-dimethylamino-2,7-dibromonaphthalene (2,7-Br_2-DMAN) was investigated using X-ray diffraction, and the bridging proton was found to be in the middle between the two amino groups. The $N\text{-}H^{+}\cdots N$ distance was found to be 2.547 Å, and the compound has a $N\text{-}H^{+}\cdots N$ bond angle of 162^{o}. X-ray diffraction of neutral 2,7-Br_2-DMAN revealed four distinct N-N distances, at 2.748, 2.750, 2.787, and 2.815 Å, suggesting neutral 2,7-Br_2-DMAN is not a symmetric molecule. Protonation relieves the steric stress the molecule is under

by forming an intramolecular hydrogen bond, making the protonated analog symmetrical.

The proton affinity of DMAN can be adjusted by substituting different substituents to the naphthalene ring. Perrin *et al.*[8,17] and Pozharskii *et al.* investigated the proton affinity of DMAN modified at the 2,4,5, and 7- position of the naphthalene ring. It was determined that having electron donating groups at the 2,7- position (*ortho*) or 4,5- position (*para*) increased the proton affinity of DMAN[17]. **Table 1.1** shows the calculated proton affinity of various substituents placed at the 2,4,5, and 7- position of DMAN in the gas phase and in DMSO. Experimental results were determined to match that of DFT calculations.

From **Table 1.1**, electron donating groups on the 2 and 7-position of naphthalene are shown to increase proton affinity. Having the 4 and 5-position of naphthalene substituted also increased the proton affinity, but not as drastic as having the same substituents on the 2 and 7-position. Hydroxy and methoxy groups were substituted at these positions and their experimental proton affinities were found to match that of DFT predictions. When the hydroxyl groups are deprotonated, leaving behind a formal charge on oxygen, this increased the calculated proton affinity even more.

With hydroxy and methoxy groups showing an increase in basicity at the 2 and 7-position, dimethylamino groups were also investigated at these two positions. Although the two dimethylamino groups are capable of donating electrons into the ring, their steric bulk decreased the basicity. Because the

dimethylamino groups are so close in space, they must orient themselves in a way that is unfavorable, even upon protonation.

		B3LYP/6-31G**		B3LYP/6-311++G**	
		PA(Gas)	PA(DMSO)	PA(Gas)	PA(DMSO)
1	H	259.6	293.5 (293.3)	253.4	287.3
2	2,7-$(OMe)_2$	274.7	303.9	268.2	298.2
4	2,7-$(OH)_2$	272.6	304.4 (304.2)	265.3	297.5
5	2,7-$(O^-)_2$	440.4	334.6 (334.8)	424.8	321.4
9	2-OH, 7-O^-	359.8	322.0 (322.7)	349.0	312.0
10	4,5-$(OH)_2$	263.5	296.8 (297.0)	256.6	290.5
12	4-OH, 5-O^-	340.5	310.1 (311.0)	331.5	301.5
13	4,5-$(O^-)_2$	419.8	312.7	406.1	301.7
OH^-	–	435.1	348.5 (348.4)	396.0	313.3 (313.3)
NH_2^-	–	445.2	369.9 (369.7)	412.5	340.0 (339.9)
H^-	–	449.7	–	405.0	–

Table 1.1[17]: DFT calculated proton affinities of methoxy and hydroxy substituted DMAN at the 2,4,5, and 7-position in the gas phase and in DMSO. Proton affinities were calculated at the B3LYP/6-31G** and B3LYP/6-311++G** level. Having the 2 and 7 position substituted increased the basicity more than the 4- and 5-position. Values in parentheses are experimental values.

Since DMAN is a rigid molecule, less rigid diamines possessing higher degrees of flexibility were investigated to see if they exhibit any LBHB characteristics. Several linear diamines and one cyclic diamine were investigated to determine if this same band persists. Linear diamines were examined in the gas phase with varying carbon chain lengths, ranging from three to five, recreating a ring. Like DMAN, all diamines were methylated to prevent any unwarranted hydrogen bonding interactions with solvent. N,N,N',N'-

tetramethylpropanediamine ($Me_2NCH_2CH_2CH_2NMe_2$) (TMPD), N,N,N',N'-tetramethylputrescine ($Me_2NCH_2CH_2CH_2CH_2NMe_2$) (TMP), and N,N,N',N'-tetramethyl-cis-1,5-diaminocyclooctane (DACO) are all methylated diamines that were all synthesized and investigated by previous group members[18-20].

A proton bound internally between two basic sites act as a weaker Brønsted acid when compared to an isolated ammonium ion. Addition of a proton to a molecule with an amine and a remote functional group can lead to profound changes in conformation, as **Eq 1** illustrates for an acyclic tertiary diamine. A proton, bound internally between two basic sites, one of which is a basic nitrogen, acts as a weaker Brønsted acid when compared to an isolated ammonium ion.

For a linear LBHB the stretching and bending vibrations can be separated. The theoretical model treats the symmetric and asymmetric N-$H^{+\cdots}$N stretches in terms of an anharmonic potential while leaving all of the other vibrations as harmonic. In a case where the bridging proton is between two equally basic sites, the midpoint corresponds to a local maximum in the potential energy surface, as shown in **Figure 1.4**. Having the local maximum in the middle suggests a low barrier. Experiments have determined the equilibrium position of the bridging proton near the midpoint between two basic sites[21].

The N-$H^{+\cdots}$N bond angle of TMPD is 155^o (comparable to the N-$H^{+\cdots}$N bond angles of the aforementioned proton sponges) and should limit the ability to separate N-$H^{+\cdots}$N stretches (which shift down in frequency) from bending motions

(which shift up in frequency). **Figure 1.6** reproduces the gas phase vibrational spectrum of the protonated and deuteronated TMPD in the fingerprint IR domain from 550 – 1800 cm^{-1}. The comparison shows at least three bands that disappear with deuteration at 1200 cm^{-1}, 1300 cm^{-1}, and between 1350 – 1400 cm^{-1} (black arrows). Although the motions associated with these bands have yet to be assigned, this demonstrates that deuteration can cause more profound changes than previously observed. The IR spectrum in the lower fingerprint domain from 300 – 650 cm^{-1} was not obtained because the N-H^{+}···N bond angle deviates too far from linearity to be treated as a linear triatomic[18-20].

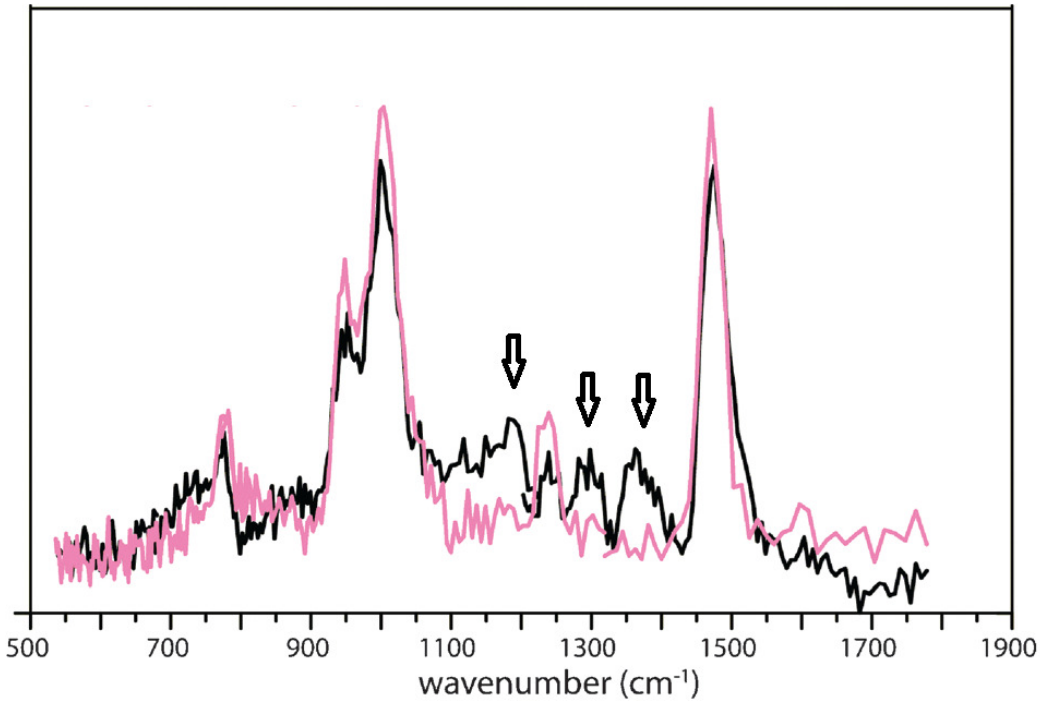


Figure 1.6[6]: Experimental gas phase vibrational spectra of protonated TMPD (black) plotted against deuterated TMPD (pink) in the fingerprint domain from 550 – 1800 cm^{-1}. Replacement of the bridging proton with a deuterium causes at least three bands to shift near 1200, 1300, and between 1350 – 1400 cm^{-1} (black arrows).

Morton *et. al.* previously investigated the protonated triflate and iodide salts of TMP[6,21]. TMP was selected for investigation because it has an $N\text{-}H^{+}\cdots N$ bond angle within 10^{o} of linearity. Calculations suggest that the $N\text{-}H^{+}\cdots N$ stretch should occur in the domain between 500 – 550 cm^{-1}. For molecules that deviate far from linearity, the effective barrier of proton transit might not correspond to that of **Figure 1.4**. The gas phase IR of protonated and deuteronated TMP in the fingerprint domain from 250 - 650 cm^{-1} is reproduced in **Figure 1.7**[6,18-20]. Comparison of the two gas phase IR spectra shows a band at 528 cm^{-1} that disappears upon deuteration. This band at 528 cm^{-1} (black arrow) is assigned to

be a motion associated with the bridging proton, more specifically the N-H^{+}···N asymmetric stretch, which falls into the range previously calculated.

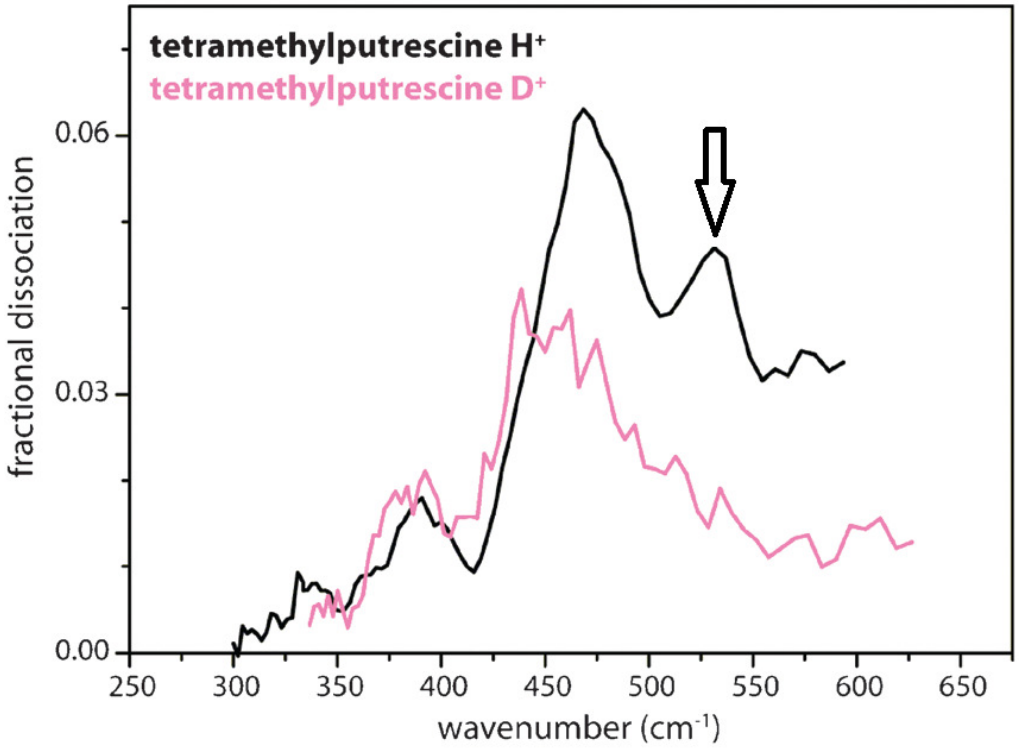


Figure 1.7[6]: Experimental IRMPD spectrum of protonated TMP (black) plotted over the experimental IRMPD spectrum of deuterated TMP (pink trace) in the fingerprint domain from 250 – 650 cm^{-1}. A band present in the protonated spectrum at 528 cm^{-1} (black arrow) disappears upon deuteration. This band is assigned to be the N-H^{+}···N asymmetric stretch.

To investigate whether the disappearance of the band at 528 cm^{-1} does indeed correspond to a motion of bridging proton, the gas phase vibrational spectrum of a control molecule (tetramethylornithine, $Me_2CH_2CH_2CH_2CH(NMe_2)COOH$) was obtained. The two nitrogens on this molecule no longer have the same basicity, and therefore should not behave like

a low-barrier hydrogen bonding molecule. **Figure 1.8** depicts the comparison between the proton bridged and deuteron bridge cation shows no a shift of band position in the fingerprint domain from 300 – 650 cm^{-1}.

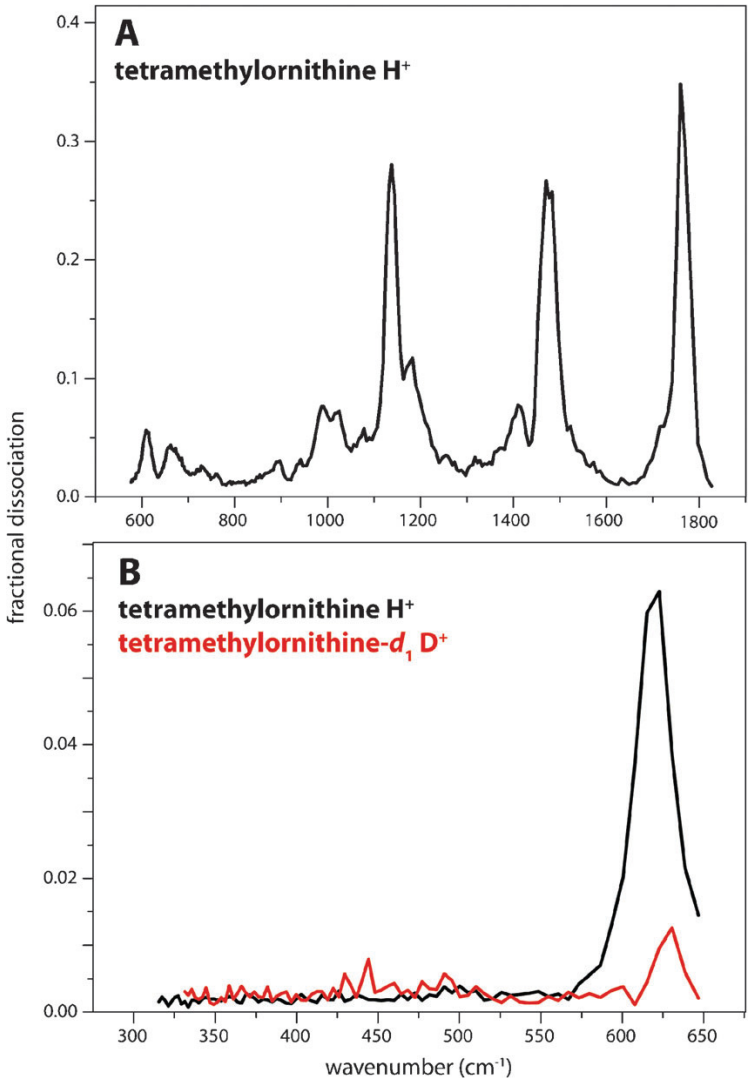


Figure 1.8[6]: Experimental IRMPD spectra of protonated (black) and deuterated (red) tetramethylornithine. **A**) Experimental IRMPD spectrum of protonated tetramethylornithine in the domain from 600 – 1800 cm^{-1}. **B**) Experimental IRMPD spectra of protonated tetramethylornithine plotted over deuterated tetramethylornithine in the fingerprint domain from 300 – 650 cm^{-1}.

INS experiments have also been performed on protonated TMP iodide and deuteronated TMP iodide salts. INS experiments are ideal for looking at motions associated with hydrogen due to their scattering cross section. Hydrogen has a scattering cross section more than 80 times that of deuterium and more than 160 times greater than that of nitrogen. Two bands disappear when the protonated spectrum is overlapped with the deuteronated spectrum at 190 cm^{-1} and at 525 cm^{-1} (black arrow), as shown in **Figure 1.9**. The former was assigned to be a phonon band and the latter to be the $N\text{-}H^{+\cdots}N$ asymmetric stretch. The $N\text{-}H^{+\cdots}N$ asymmetric stretch showed up almost in the same domain as in the gas phase IR spectrum[18].

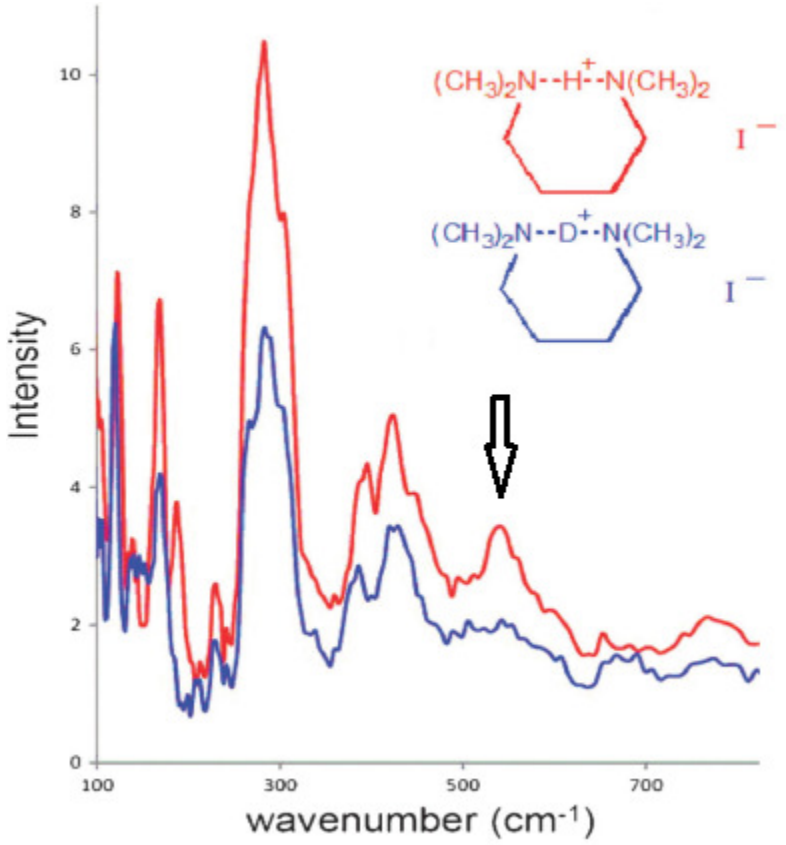


Figure 1.9[18]: INS spectra of protonated TMP (red) plotted against deuterated TMP (blue) in the domain from 100 – 900 cm^{-1}. Two bands in the protonated spectrum (190 and 525 cm^{-1}) shift when compared to the deuterated spectrum. These two bands are assigned to be motions associated with the bridging proton.

Figure 1.10 reproduces the gas phase vibrational spectra of the protonated and deuteronated TMP in the mid-IR domain from 550 – 1800 cm^{-1}. Having a bridging deuteron in place of the bridging proton shows an additional band disappearance at 1200 cm^{-1}. This band was assigned to be the N-$H^{+\cdots}$N in-plane and out-of-plane bending modes. Calculations predicted the N-$H^{+\cdots}$N bending modes to appear around 1600 cm^{-1}, but the small shoulder at 1600 cm^{-1} in the experimental spectrum does not disappear or shift upon deuteration.

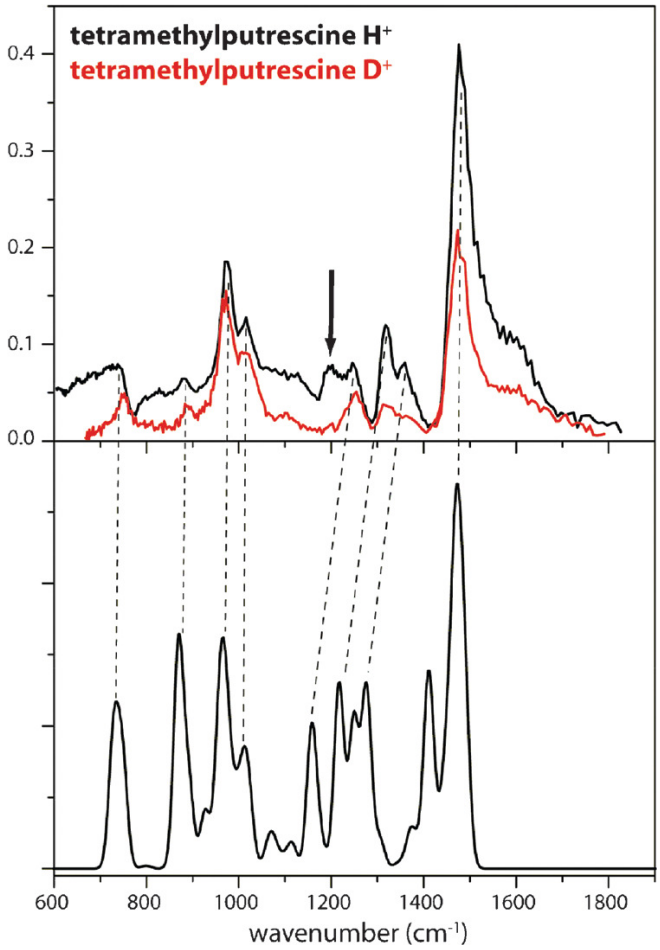


Figure 1.10[6]: The top panel reproduces the experimental IRMPD spectra of protonated (black) and deuterated (red) TMP in the domain from 600 – 1800 cm^{-1}. A band at 1200 cm^{-1} disappears upon deuteration and is assigned to be a $N\text{-}H^{+}\cdots N$ bending mode. The bottom panel shows the DFT predicted IR spectrum calculated at the B3LYP/6-31G** level.

Protonated TMP iodide salt was examined using X-ray diffraction and ssNMR. X-ray diffraction of protonated TMP iodide salt places the N-N distance at 2.655 Å, which is in good agreement with predictions[21]. The bridging N-H distance was determined using $^{15}N\text{-}^{1}H$ dipolar coupling experiments and gave an average N-H distance of 1.324, halfway between the two dimethylamino groups.

All of the diamines investigated up to this point have significantly less rigidity compared to DMAN. Another diamine (DACO) was examined by a previous group member[17] and DACO differs from the linear chain diamines because the dimethylamino groups rest on a cyclooctane ring. This provides more rigidity compared to linear chain diamines, but still not as rigid as DMAN. **Figure 1.11** depicts the experimental IRMPD spectrum of protonated DACO investigated in the fingerprint domain from 600 – 1800 cm^{-1}. The protonated IRMPD spectrum is reproduced in black and the deuterated IRMPD spectrum is shown in red. Overlap between the two IR spectra shows a band at 1650 cm^{-1} (black arrow) that disappears upon deuteration. This band is assigned to be a N-$H^{+}\cdots$N bending mode. Calculations at the B3LYP/6-31G** level place the unscaled bending modes at 1635 and 1648 cm^{-1}.

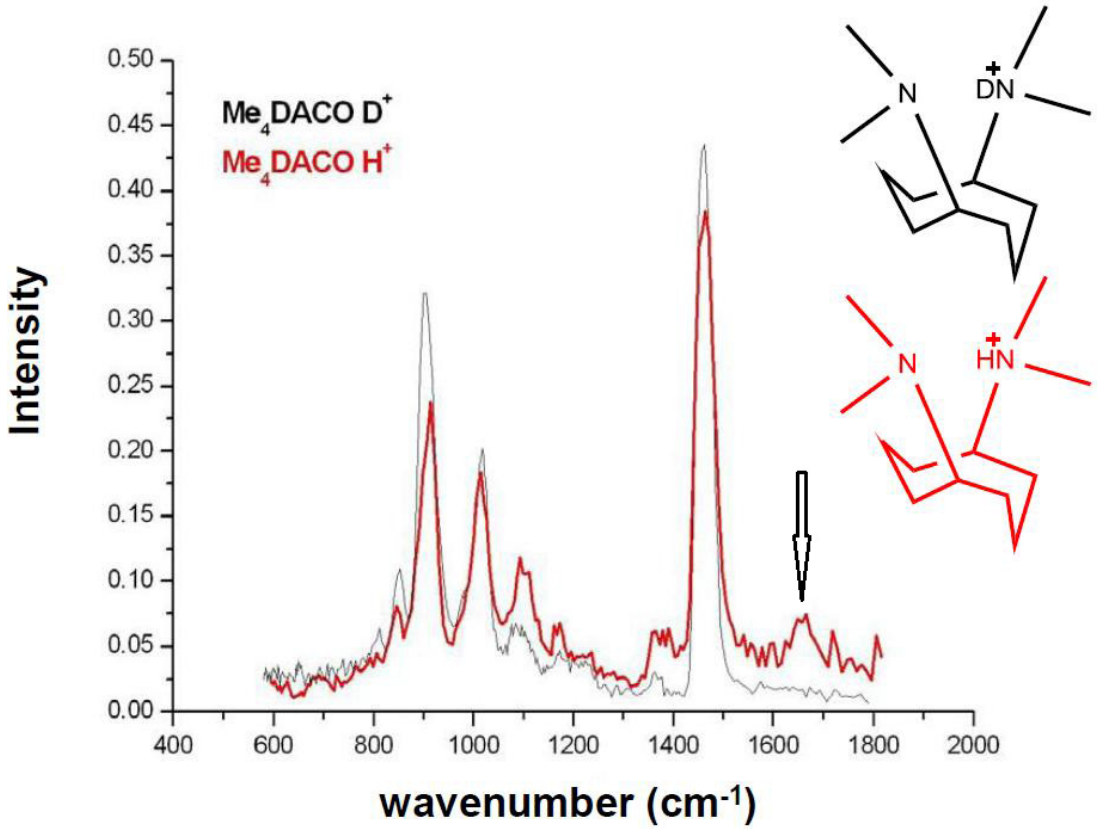


Figure 1.11[18]: Experimental IRMPD spectra of protonated DACO (black) plotted against deuterated DACO (red) in the domain from 600 – 1800 cm^{-1}. Upon isotopic replacement of the bridging proton with deuterium, a band around 1650 cm^{-1} disappears, and is assigned to be an $N\text{-}H^{+\cdots}N$ bending mode.

This chapter will present infrared data on rigid tri- and tetraamines, capable of forming intramolecular hydrogen bonds upon protonation. Since protonated TMP was determined to be a LBHB molecule, this chapter will examine tri- and tetraamines that resemble TMP. All of the diamines mentioned up to this point were capable of acquiring only one proton. Two molecules, tetraamine (**3**) and triamino alcohol (**4**) are capable of forming diprotonated

dications upon acquisition of a second proton. Before molecules **3** and **4** can be discussed, doubly charged ions warrant some discussion.

Multiply charged ions exhibit the same characteristics as singly charged ions, but their thermochemistry is altered due to the internal electrostatic repulsion of the identical charges[22]. Small ions bearing multiple charges are unstable due to the close proximity of like charges. As the size of a molecule decreases, the charge repulsion between the two like charges increases and may cause the ion to dissociate. The study of multiple charged ions is of interest because of the crucial role they play in chemistry and biochemistry[23]. For example, doubly charged ions are used to detect smaller ions that are otherwise undetectable. Armstrong *et al.* observed the binding of iodide and other anions to a doubly-charged cation which would otherwise be difficult to detect[24]. Symmetric and unsymmetric dications were investigated and their results showed that unsymmetrical dications lowered the detection limit for anions.

Doubly charged ions formed from a single molecule can either carry two of the same charge, or two opposite charges. An example of a doubly charged ion caring the same charge is a sulfate dianion[25]. The number of water molecules in a cluster ranges from 5 to 16, with 12 being the most abundant. Without the aid of water, the dianion was undetectable. Another example of a doubly charged ion bearing the same charge is succinylcholine. Succinylcholine possesses two quaternary amine groups which bears the positive charge. Succinylcholine has a *m/z* 145, but is not a result of protonation of a neutral compound[26].

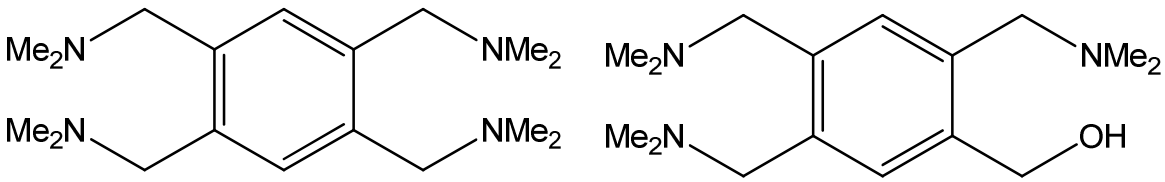
Betaine is another example of a molecule carrying two opposite charges in its natural state. O'Hair *et al.* found that as betaine was ionized by ESI, it formed clusters containing multiple charges. The smallest of the doubly charged clusters contained 13 betaine zwitterions and has two protons attached to the cluster. In order for these clusters to be stable and organized in the gas phase, dipole-dipole interactions and hydrogen bonding must compensate for the coulombic repulsion of the two cations[23]. Although the diprotonated betaine dimer was never detected, DFT calculations suggest that it can exist in the gas phase[23]. The doubly charged betaine dimer could exist in a hydrogen bond donor-acceptor motif commonly observed between carboxylic acids. The dissociation of a diprotonated betaine dimer into two monoprotonated betaines is predicted to be an exothermic process, releasing approximately 26 kcal/mol[23].

Recently, it was discovered that 1,12-diaminododecane is capable of forming a diprotonated dication. An ESI-MS of showed the presence of protonated 1,12-diaminododecane (*m/z* 201.2325), as well as diprotonated 1,12-diaminododecane (*m/z* 101.1184). The intensity of the diprotonated dication is approximately one-fifth the intensity of the protonated parent ion. The base peak is not that of the protonated parent ion, but loss of ammonia from the parent ion (*m/z* 184.2061). The loss of ammonia can occur from an internal S_N2 nucleophilic displacement, which results in a 13-membered ring. Having two charges as far away as possible on a linear 12 carbon chain help stabilize charge repulsion.

Synthesis and Characterization / Methods

General

DFT calculations are utilized to predict the most stable conformation as well as predict gas phase infrared spectra. Comparison between predicted and experimental infrared spectra provides a clue to how ions interact in a solvent-free environment.



3 **4**

Figure 1.12: Aromatic polyamines **3** and **4** that were selected for investigation due to their resembalance to tetramethylputrescine.

IRMPD spectra were obtained at the FOM Institute for Plasma Physics Rijnhuizen Nieuwegein, The Netherlands. Ions were generated *via* ESI and isolated in a FT-ICR cell for several seconds. ESI makes it possible to study molecules with high charge densities[22]. The ions are then mass selected and subsequently irradiated for several seconds with pulses generated from a tunable or a fixed frequency laser. The fingerprint domain from 300 – 1800 cm^{-1} utilized the free electron laser for infrared experiments (FELIX) as a tunable IR source. A Nd:YAG laser pumped Optical Parametic Oscillaor tunable laser was utilized for

investigating the IR domain from 2350 – 3900 cm^{-1}. To increase the dissociation ion yield, ions were irradiated for several milliseconds using a fixed frequency CO_2 laser after each OPO pulse. Experimental IRMPD spectra were obtained *via* action spectroscopy where both the parent ion abundance and the dissociated daughter ion dissociated daughter ions are plotted against the scanned frequency[27-29]. A schematic representation of the FELIX interface with an ICR is reproduced in **Figure 1.13**.

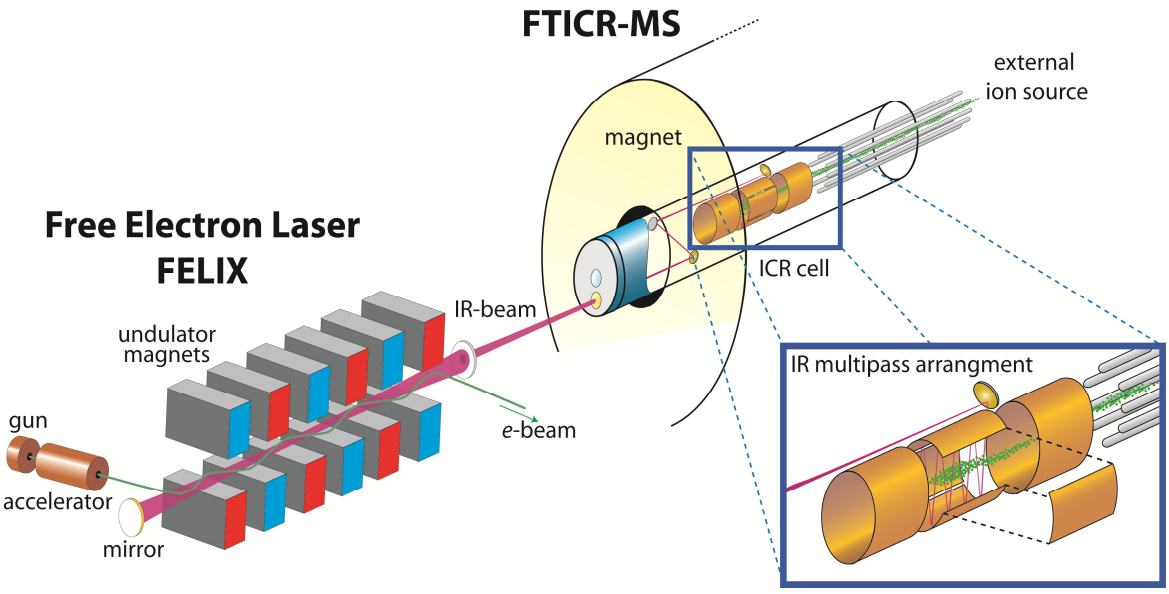
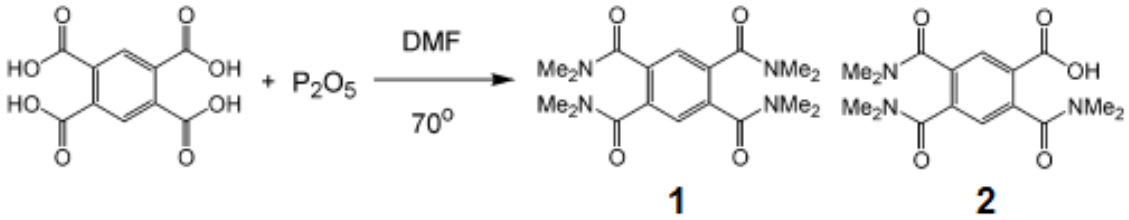


Figure 1.13[27]: A schematic representation of FELIX. A tunable free electron laser is used to obtained IRMPD spectra of gaseous ions in the gas phase, whether it is in the fingerprint or the CH/NH/OH stretching domain.

Tandem MS/MS experiments were performed using a linear triple quadrupole mass spectrometer at 4 and 10V lab frame energy with Ar as a collision gas at the Faculté de Pharmacie, Universite Paris-Sud XI, Châtenay-

Malabry. The linear quadrupole consists of three quadrupoles: Q1, q2, and Q3. Q1 uses alternating AC and DC electric fields to isolate an ion with a specific mass-to-charge ratio. The ions then move to q2 where only rf is used. The isolated ions collide with neutral collision gas in q2. The fragment then passes to the third quadrupole where all fragmented ions are scanned. MS/MS experiments can be performed by isolating the daughter ion and undergoing further dissociation with neutral collision gas.

N1,N1,N2,N2,N4,N4,N5,N5-octamethylbenzene-1,2,4,5-tetracarboxamide (1)



Scheme 1.1: Synthesis of N1,N1,N2,N2,N4,N4,N5,N5-octamethylbenzene-1,2,4,5 tetracarboxamide (**1**) starting from pyromellitic acid. Incomplete conversion of the carboxylic acid to the amide resulted in a byproduct, 2,4,5-tris(dimethylcarbamoyl)benzoic acid (**2**).

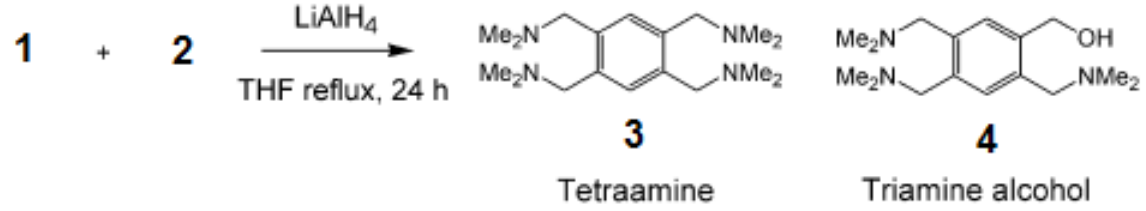
N1,N1,N2,N2,N4,N4,N5,N5-octamethylbenzene-1,2,4,5-tetracarboxamide (**1**) was synthesized by a previous graduate student, Sevana Khodagholian, as previously described in literature as shown in **Scheme 1.1** [30,31].

To 2.8 grams (0.019 mol) of P_2O_5 in 100 ml N,N-dimethylformamide in a 250 mL round bottom flask equipped with a stir bar was added 10 grams (0.039 mol) of pyromellitic acid. The resulting solution is heated for 12 hours at 65°C

with a reflux condenser attached. After cooling to room temperature, the yellow solid was filtered and recrystallized using absolute ethanol. A yield of 1.8 grams (13%) was obtained. Incomplete conversion of the carboxylic acids to amides produced 2,4,5-tris(dimethylcarbamoyl)benzoic acid (**2**).

APCI-ESI MS: MH^+: 363.2061 (calc. 363.2026)

Tetrakis(dimethylaminomethyl)benzene (3)



Scheme 1.2: Synthesis of tetrakis(dimethylaminomethy)benzene (**3**) starting from a crude sample containing both N1,N1,N2,N2,N4,N4,N5,N5-octamethylbenzene-1,2,4,5 tetracarboxamide (**1**) and 2,4,5-tris(dimethylcarbamoyl)benzoic acid (**2**). Reduction of 2,4,5-tris(dimethylcarbamoyl)benzoic acid (**2**) using $LiAlH_4$ produced (2,4,5-tris((dimethylamino)methyl)phenyl)methanol (**4**).

Tetrakis(dimethylaminomethy)benzene (**3**) was synthesized from **1** upon reduction using $LiAlH_4$, as shown in **Scheme 1.2**. A 50 mL round bottom flask was charged with 60 mg lithium aluminum hydride (0.0016 mol) in 25 mL dry THF and equipped with a stir bar to which was added 200 mg of **1** drop wise in THF dropwise at 0^{o}C. The reaction mixture was allowed to warm to room temperature before heating at 70^{o}C for 12 hours with a reflux condenser attached.

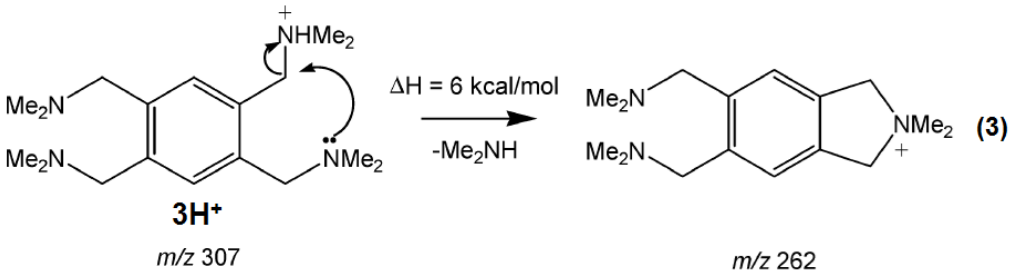
After 12 hours of stirring and heating, the reaction was cooled to room temperature. While stirring, 60 μL of DI H_2O, 60 μL 15% KOH (aq), and 180 μL DI H_2O was added dropwise. The yellow solution was allowed to stir at room temperature for an additional 30 minutes. The inorganic salts were then removed by vacuum filtration and the filtrate treated with 30 mL of concentrated aqueous KOH and transferred to a separatory funnel. The upper (organic) layer was separated**,** and the lower (aqueous) layer was extracted with two 10 mL portions of THF. The combined organic layers were dried over 2 grams of anhydrous sodium sulfate. Sodium sulfate was then removed by gravity filtration and THF removed under reduced pressure.

APCI-ESI MS: MH^+: 307.2855 (calc. 307.2855)

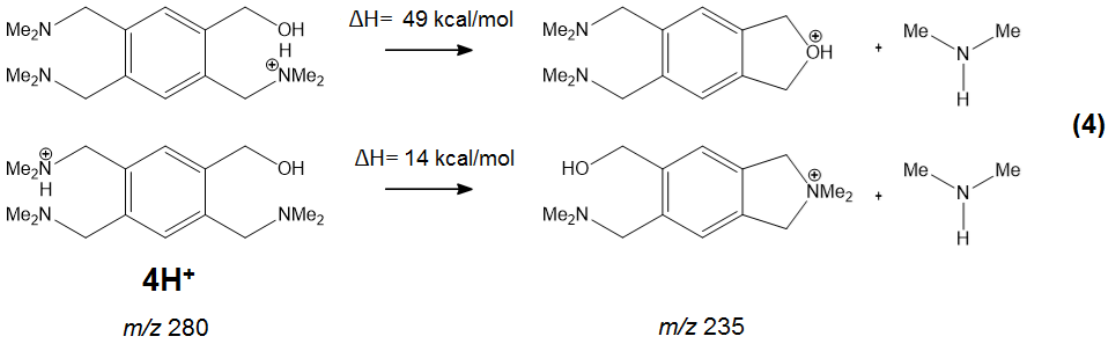
APCI-ESI MS: MH_2^{+2}: 154.1483 (calc. 154.1463)

Theory / Calculation

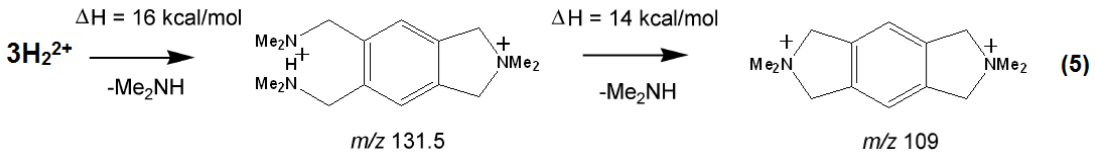
Geometry optimization and harmonic vibrational frequencies were calculated for protonated and diprotonated **3** and **4** using the Gaussian 09 program suite at the B3LYP/6-31G** and B3LYP/6-311++G** levels. DFT calculations provide a clue of the thermochemistry and different potential decomposition pathways. Monoprotonated cations **3** and **4** can eliminate neutral dimethylamine, but first must rotate about the C-C bond, breaking the ionic hydrogen bond in the process as depicted in **Eq 3**.



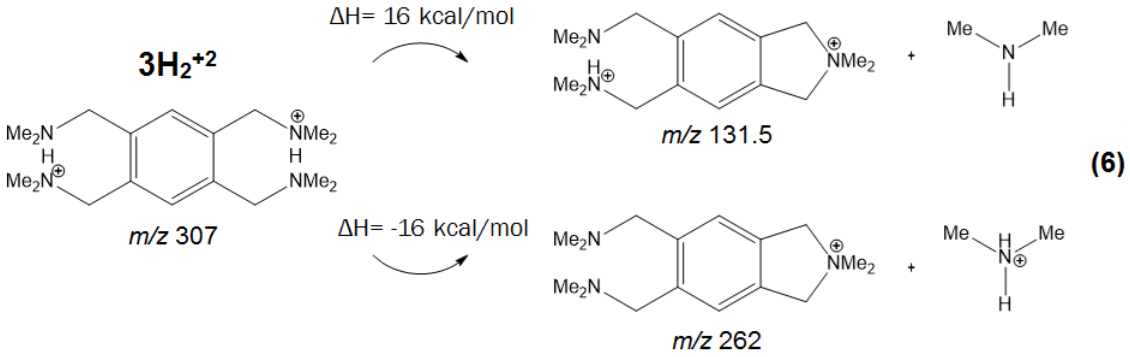
Monoprotonated cation **4** can also expel two separate molecules of dimethylamine, depending on the nucleophile. The reaction sequence is outlined in **Eq 4**. All processes are calculated to be endothermic if the number of electric charges in the fragment ion is the same as in the parent ion.



The dissociation of the diprotonated dication **3** can follow two pathways. The first pathway eliminates neutral dimethylamine, leaving behind a doubly charged daughter ion as shown in **Eq 5**.

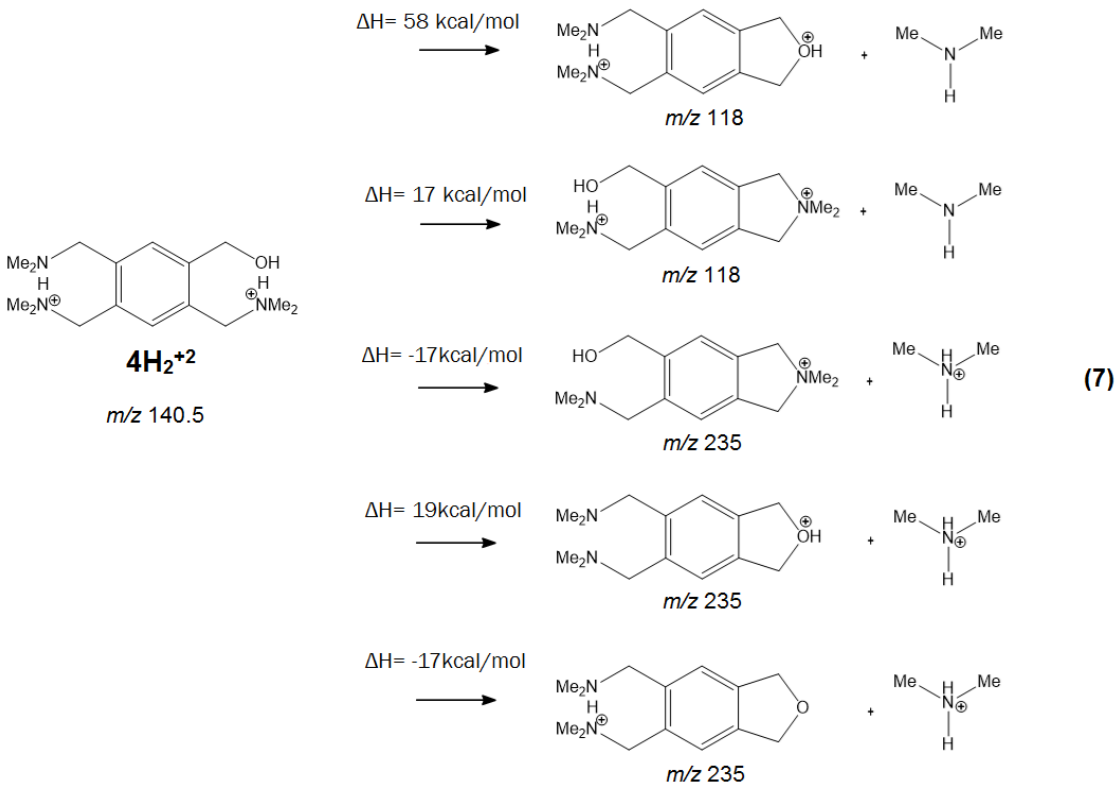


The dissociation process is calculated to be endothermic, because the daughter ion has to bear two like charges. If the expelled neutral dimethylamine can shuttle to the other side of the benzene ring and acquire the bridging proton, that process would produce two singly charged daughter ions. Having the two charges distributed between two daughter ions is calculated to be the most thermodynamically stable process with a ΔH = -16 kcal/mol, as shown in **Eq 6**.



(6)

Collisional dissociation of diprotonated **4** can follow many pathways depending on the actions of the expelled dimethylamine. **Eq 7** depicts the possible outcomes. Having two charges on one molecule is predicted to be unfavorable, especially if the oxygen is bearing one of the charges. Division of charges between two molecules is predicted to be the most favorable pathway, except in the case where a protonated oxygen is produced. The two most favorable dissociation pathways of ion $\mathbf{4H_2}^{+2}$ both produced a protonated dimethylamine and has the nitrogen bearing the positive charge.



Results

ESI and Linear Triple Quadrupole MS

Diamines with both equal and unequal proton affinities have been previously investigated to determine if they are capable of forming LBHBs [15-17]. The same criteria used for determining LBHBs can be applied to ions **3**, **4** and the daughter ions produced from fragmentation of $\mathbf{3H_2^{+2}}$. **Figure 1.12** depicts the polyamines **3** and **4,** which were selected for study due to their strong resemblance to tetramethylputrescine on both sides of the benzene ring, separated by two sp^2 hybridized carbons.

The gas phase IR spectra between 600 – 1800 cm^{-1} of mono- and diprotonated polyamines were examined. The APCI-ESI mass spectrum (MS) depicting ions **3** and **4** are reproduced in **Figure 1.14**. The MS shows not only the monoprotonated cations, but also diprotonated dications. The ^{13}C natural abundance peaks of the monoprotonated ions are separated by 1 amu from the much more intense all-^{12}C peaks. In the case of the diprotonated dications, the ^{13}C natural abundance peak is separated by ½ amu because a mass spectrometer plots the mass-to-charge ratio (and the electrical charge is 2).

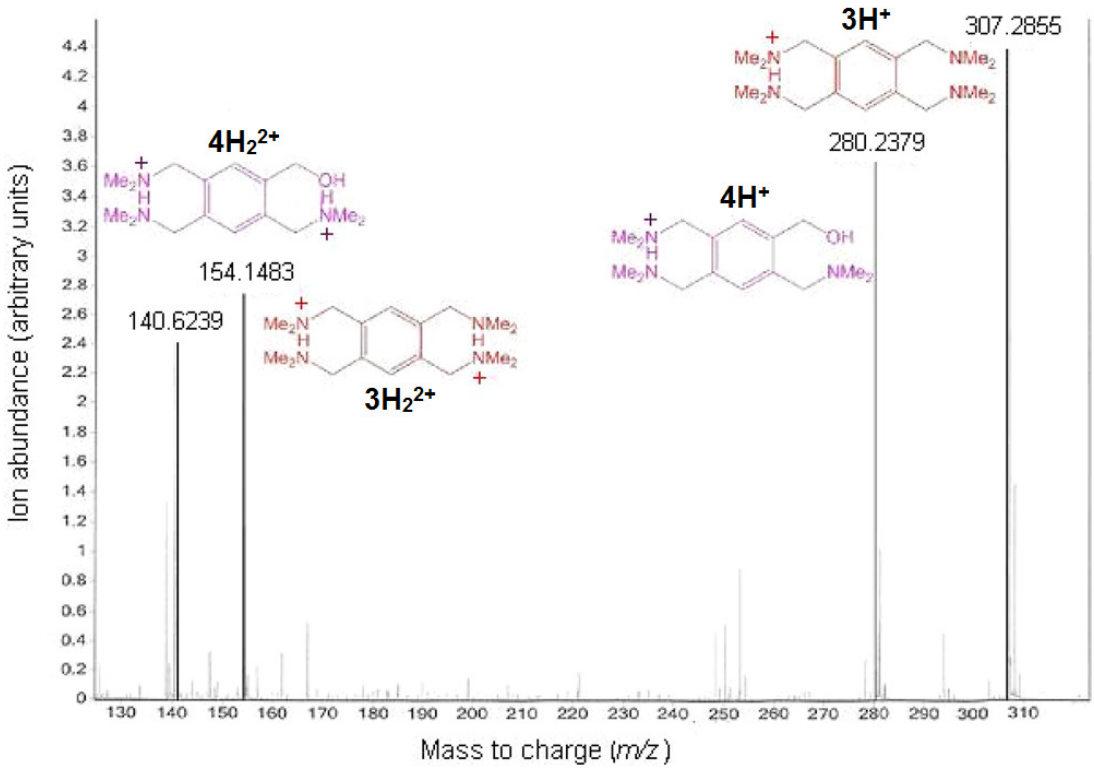
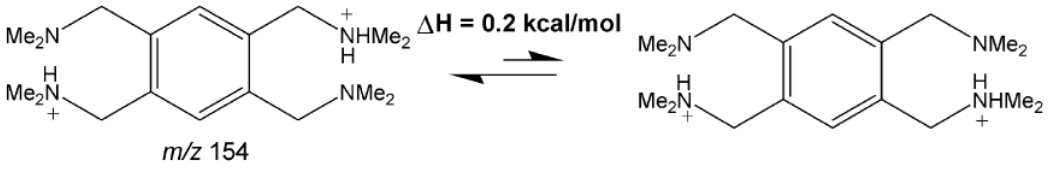


Figure 1.14: APCI-ESI MS depicting protonated and diprotonated cations of *tetrakis*(dimethylaminomethy)benzene (**3**) and (2,4,5-*tris*[(dimethylamino)methyl)phenyl]methanol (**4**).

3H$^+$ has a calculated N-H$^{+\cdots}$N bond angle of 165^o, 5^o farther away from linearity than protonated tetramethylputrescine. All the amino groups on this molecule have equal proton affinity. In the case of **3H$_2^{+2}$**, the N-H$^{+\cdots}$N bond angle on one side of the aromatic ring differs slightly from the other. The calculated N-H$^{+\cdots}$N bond angle on one side is identical to that of its protonated analog at 165^o. The other N-H$^{+\cdots}$N bond angle deviates more from linearity, having an N-H$^{+\cdots}$N bond angle of 162^o. It is possible for **3H$_2^{+2}$** to exist as two tautomers depending on the location of the bridging protons. Placing the charges at the farthest point possible was calculated to be more stable by 0.2 kcal/mol, as shown in **Eq. 8**. Having the proton transfer back and forth between the nitrogens on one side of the benzene will affect the basicity of the nitrogens on the other side of the ring.

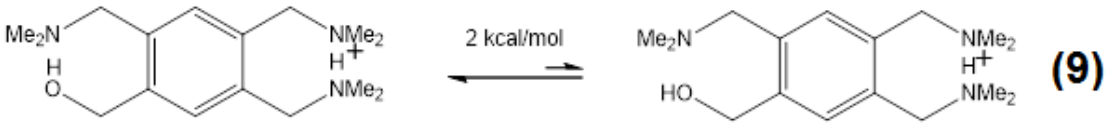


m/z 154

(8)

Tandem MS/MS experiments using a linear triple quadrupole mass spectrometer at 4 and 10 V lab frame energy of **3H$_2^{+2}$** is shown in **Figure 1.15**. The doubly charged fragment ion at *m/z* 131.9 corresponds to a daughter ion produced from the expulsion of neutral dimethyl dimethylamine. If the expelled neutral dimethylamine can shuttle to the other side of the benzene via an ion-neutral complex, it can acquire the bridging proton and produce a resulting

daughter ion *m/z* 262.5. This process contributes less much less than 1% of the fragment ions. The doubly charged daughter ion at *m/z* 109.4 corresponds to a second daughter ion from expulsion of two molecules of neutral dimethylamine, which has approximately 70% of the intensity of the expulsion of one neutral dimethylamine. Clearly the exothermic decomposition is much less probable at this collision energy then the endothermic decompositions.



(9)

Ion **4H$^+$** can exist as two tautomers as shown in in **Eq 9**. The more stable of the two tautomers forms an intramolecular hydrogen bond between the two dimethylamino groups and also forms an intramolecular hydrogen bond between the alcohol and the dimethylamino group on the other side of the benzene ring. The amount of energy required to break the intramolecular hydrogen bond between the alcohol and the dimethylamino group was calculated to be 2 kcal/mol.

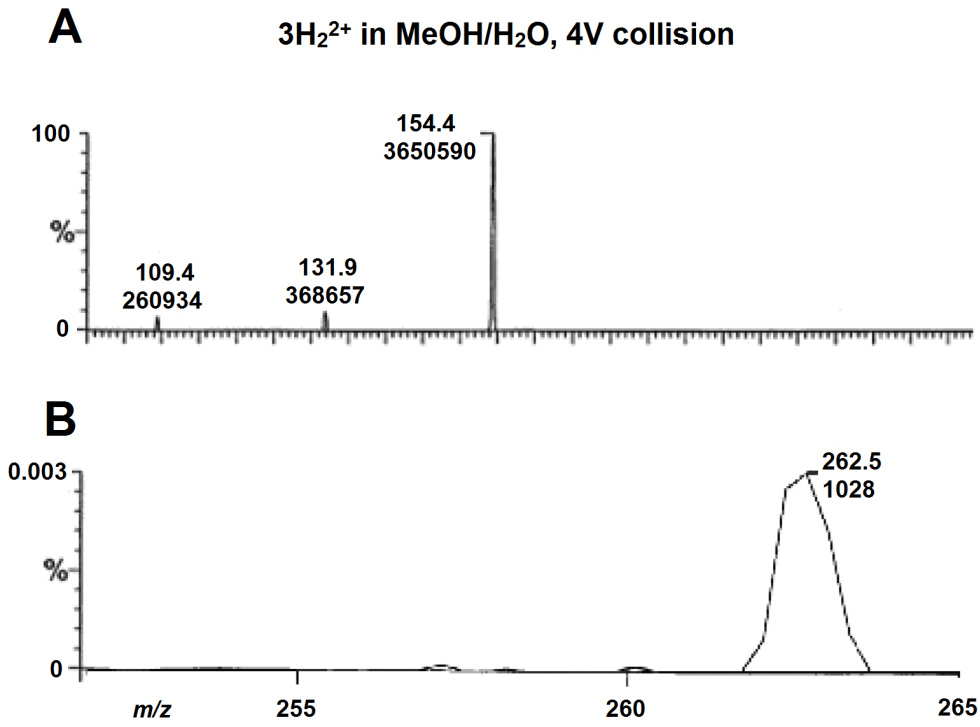


Figure 1.15: Linear tandem MS/MS experiments using a lienear triple quadrupole mass spectrometer at 4 V lab frame energy. **A)** Diprotonated dication **3** (*m/z* 154.4) dissociated upon collision with Ar gas to produce two doubly charged daughter ions (*m/z* 131.9 and 109.4). **B)** Diprotonated dication **3** (*m/z* 154.4) dissocated upon collision with Ar gas producing a singly charged daughter ion (*m/z* 262.2).

Incomplete conversion of pyromellitic acid to **1** follow by reduction using $LiAlH_4$ leads to different dimethylamino alcohol variants shown in **Figure 1.16**.

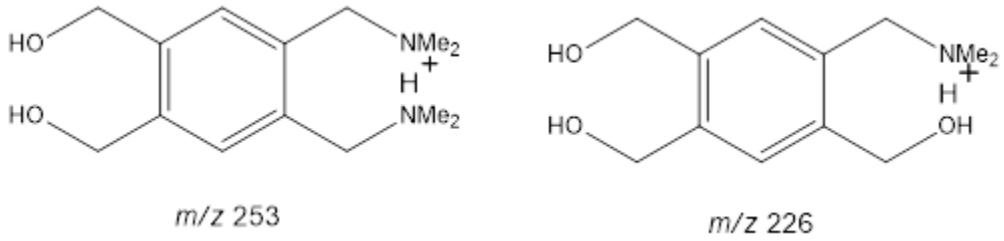


Figure 1.16: Incomplete conversion of molecules **1,2** to **3,4** results in two amine alcohol derivatives with *m/z* 253, which contains two diamino and two alcohol groups, and *m/z* 226, which contains one amino and three alcohol groups.

Figure 1.17 shows the mass spectrum containing both diaminodiol (blue) and aminotriol (red) while the tetraalcohol variant was not observed. The crude reaction mixture also contains other impurities that were not identified.

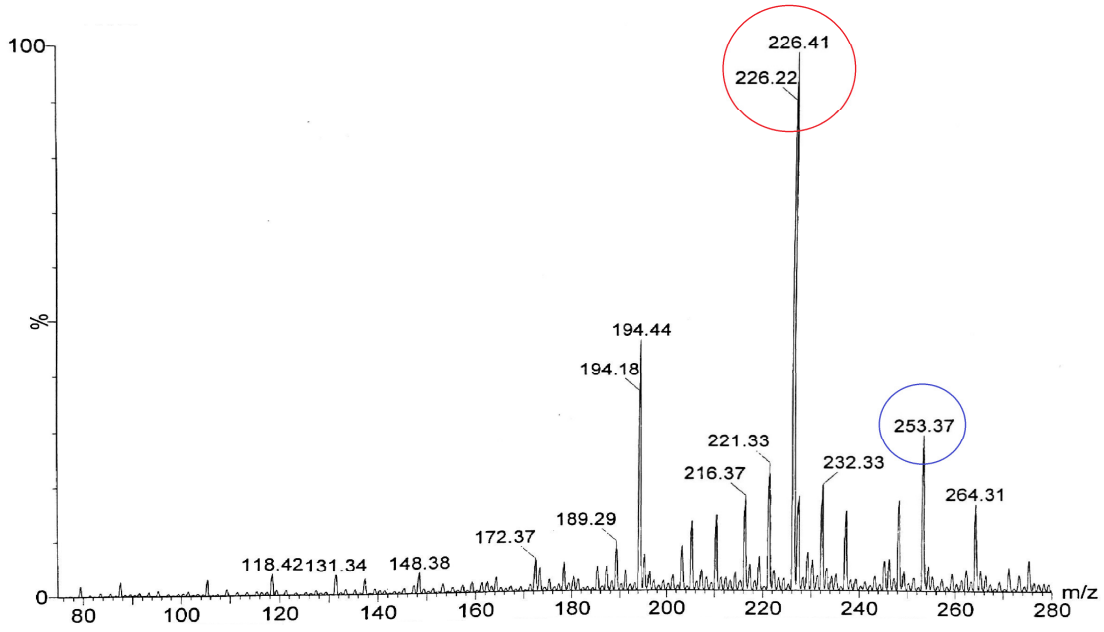


Figure 1.17: ESI mass spectrum of the incomplete conversion of pyromellitic acid to **1**, followed by reduction using $LiAlH_4$, resulting in diaminodiol (*m/z* 253) and aminotriol (*m/z* 226).

IRMPD

Similar to tandem MS/MS experiments, IR irradiation of a doubly charged ion primarily expels two molecules of neutral dimethylamine to produce doubly charged daughter ions. The IRMPD spectrum of **3H$^+$** is plotted against **3D$^+$** in **Figure 1.18**. Calculated bond angles of mono- and diprotonated **3** and **4** along with N-H$^{+\cdots}$N bond distances are outlined in **Table 1.2**.

Table 1.2: Calculated N-H$^{+\cdots}$N bond angle and distances of mono- and di-protonated ions **3** and **4** at the B3LYP/6-311++G** level.

Ions	NHN (θ)	N-H$^{+\cdots}$N (Å)	N-O (Å)	N-H$^+$ (Å)	H$^{+\cdots}$N (Å)	H$^{+\cdots}$O (Å)
3H$^+$	165^o	2.75	-	1.08	1.68	-
3H$_2^{+2}$	165^o, 162^o	2.74, 2.69	-	1.08, 1.09	1.67, 1.63	-
4H$^+$	165^o	2.74	-	1.09	1.67	-
4H$_2^{+2}$	165^o	2.74	2.71	1.09, 1.04	-	1.74

The experimental IRMPD spectra of the monoprotonated and monodeuterated cations of **3** look almost identical, except for the shoulder near 1510 cm^{-1}. DFT calculations predict the N-H$^{+\cdots}$N bending modes to occur around 1600 cm^{-1}. Since this shoulder disappears upon deuteration, the shoulder near 1510 cm^{-1} could correspond to a bending mode associated with the bridging proton. The fingerprint domain below 750 cm^{-1} was not investigated, but one

would expect to find the N-H^{+}···N asymmetric stretch in that domain due to the similarity between **3** and protonated TMP.

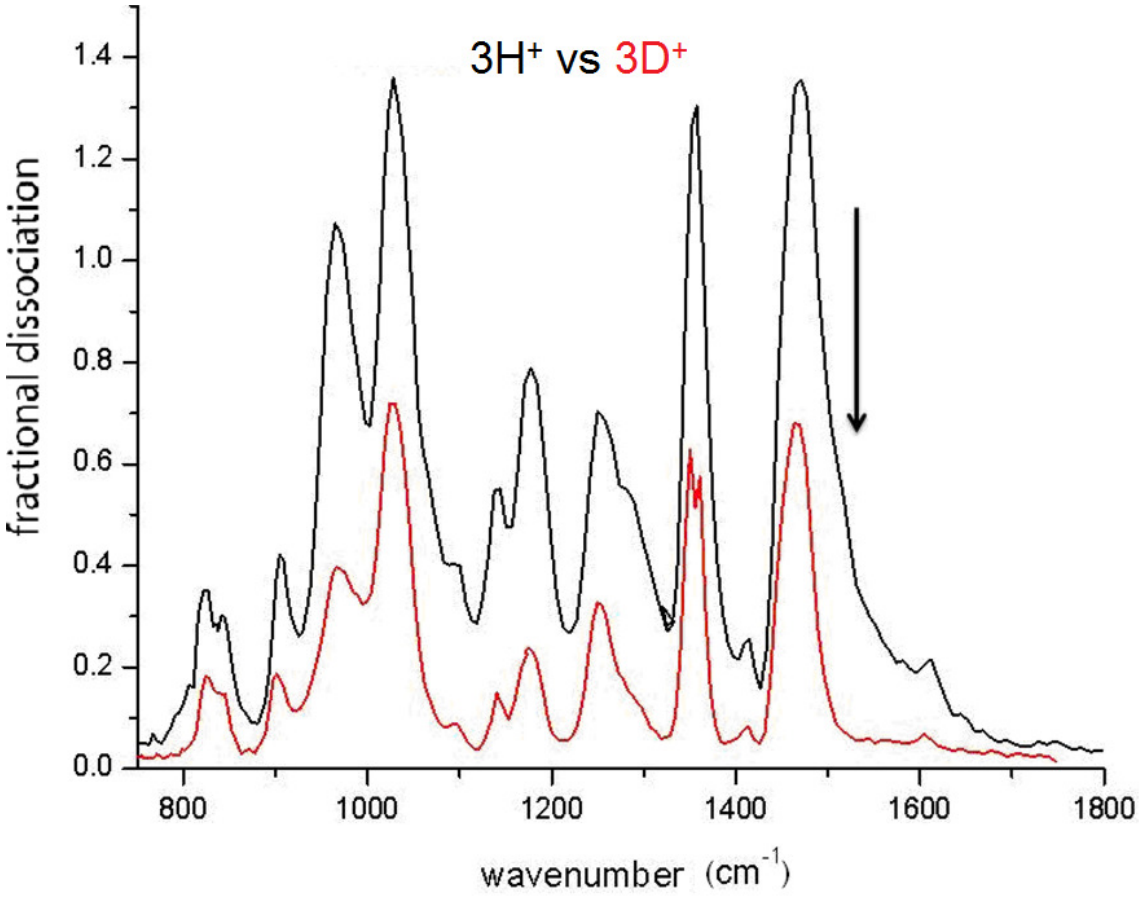
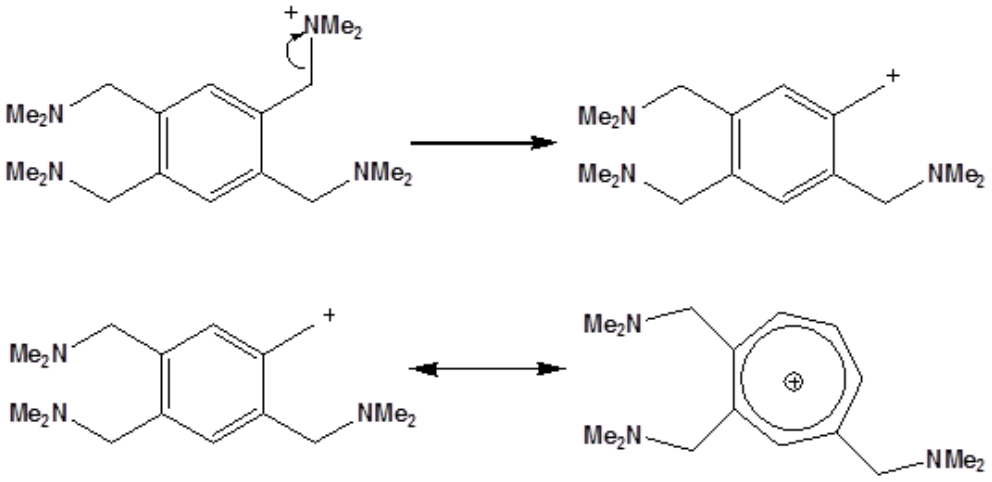


Figure 1.18: Experimental IRMPD spectra of monoprotonated cation **3** (black) plotted against monodeuterated cation **3** (red) in the mid-IR domain from 750 – 1800 cm^{-1}. The black arrow indicates a shoulder that is observed in the monoprotonated spectrum around 1510 cm^{-1} that is not present in the monodeuterated spectrum.

After the acquisition of a proton, it is possible for **3H^{+}** to lose neutral dimethylamine and undergo rearrangement to produce a tropylium ion, as depicted in **Eq 10**.



(10)

The process of forming the tropylium ion is an endothermic process with a calculated ΔH = 62 kcal/mol to form a doubly charged tropylium ion and a ΔH = 39 kcal/mol to form the singly charged tropylium ion.

Figure 1.19 depicts the experimental vs calculated IRMPD plots of two doubly charged ions obtained from the loss of neutral dimethylamine. The calculated spectrum on the left corresponds to the end result of internal nucleophilic attack to displace neutral dimethylamine of doubly charged ion **3**, while the doubly charged troplyium ion on the right results from the loss of neutral dimethylamine and undergoing rearrangement. Even though the doubly charged tropylium is calculated to have a high heat of formation, it is worth examining. The experimental IRMPD spectrum (the same in both cases) is plotted against DFT-predicted IR in **Figure 1.19**. The comparison is done in an attempt to see which ion is observed in the gas phase. It is uncertain whether the parent ion contains only one structure. The doubly charged ion on the left has two bands at 1050 and 1150 cm^{-1} in the calculated spectrum that are not observed

experimentally. The doubly charged tropylium ion on the right has an intense band at 1530 cm^{-1} that is not seen in the experimental spectrum. On balance, though, it seems that the dozen most intense bands of the unrearranged ion provide a better match to the experimental IRMPD spectrum.

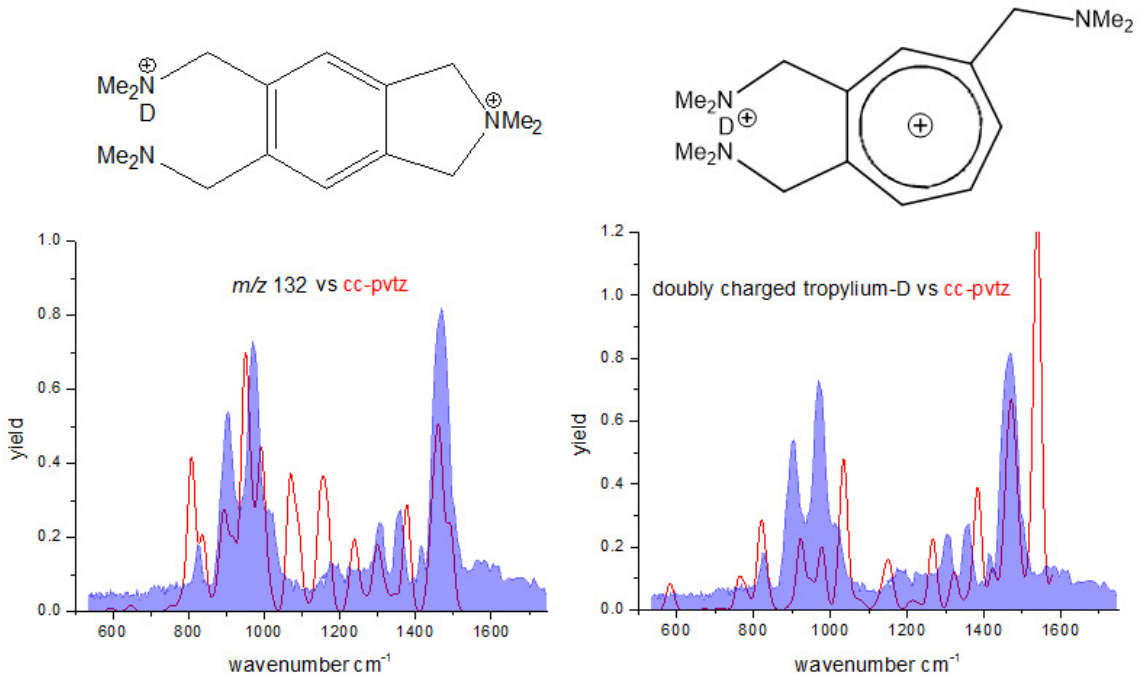


Figure 1.19: Experimental IRMPD spectra (blue silhouette) of a doubly charged ion with *m/z* 132 plotted against two different isomers calculated by DFT (red). The doubly charged ion on the left resulted from an internal nucleophilic displacement, losing neutral dimethylamine, and the doubly charged ion on the right resulted from loss of neutral dimethylamine with rearrangement. The doubly charged ion on the right was calculated to have a much higher heat of formation.

Figure 1.20 depicts the IRMPD spectra of $\mathbf{3H_2^{+2}}$ plotted against $\mathbf{3D_2^{+2}}$ in the mid-IR domain. These spectra are vastly different than that of $\mathbf{3H^{+}}$. Isotopic replacement of the two protons with deuterons shifts at least six bands (black arrows) between 1100 and 1250 cm^{-1}. Because these bands disappear upon

deuteration, they are assigned to be motions associated with the bridging protons. Two of the bands could correspond to the N-$H^{+\cdots}$N asymmetric stretches, while the other four bands could correspond to the in-plane and out-of-plane bending modes. These bands all disappear in the domain similar to that of *N,N,N',N'*-tetramethylpropanediamine, which was thought to not behave like a LBHB molecule. As previously mentioned, tetramethylpropanediamine has an N-$N^{+\cdots}$N bond angle of 155^{o} and had three bands that shift in that similar domain near 1200 cm^{-1}, 1300 cm^{-1}, and 1350 cm^{-1}. Tetramethylpropanediamine has an N-$H^{+\cdots}$N bond angle that deviates too far from linearity and therefore does not exhibit the characteristics of a linear LBHB.

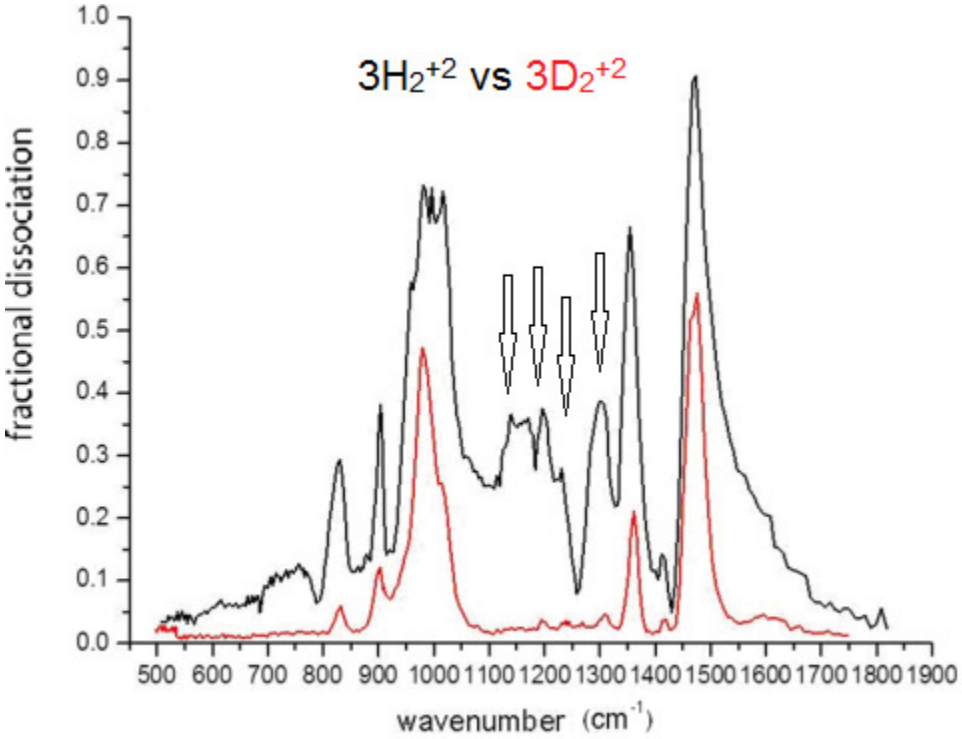


Figure 1.20: Experimental IRMPD spectra of diprotonated dication **3** (black) plotted against dideuterated dication **3** (red) in the mid-IR domain from 500 – 1800 cm^{-1}. At least six bands seem to disappear upon isotopic replacement. Those bands could account for two asymmetric stretches and four in-plane and out-of-plane N-$H^{+\cdots}$N bending modes.

Figure 1.21 reproduces the IRMPD spectra of a doubly charged daughter ion (*m/z* 131.5) resulting from the loss of neutral dimethylamine. The red trace is the same silhouette as that of **Figure 1.19**. In this case, the expelled dimethylamine did not shuttle to the other side of the benzene ring to acquire the bridging proton. DFT predicts the amino groups of the singly charged analog (without the bridging proton) to have equal basicity. The N-$H^{+\cdots}$N bond angle of *m/z* 131.5 is calculated to be 161°. Upon isotopic replacement with a deuteron, a broad band near 1175 cm^{-1} disappears. This band disappearance occurs near

the same domain of that of the in-plane and out-of-plane bending modes of TMP. Although this band cannot be definitively assigned, it is reasonable to assume this may correspond to the in-plane and out-of-plane bending modes. The fingerprint domain below 800 cm^{-1} was not investigated, but one would expect to find the $N\text{-}H^{+}\cdots N$ asymmetric stretch in that domain.

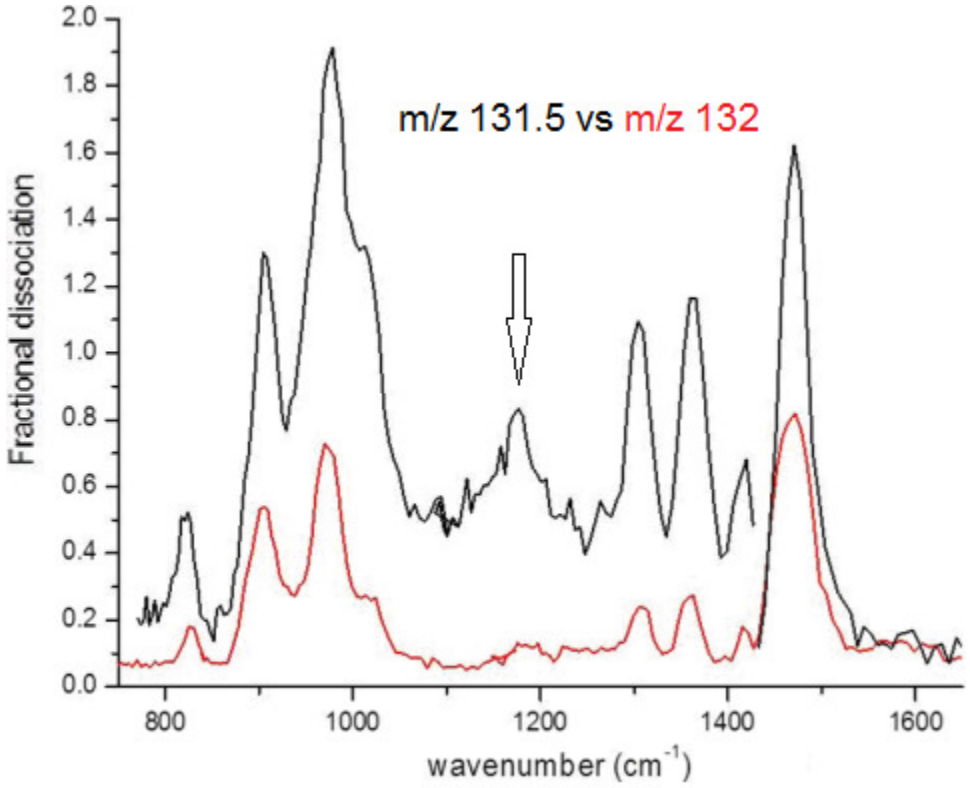


Figure 1.21: Experimental IRMPD spectra of the doubly charged daughter ion *m/z* 131.5 (black) plotted against the doubly charged daughter ion *m/z* 132 with a bridging deuterium (red) in the fingerprint domain from 800 – 1600 cm^{-1}. One broad band seems to disappear upon isotopic replacement and that band could corresponding to the in-plane and out-of-plane $N\text{-}H^{+}\cdots N$ bending modes.

Figure 1.22 shows the experimental IRMPD spectrum of **$4H^{+}$** in the CH/NH/OH stretching domain from 2800 – 3800 cm^{-1}. The IR spectrum shows

one band around 3640 cm^{-1}. The C-H and N-H stretches are not observed in this spectrum, but the O-H stretch is observed around 3640 cm^{-1}. This spectrum demonstrates protonation of this molecule does not affect the OH stretching frequency.

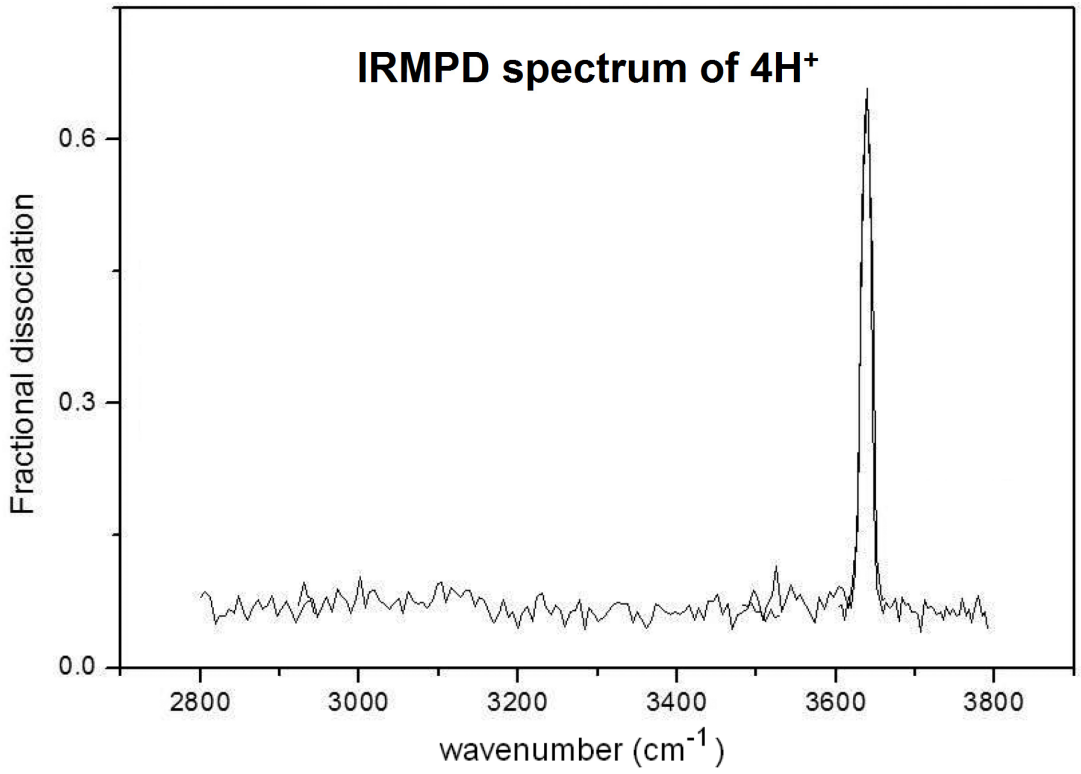


Figure 1.22: Experimental IRMPD spectrum of **$4H^+$** in the CH/NH/OH stretching domain from 2800 – 3800 cm^{-1}. C-H and N-H stretches are not observed but the O-H stretch is around 3640 cm^{-1}.

Conclusions

In summary, the results discussed in this chapter examined intramolecular hydrogen bond formation between two dimethylamino groups in molecules that resemble TMP. The diamines selected for study all contain tertiary amines, so to

avoid unwanted hydrogen bonding interactions. Previous group members thoroughly investigated gas phase and solid infrared signature of protonated TMP. Results obtained from previous investigations were applied to the current experiment and expanded upon.

Molecule **3**, possessing four dimethylamino groups, resembles that of TMP except that **3** is attached to a benzene ring and contains two sp^2 centers at carbons #3 and #6. Molecule **3** is capable of acquiring two protons forming two intramolecular hydrogen bonds. Since TMP was determined to satisfy the criteria to be a LBHB molecule, we anticipate molecule **3** would behave similarly to protonated TMP. IRMPD spectra comparison between $\mathbf{3H^+}$ and $\mathbf{3D^+}$ looks to be almost identical except for a shoulder in the protonated spectrum, around 1510 cm^{-1}. The shoulder is assigned to be a bending mode associated with the bridging proton. IRMPD spectra comparison between $\mathbf{3H_2^{+2}}$ and $\mathbf{3D_2^{+2}}$ shows at least six bands that shift. Two of those bands could correspond to the $N\text{-}H^{+\cdots}N$ stretches, while the other four could be the in-plane and out-of-plane bending modes.

Molecule **4** turned out to be a byproduct from the synthesis of molecule **3**. One side of molecule **4** resembles TMP while the other side does not. The asymmetry of molecule **4** should render the dimethylamino groups inequivalent and incapable of forming a LBHB due to the difference in basicity of the dimethylamino groups. We would not expect to find the $N\text{-}H^{+\cdots}N$ asymmetric stretch around the 500 cm^{-1} domain due to its asymmetry.

In the gas phase, both molecules **3** and **4** were found to form both mono- and di-protonated ions. DFT calculations were employed to help predict the dissociation pathways of mono- and diprotonated ions **3** and **4**. DFT geometry optimization and frequency calculations were performed using B3LYP/6-31G** and B3LYP/6-311++G** levels. The reported ΔH values derived from using calculations at the B3LYP/6-311++G** level. DFT predicts the dissociation of ions $\mathbf{3H^+}$ and $\mathbf{4H^+}$ to one singly charged ion and one neutral daughter fragment to be an endothermic process. The predicted dissociation pathways of $\mathbf{3H_2^{+2}}$ and $\mathbf{4H_2^{+2}}$ were calculated to be endothermic if the two charges reside on one daughter ion, but exothermic if the two charges are divided amongst the two daughter ions. Despite this fact, the principal dissociation pathways give doubly charged ions.

The IR spectrum of $\mathbf{3H_2^{+2}}$ shows at least six bands that are not present in the spectrum of the corresponding deuterated analog. The six bands could account for the two N-H^+···N asymmetric stretches, and the four in-plane and out-of-plane bending modes. Because we believe the N-H^+···N asymmetric stretches are already observed, the fingerprint domain below 800 cm^{-1} would not hold any information regarding the motion of the bridging proton. The loss of neutral dimethyl amine from the internal attack of ion $\mathbf{3H_2^{+2}}$ produces a daughter ion with *m/z* 131.5. This ion is interesting because the two dimethylamino groups should be of equal basicity. Comparison between *m/z* 131.5 and *m/z* 132 shows a band that is present in the protonated spectrum but not in the deuterated spectrum at 1175 cm^{-1}. This band is assigned to be a N-H^+···N bending mode so if we were to

investigate the fingerprint domain below 800 cm^{-1}, we would expect to find the N-$H^{+\cdots}$N asymmetric stretch.

3H$^+$ and daughter ion *m/z* 131.5 both satisfy criteria to be designated LBHB molecules. Like that of protonated TMP, **3H$^+$** and *m/z* 131.5 has a band around 1200 cm^{-1}. This band is assigned to be the in-plane and out-of-plane bending modes. It is possible for the two bending modes to be close enough that they superimpose and give rise to one band.

References

1. K. Kawaguchi, E. Hirota “Infrared Diode Laser Study of the Hydrogen Bifluoride Anion: FHF^-“ *J. Chem. Phys.* **1986**, 84, 2953 – 2960.

2. J. E. Del Bene, M. J. T. Jordan “A Comparative Study of Vibrational Anharmonicity in the Bihalide Anions XHX^-: X=F, Cl, Br” *Spectrochimica Acta Part A* **1999**, 55, 719 – 729.

3. K. R. Asmis, N. L. Pivonka, G. Santambrogio, M. Brummer, C. Kaposta, D. M. Neumark, L. Woste “Gas-Phase Infrared Spectrum of the Protonated Water Dimer” *Science* **2003**, 299, 1375 – 1377.

4. D. T. Moore, J. Oomens, L. van der Meer, G. von Helden, G. Mejer, J. Valle, A. G. Marshall, J. R. Eyler “Probing the Vibrations of Shared, OH^+O-Bound Protons in the Gas Phase” *ChemPhysChem* **2004**, 5, 740 – 743.

5. M. U. Ehsan, Y. Bozai, W. L. Pearson III, N. A. Horenstein, J. R. Eyler “Infrared Multiple Photon Dissociation Spectrum of Protonated Bis(2-methoxyethyl)ether Obtained with a Tunable CO_2 Laser” *Phys. Chem. Chem. Phys.* **2015**, in press DOI: 10.1039/c5cp01752a.

6. G. J. O. Beran, E. L. Chronister, L. L. Daemen, A. R. Moehlig, L. J. Mueller, J. Ooens, A. Rice, D. R. Santiago-Dieppa, F. S. Tham, K. Theel, S. Yaghmaei, T. H. Morton “Vibrations of a Chelated Proton in a Protonated tertiary Diamine” *Phys. Chem. Chem. Phys.* **2011**, 13, 20380 – 20392.

7. M. D. Pluth, R. G. Bergman, K. N. Raymond “Proton-Mediated Chemistry and Catalysis in a Self-Assembled Supramolecular Host” *Acc. Chem. Res.* **2009**, 42, 1650 – 1659.

8. C. L. Perrin, B. K. Ohta “Symmetry of N-H-N Hydrogen Bonds in 1,8-Bis(dimethylamino)naphthalene·H+ and 2,7-Dimethoxy-1,8-bis(dimethylamino)naphthalene·H+” *J. Am. Chem. Soc.* **2001**, 123, 6520 – 6526.

9. R. W. Alder, P. S. Bowman, W. R. S. Steele, D. R. Winterman “The Remarkable Basicity of 1,8-Bis(dimethylamino)naphthalene” *Chem . Commun.* **1968**, 723 – 724.

10. Y. Masuda, Y. Mori, K. Sakurai “Effects of Counterion and Solvent on Proton Location and Proton Transfer Dynamics of N-H···N Hydrogen Bond

of Monoprotonated 1,8-Bis(dimethylamino)naphthalene" *J. Phys. Chem. A.* **2013**, 117, 1576 – 10587.

11. B. Brzezinski, E. Grech, Z. Malarski, L. Sobczyk "Infrared Spectra and Protonation of 1,8-Bis(dimethylamino)naphthalene in Acetonitrile" *J. Chem. Soc., Faraday Trans.*, **1990**, 86, 1777 – 1780.

12. J. A. Kanters, A. Schouten, A. J. M. Duisenberg "Temperature Effect on the Structure of the Complex 1,8-Bis(dimethylamino)naphthalene-Chloranilic Acid* (2/1) Dihydrate" *Acta Cryst.*, **1991**, C47, 2148 – 2151.

13. K. Wozniak, H. He, J. Klinowski, E. Grech "^{1}H and ^{13}C Solid-State NMR Studies of 1-8-Bis(dimethylamino)naphthalene and Its Complexes" *J. Phys. Chem.* **1995**, 5, 1403 – 1409.

14. V. A. Ozeryanskii, A. F. Pozharskii, A. J. Bienko, W. Sawka-Dobrowolska, L. Sobczyk "$[NHN]^{+}$ Hydrogen Bonding in Protonated 1,8-Bis(dimethylamino)-2,7-dimethoxynaphthalene. X-ray Diffraction, Infrared, and Theoretical ab Initio and DFT Studies" *J. Phys. Chem. A.* **2005**, 109, 1637 – 1642.

15. W. Sawka-Dobrowolska, E. Grech, B. Brezezinski, Z. Malarski, L. Sobczyk "Crystal and Molecular Structure and IR Spectrum of 1,8-Bis(dimethylaminomethyl)naphthalene Monoperchlorate" *J. Mol. Struct.* **1994**, 319, 11 – 16.

16. A. F. DeBlase, S. Bloom, T.Lectka, K. D. Jordan, A. B. McCoy, M. A. Johnson "Origin of the Diffuse Vibrational Signature of a Cyclic Intramolecular Proton Bond: Anharmonic Analysis of Protonated 1,8-Disubstituted Naphthalene ions" *J. Chem. Phys.* **2013**, 139, 024301.

17. V. A. Ozeryanskii, A. A. Milov, V. I. Minkin, A. F. Pozharski "1,8-Bis(dimethylamino)naphthalene-2,7-diolate: A Simple Arylamine Nitrogen Base with Hydride-Ion-Comparable Proton Affinity" *Angew. Chem. Int. Ed.* **2006**, 45, 1453 – 1456.

18. A. R. Moehlig "Peptides, Diamines, and Nucleobases: Investigations in Ion Chemistry" University of California, Riverside, **2011**.

19. S. Yaghmaei “In Search of a Low Barrier Hydrogen Bond in Proton Bridged Diamines” University of California, Riverside, **2008**.

20. S. Khodagholian “Low Barrier Hydrogen Bonds in Acyclic Tertiary Diamines” University of California, Riverside, **2010**.

21. S. Yaghmaei, S. Khodagholian, J. M. Kaiser, F. S. Tham, L. J. Mueller, T. H. Morton “Chelation of a Proton by an Aliphatic Tertiary Diamine” *J. Am. Chem. Soc.* **2008**, 130, 7836 – 7838.

22. S. Gronert “Determining the Gas-Phase Properties and Reactivities of Multiply Charged Ions” *J. Mass Spectrom.* **1999**, 34, 787 – 796.

23. L. Feketeova, R. A. J. O’Hair “Multiply Protonated Betaine Clusters are Stable in the Gas Phase” *Chem. Commun.* **2008**, 4942 – 4944.

24. C. Xu, H. Guo, Z. S. Breitbach, D. W. Armstrong “Mechanism and Sensitivity of Anion Detection Using Rationally Designed Unsymmetrical Dications in Paired Ion Electrospray Ionization Mass Spectrometry” *Anal. Chem.* **2014**, 86, 2665 – 2672.

25. A. T. Blades, P. Kebarle “Study of the Stability and Hydration of Doubly Charged Ions in the Gas Phase: SO_4^{2-}, $S_2O_6^{2-}$, $S_2O_8^{2-}$, and Some Related Species” *J. Am. Chem. Soc.* **1994**, 116, 10761 – 10766.

26. J. J. Roy, D. Boismenu, H. Gao, O. A. Mamer, F. Varin “Measurement of Succinylcholine Concentration in Human Plasma by Electrospray Tandem Mass Spectrometry” *Anal. Biochem.* **2001**, 290, 238 – 245.

27. N. C. Polfer, J. Oomens “Reaction Products in Mass Spectrometry Elucidated with Infrared Spectroscopy” *Phys. Chem. Chem. Phys.* **2007**, 9, 3804 – 3817.

28. D. Oepts, A. F. G. van der Meer, P. W. van Amersfoort “The Free-Electron-Laser user facility FELIX” *Infrared Phys. Technol.* **1995**, 36, 297 – 308.

29. N. C. Polfer “Infrared Multiple Photon Dissociation Spectroscopy of Trapped Ions” *Chem. Soc. Rev.* **2010**, 40, 2211 – 2221.

30. H. Schindlbauer “Reaktionen mit Dimethylformamid, 1. Mitt.: Synthese von Carbonsauredimethylamiden” *Monatsh. Chem.***1968**, 99, 1799 – 1807.

31. H. Schindlbauer “Reaktionen mit Dimethylformamid, 2. Mitt.: Uber die Bildung von Carbonsauredimethylamiden” *Monatsh. Chem.***1969**, 100, 1413 – 11422.

Chapter II:

1-Methylcytosine Proton-Bound Dimers

Introduction

DNA contains all of the information required for life to exist. Formation of the DNA double-helix was first discovered by Watson and Crick in 1953[1]. The DNA double-helix comprises of complementary pairs of nucleobases: guanine, cytosine, adenine, thymine, and their *N*-methylated analogues. **Figure 2.1** depicts cytosine forming a three-point hydrogen bond with guanine and thymine forming a two-point hydrogen bond with adenine on anti-parallel strands. With a greater number of hydrogen bonding interactions, the cytosine : guanine base pair (C•G) is bound more tightly compared to the adenine : thymine base pair. It has been over 60 years since the discovery of the DNA double-helix, but research on DNA nucleobases still continues to be an area of intense interest.

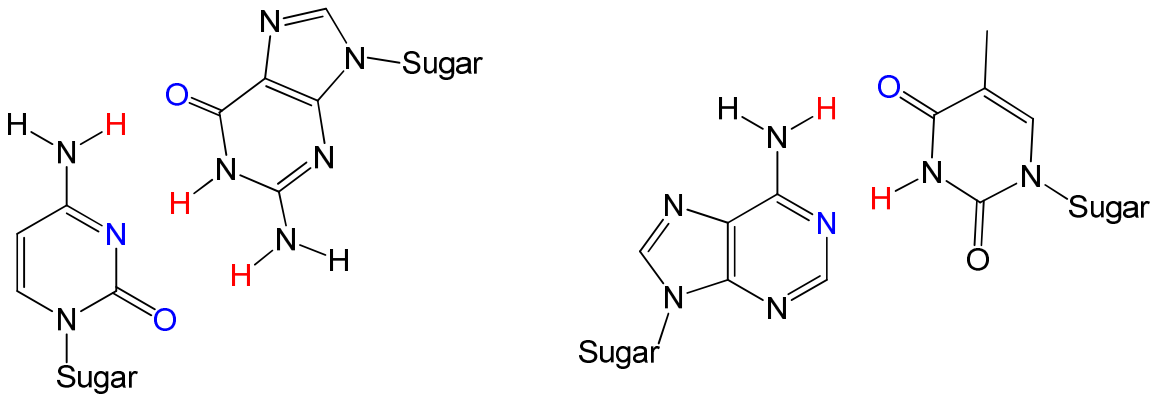


Figure 2.1: Pairing of DNA nucleobases cytosine with guanine and adenine with thymine. Hydrogen bond donor groups are shown in red while the hydrogen bond acceptor groups are in blue.

The aforementioned pairing of cytosine with guanine and adenine with thymine utilizes traditional Watson-Crick base pairing. In the case of guanine, interactions on the Hoogsteen face (**Figure 2.2**) have been observed to form a secondary structure known as the G-quadruplex[2-5].

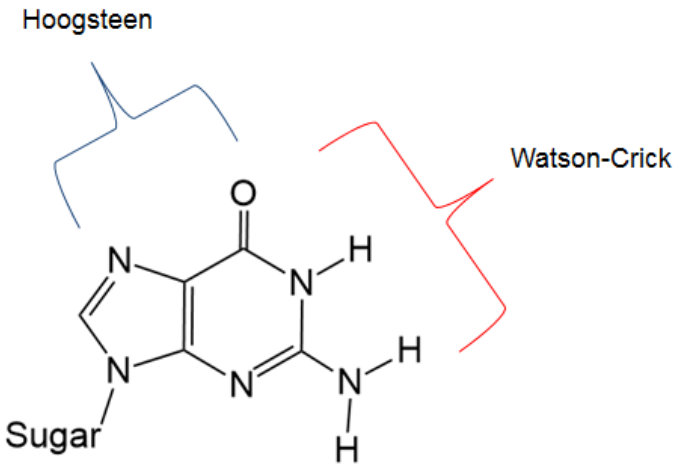


Figure 2.2: Watson-Crick and Hoogsteen faces on guanine available for base paring.

Four guanine residues can utilize both Hoogsteen and Watson-Crick faces to form the G-quaduplex (**Figure 2.3**). A typical G-quadruplex usually contains three layers of $(guanine)_4$ base pairs, but fewer layers have also been observed[6]. The formation of the G-quaduplex can occur on segments of DNA that are primarily guanine rich. A vast area of research has been dedicated to the stabilization of the G-quaduplex.

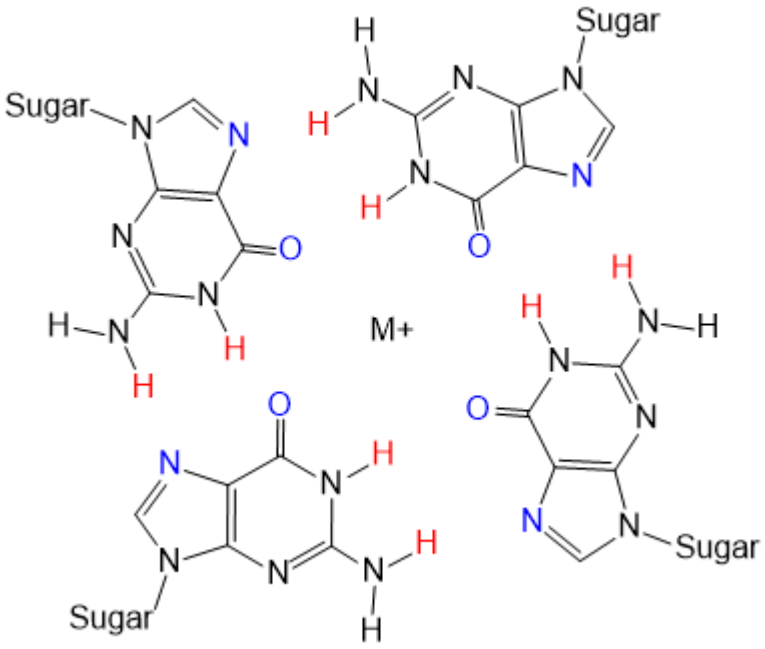


Figure 2.3: Formation of the G-quadruplex consisting of four guanine residues. Hydrogen bonding occurs on the Watson-Crick and Hoogsteen faces with the hydrogen bond donors shown in red and the hydrogen bond acceptors in blue. A metal cation, usually potassium, serves as a template in the middle of the G-quadruplex.

Cancerous cells can avoid normal regulatory steps, such as apoptosis, and lead to uncontrollable cell growth. Weinberg *et al.* described six acquired capabilities associated with cancer cells[7] which includes: self-producing growth signals, insensitivity to antigrowth signals, evasion of apoptosis, limitless reproductive potentials, angiogenesis, and invasion of healthy tissues and metastasis. Due to the formation of this secondary structure as a persistent feature within telomeres of malignant cells, a lot of effort has been devoted to the stabilization/destabilization of the G-quadruplex in hope of halting the expression of cancerous genes.

Hurley *et al.* have classified four kinds of G-quadruplexes [8]. The different classes are based on the number of G-quadruplexes, strand orientation, and the number of bases in the loop regions. An example of a G-quadruplex notation is 3'-[1:2:1]-5'. The reading of the G-quadruplex notation begins at the 3'-end and ends at the 5'-end. The number of integers separated by colon will determine the number of G-quadruplexes along with the number of loops, while the integers themselves tell the number of bases that separate the quadruplexes.

A class I G-quadruplex is the simplest of the four. One G-quadruplex predominates but can form multiple loop isomers. **Figure 2.4A** depicts an example of a Class I G-quadruplex. Class I contains a total of three quadruplexes (quadruplex 1 contains 1 base, quadruplex contains 2 bases, and quadruplex 3 contains 1 base), and completes at the 5'-end. An example of a Class I G-quadruplex is the c-Myc gene[8]. Four loop isomers are possible with c-Myc with varying loop bases with 3'-[1:2:1]-5' being the predominant isomer[9]. Class II differs from Class I in that two separate G-quadruplexes are formed. An example of a Class II G-quadruplex is c-Kit. The spacing between the two G-quadruplexes is separated by 33 base pairs, but both G-quadruplexes are formed independently from one another[8]. The notation is read the same way as for Class I, starting from the 3'-end and ending at the 5'-end. A Class III G-quadruplex also forms a pair of G-quadruplexes, like Class II, but instead of being separated by a large sequence of bases, the two G-quadruplexes are formed in tandem. Having two sets of G-quadruplexes form in tandem is more stable than having the G-

quadruplexes form separately[8]. Examples of genes that form Class III G-quadruplexes are c-Myb[10] and hTERT[11]. A class IV G-quadruplex is a case where multiple G-quadruplexes exist due to overlapping of the guanine residues. An example of a Class IV G-quadruplex is Bcl-2. This gene can adopt more than one kind of G-quadruplex with a variety of loops and bases within the loops.

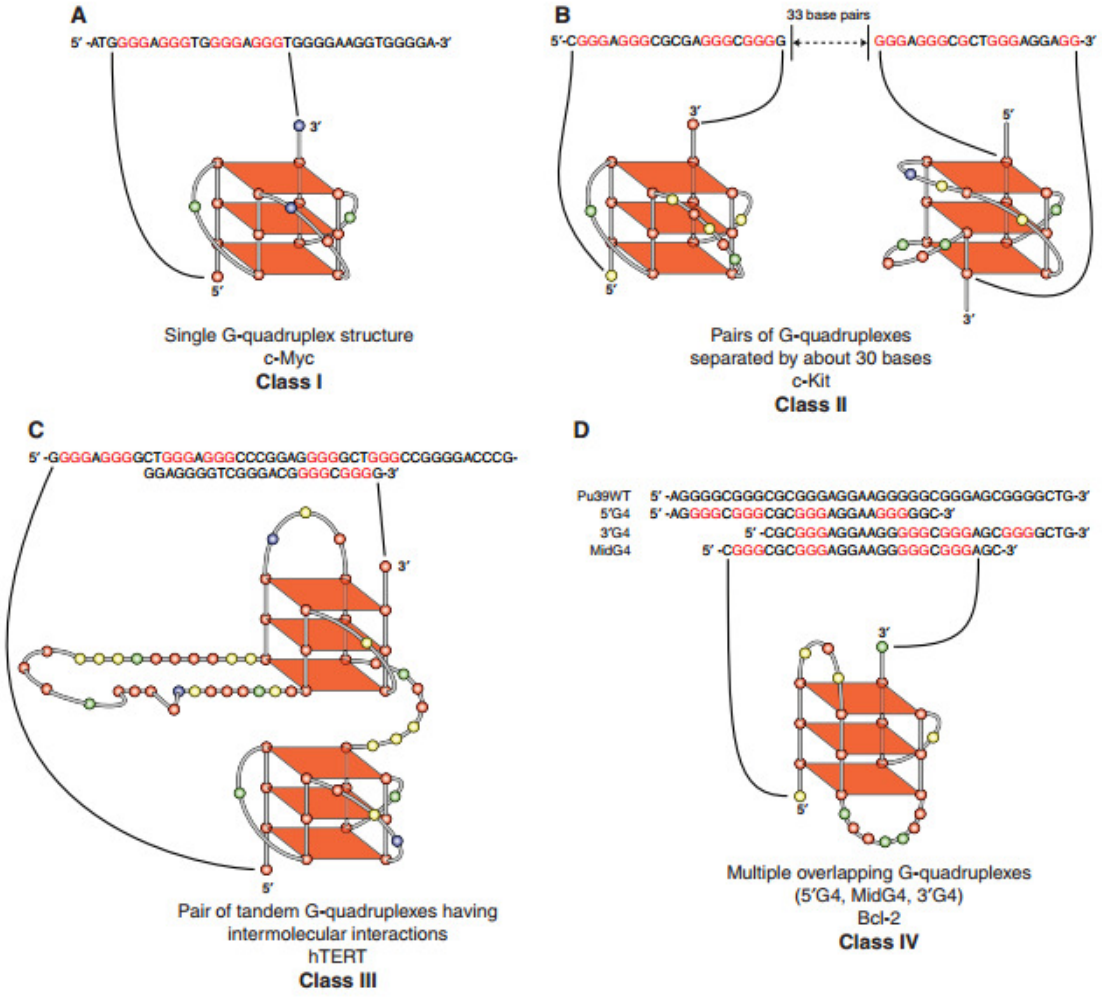


Figure 2.4[8]: Four classes of G-quadruplexes located within several oncogene sequences. **A)** Class I G-quadruplex within c-Myc. **B)** Class II G-quadruplexes within c-Kit. **C)** Class III G-quadruplexes within hTERT. **D)** Class IV G-quadruplex within Bcl-2.

Over the years, the stability of the G-quadruplex has received much attention, while the importance of the complementary *i*-motif is only starting to be recognized. Binding to and stabilization of the *i*-motif by small organic molecules (especially in promoter regions) may potentially inhibit transcription and the expression of oncogenes. The synthesis of these small molecules will be discussed later on.

Background

As double-stranded DNA unwinds and becomes supercoiled (as in the case of transcription or replication) it can adopt the aforementioned secondary structures. The G-quadruplex is known to exist under physiological conditions in DNA strands that are G-rich, and have been shown to be present in human cells[12]. Although the complementary C-rich strand should form the *i*-motif, it has not yet been observed *in vivo*. The *i*-motif consists of at least **6** PBDs of cytosine intercalated with one another oriented 90^{o} in respect to the PBDs above and below, utilizing two ordinary and one ionic hydrogen bond (**Figure 2.5**)[13-15]. Due to charge repulsion, it is possible for the top or bottom cytosine-H^{+}-cytosine base pairs to dissociate, making the protonated and neutral cytosine residue accessible to ligand binding[16]. Having a ligand bind to the residues may increase the stability of the *i*-motif. The dissociation enthalpy of the ionic hydrogen bond between cytosine and protonated cytosine has been calculated by Rodgers *et al.* with values ranging from 160 kJ/mol[13] to 173 kJ/mol[14,17].

Sequence repeats of cytosine and guanine at the ends of eukaryotic chromosomes are known as telomeres. Telomeres serve as a safety cap at the end of a chromosome to protect the genetic material during cell replication. After repeated cell replication cycles, parts of telomere begin to erode in cells that do not express the enzyme known as telomerase. Telomerase is known to extend the telomere, and this enzyme has been shown to be active in cancers[18].

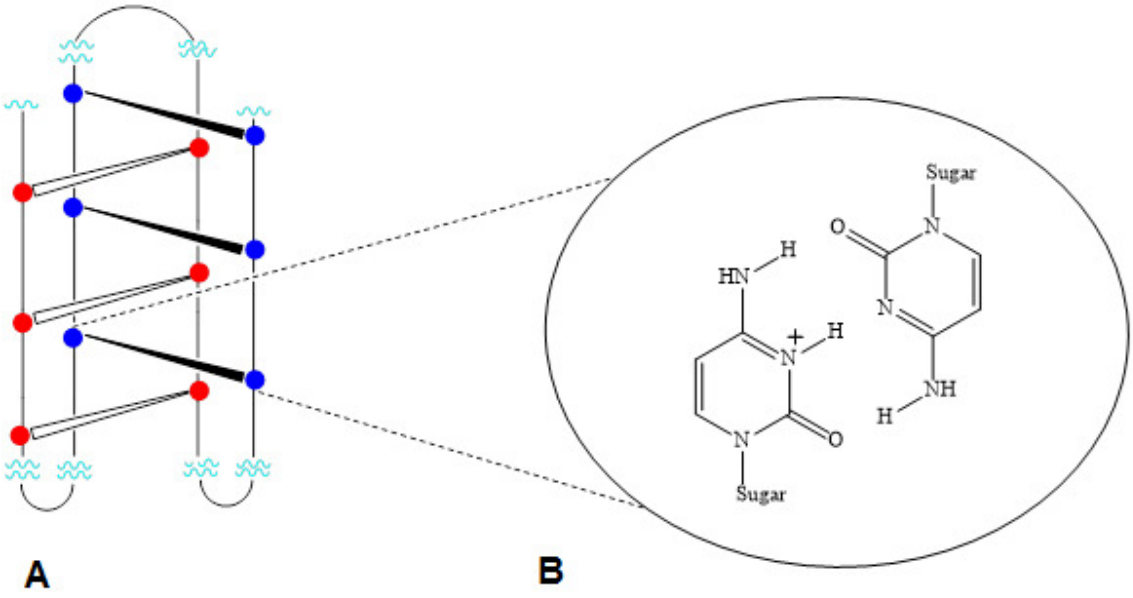


Figure 2.5: **A)** Schematic representation of the *i*-motif. White and black dashes represent six pairs of cytosines forming a PBD. **B)** PBD formation between protonated and neutral cytosine. The PBD utilizes two ordinary hydrogen bonds and one ionic hydrogen bond to stabilize the complex.

Like the G-quadruplex, *i*-motifs are placed into different classes based on the number of bases in the loop region. Hurley *et al*. identified two classes of i-motif[8] (**Figure 2.6**). *i*-Motif complexes containing small loops are designated

Class I while larger loops are designated Class II. Three examples of Class I *i*-motifs are VEGF, RET, and Rb. These three sequences have relative small number of bases that make up the loop region. Two examples of Class II *i*-motifs are c-Myc and bcl-2. Reading of *i*-motif notation is similar to G-quadruplex notation. An example of *i*-motif notation is 5'-(2:3-4:2)-3'. Reading of this notation begins at the 5'-end and ends at the 3'-end. The number of integers separated by colons represents the number of loops. In this example, there are three loop regions. The first loop consists of two bases, the second loop can either contain three or four bases, and the third loop contains two bases.

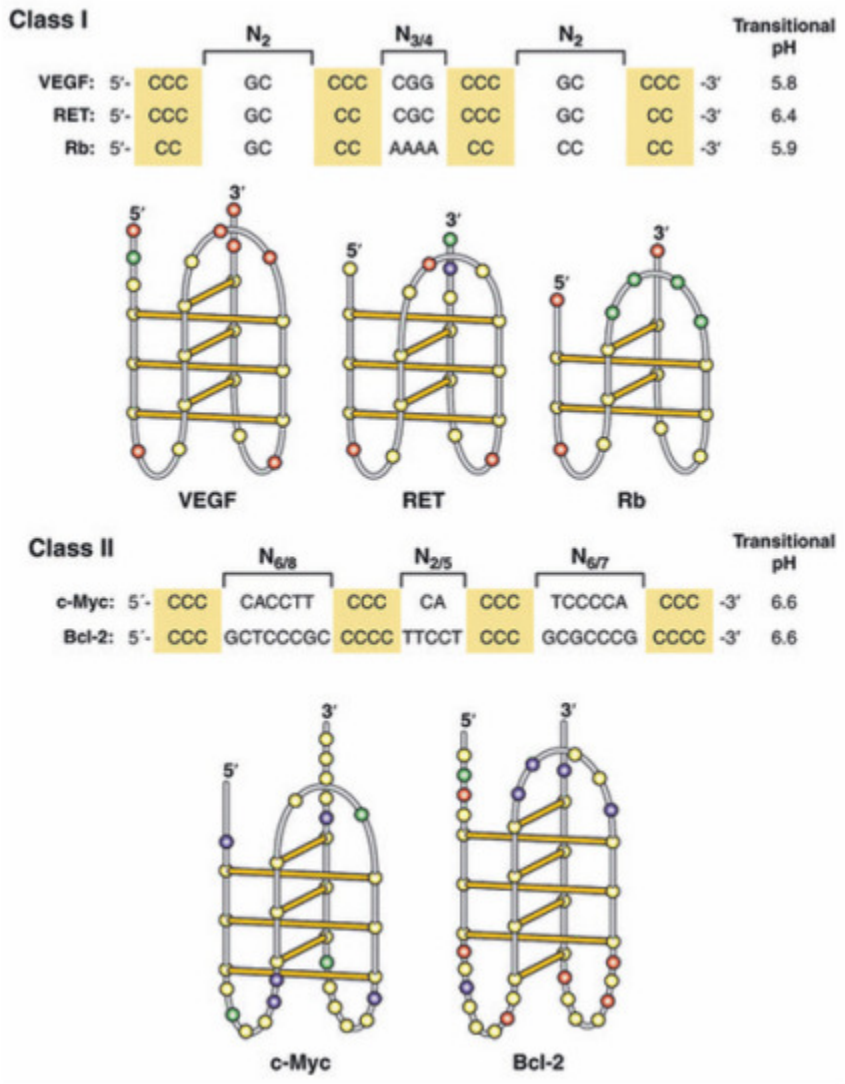


Figure 2.6[8]: Two classes of *i*-motif located within several oncogenes *cis*-promoter sequences. Class I *i*-motifs have smaller loop regions when compared to Class II *i*-motifs.

In 1993, Gueron *et al.* discovered a tetrameric complex consisting of four 5'-d(TCCCCC) strands in acidic solution[19]. To form this tetrameric complex, half of the base pairs on one strand must acquire protons and form a duplex with another strand. Another set of strands then form a separate duplex structure with

its PBDs intercalating between the PBDs of the initial duplex. Two strands of the oligomers are orientated in a parallel and an anti-parallel orientation (**Figure 2.7**). NMR analysis showed the structure to be symmetrical and each of the oligomers to be identical, at least on the NMR timescale. Five proton resonances between 15 – 16 ppm are observed and assigned to be the bridging protons due to their chemical shifts. There are ten bridging protons in the tetrameric complex, but due to symmetry five bands are observed. Gueron *et al.* also investigated a smaller tetrameric complex consisting of four d(TCC) strands[20]. Their results confirmed *i*-motif formation formed between intercalating base pairs between two sets of parallel strands. Two proton resonances are again observed between 15 – 16 ppm, corresponding to the bridging protons.

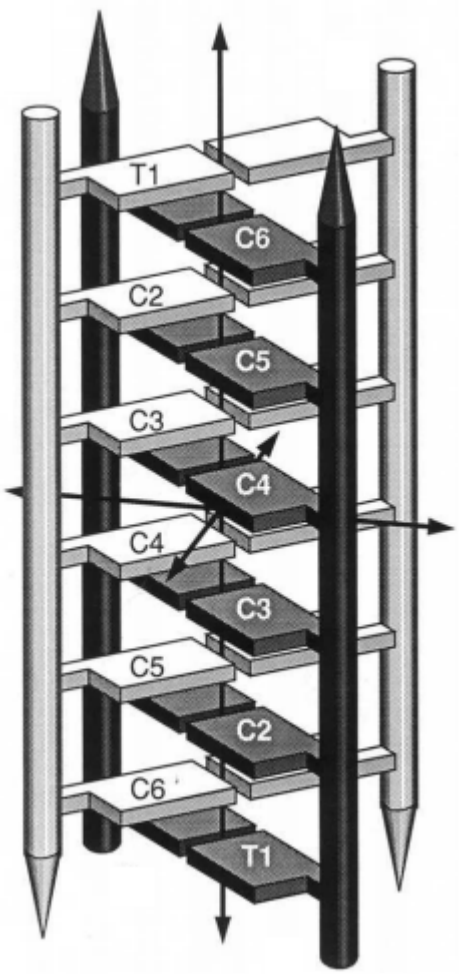


Figure 2.7[19]: Tetrameric complex formation between four strands of d(TCCCCC).

The *i*-motif consists of six (or more) cytosine PBDs intercalated with one another. Different *i*-motif sequences will have different bases associated with the loop region. Certain aspects can affect the stability of the *i*-motif, such as the nature of the phosphate backbone. Lacroix *et al.* investigated the effects of modification of the phosphate backbone of oligomers containing high number of cytidine residues[21]. Another aspect that alters the stability of the *i*-motif is the sugar conformation. DNA and RNA sugars differ at the 2'-position. Gehring *et al.*

investigated the stabilities of nucleotide sequences with DNA and RNA sugars using UV melting curves[22]. Having an OH group at the 2'-position oriented toward the minor groove (down) destabilized the RNA *i*-motif when compared to the corresponding DNA *i*-motif. To test the effect of the OH group, nucleotide sequences containing an arabinofuranose sugar, where the OH group is oriented toward the major groove (up), was used in place of the traditional furanose sugar and determined to be more stable. The introduction of a methoxy group at the 2'-position, oriented toward the minor groove, completely inhibit *i*-motif formation[23].

To achieve a better understanding of the *i*-motif, we must first understand the simplest component of the *i*-motif. The PBD of cytosine represents the core of which the *i*-motif consists. Oomens, Moehlig, and Morton[31] were first to report the IRMPD spectra of hemiprotonated 1-methylcytosine dimer and a deuterated analogue in the gas phase, and Rodgers *et al.* subsequently investigated the gas phase infrared signatures of cytosine proton-bound homo- and heterodimers[13]. Different homo- and heterodimer tautomer conformations were calculated at the B3LYP/6-31G** level and in each case, only a single tautomer is predicted to exist in the gas phase.

One of the first cytosine dimers to be identified was that of cytosine-5-acetic acid[24]. This dimer is unique in a sense that the proton associated with the acetic acid group at the 5-position of one of the bases is transferred to the nitrogen at the *N3*-position, resulting in a zwitterion, while the other cytosine-5-acetic acid molecule remains unaltered. Since the zwitterion base has a nitrogen

at the 3-position, this base has a hydrogen bond D (donor)-D-A (acceptor) motif. The unaltered cytosine-5-acetic acid base, when rotated 180°C, has a hydrogen bonding motif of A-A-D, which complements the zwitterion base. **Figure 2.8** depicts dimer formation between two cytosine-5-acetic acid molecules.

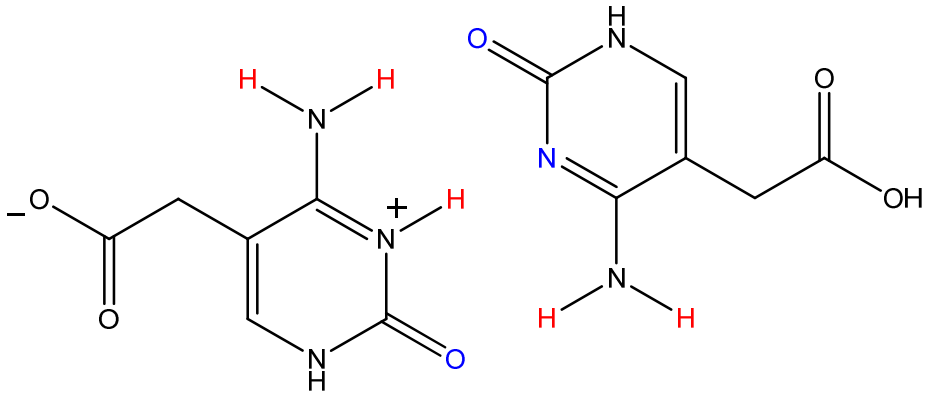


Figure 2.8: Dimer formation between two cytosine-5-acetic acid molecules. Hydrogen bond donors are depicted in red, while hydrogen bond acceptors are shown in blue.

1-Methylcytosine PBD salts have been studied in condensed phases for decades. The 1-position of cytosine is methylated to avoid any unwanted hydrogen bonding interaction that may occur at that position (because the sugar is at that position in the polymer of DNA). The X-ray crystal structure of 1-methylcytosine PBD iodide salt was first published in 1978[25]. The bridging proton at the *N3*-position is inferred to be shared unequally due to the unequal distances between the amino hydrogens and the carbonyl oxygens. The hydrogen bond distance between the amino group on the protonated 1-methylcytosine and the carbonyl oxygen (1.84 Å) on the uncharged 1-methylcytosine is shorter by 0.18 Å

when compared to the other hydrogen bond between the uncharged amino hydrogen and carbonyl oxygen (2.02 Å). A DFT calculation at the B3LYP/cc-pVTZ level predicts the hydrogen bond distance between the amino group on the protonated 1-methylcytosine and the carbonyl oxygen to be 1.64 Å and the flanking hydrogen bond distance to be 1.95 Å. Sindona *et al.* reported a symmetrical structure of 1-methylcytosine PBD perchlorate salt (**Figure 2.9**) when the recrystallization process was performed using calcium perchlorate[26]. The pyrimidine rings are equivalent and the bridging proton cannot be assigned to either one of the bases.

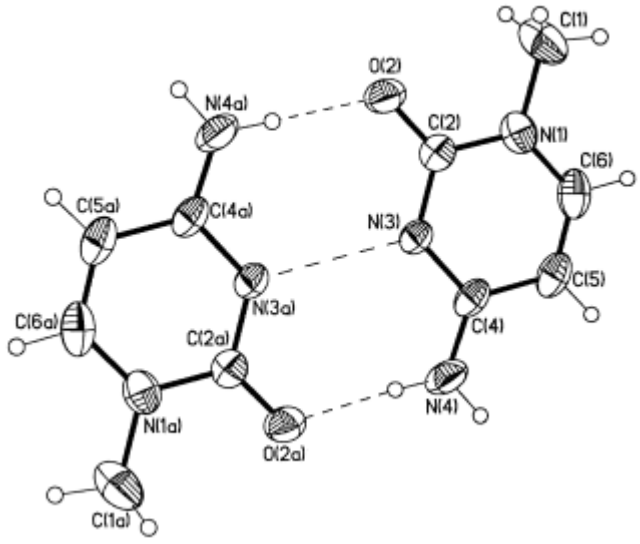


Figure 2.9[26]: X-ray cyrstal structure of 1-methylcytosine PBD perchlorate salt.

Different 1-methylcytosine PBD iodide salts have been prepared and results have varied[25,27,28]. The X-ray crystal structure of 1-methylcytosine PBD iodide salt obtained by Steinborn *et al.* shows a symmetrical structure at 298 K. A

year later, Freisinger *et al.* obtained an X-ray crystal structure of the same salt, but at 163 K. Freisinger's results disputed that of Seinborn's by showing the PBD iodide salt exists as an asymmetric unit. Both crystals share the same parameters, so it might be that temperature plays a role in the formation of the crystals (or that Steinborn's solution of the diffraction pattern was incorrect). Biré *et al.* observed a phase change as a function of temperature in $[V_{10}O_{28}](C_4H_5N_3O)_3(C_4H_6N_3O)_3{\cdot}10H_2O$ at 210 K[29]. In any event, as will be detailed below, our single crystal X-ray studies of the PBD iodide salt confirm Freisinger's structure, both for protonated and deuterated analogues.

A unique infrared signature is observed for the PBD tetrafluoroborate salt of cytidine[30]. When the infrared spectrum of the PBD tetrafluoroborate salt of cytidine is compared to the infrared spectra of neutral and protonated cytidine tetrafluoroborate salt, a band is observed at 1890 cm^{-1} in the spectrum of the PBD tetrafluoroborate salt. Since this band is not present in either spectra of the neutral or protonated monomer, it was assigned to be the motion associated with the bridging proton. That same feature is also present in our own infrared spectrum of 1-methylcytosine PBD iodide salt, but more recent gas phase IR studies have shown a different band to be the motion associated with the bridging proton.

The gas phase infrared spectrum, in the fingerprint domain between 400 – 1800 cm^{-1} of 1-methylcytosine PBD was investigated by Oomens, Moehlig, and Morton[31]. **Figure 2.10** reproduces the predicted IR spectrum, calculated at

B3LYP/6-31G**, plotted against the experimental IRMPD spectrum. The comparison in panel **B** shows a good fit of peaks, except for a single band at 1570 cm^{-1} (red arrow). The gas phase IR of the d_5-analog of 1-methylcytosine PBD, where all five of the exchangeable hydrogens are exchanged for deuteria, was also obtained in the same domain. The comparison, shown in panel **D**, between calculated and experimental IR spectra of 1-methylcytosine PBD-d_5 also show a good fit of peaks. The band at 1570 cm^{-1}, that was present in the completely protonated spectrum of 1-methycytosine PBD, is no longer present in the spectrum of 1-methycytosine deuterated at the exchangeable positions. This band is assigned to be the motion associated with the bridging proton at the *N3*-position.

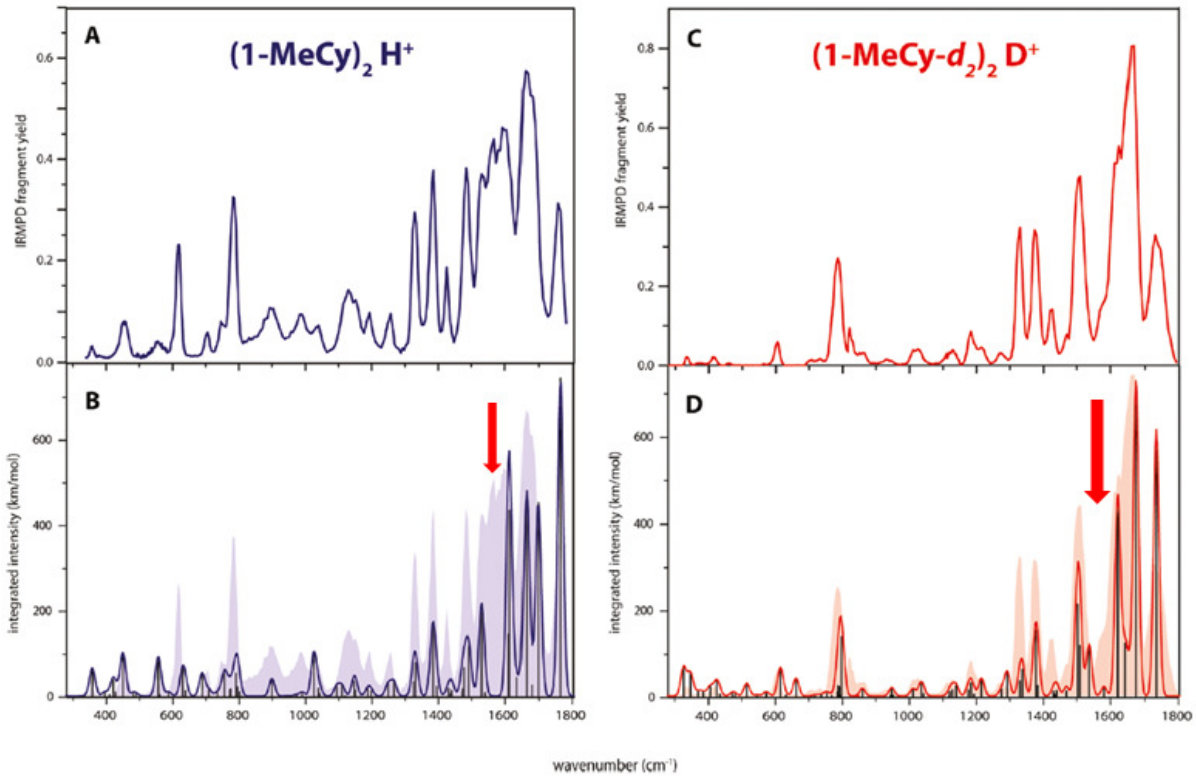


Figure 2.10[31]: Experimental and predicted IRMPD spectra of 1-methylcytosine PBD in the fingerprint domain from 400 – 1800 cm^{-1}. **A**) Experimental IRMPD spectrum of 1-methylcytosine PBD. **B**) Overlap between experimental (silhouette) and predicted IRMPD (foreground) spectra of 1-methylcytosine PBD. **C**) Experimental IRMPD spectrum of 1-methylcytosine PBD-d_5. **D**) Overlap between experimental (silhouette) and predicted IRMPD (foreground) spectra of 1-methylcytosine PBD-d_5.

This chapter will describe infrared spectra of protonated 1-methylcytosine monomers, 1-methylcytosine proton-bound homo-, and proton-bound heterodimers. The experimental gas phase infrared spectra of protonated 1-methylcytosine in the fingerprint and CH/NH/OH stretching domain are plotted against anharmonic vibrational modes calculated at the B3LYP/cc-pVTZ level. Proton-bound homo- and heterodimers IR spectra of 1-methylcytosine and its derivatives are investigated in the fingerprint and CH/NH/OH stretching domain.

Solid State NMR experiments is performed on 1-methylcytosine PBD iodide and triflate salts. X-ray diffraction and INS experiments were conducted on the 1-methylcytosine PBD iodide salt and its deuterated analogues.

Theory / Calculations

Gas phase harmonic DFT calculations used the Gaussian09 program suite while anharmonic calculation used a remote cluster, access to which has been provided by XSEDE. Geometry optimization and harmonic and anharmonic frequencies were calculated at the B3LYP/aug-CC-PVTZ level for *N3*- and *O*-protonated monomers of 1-methylcytosine, and a scaling factor of 0.97 was applied to all harmonic vibrational frequencies above 800 cm^{-1}. Predicted INS spectra were calculated at the B3LYP/6-31G** level using the Gaussian03 program and converted to INS spectra using ACLIMAX. All plots were created using OriginPro 10.1 (Student Edition). A line broadening of 10 cm^{-1} was applied to all plots. It was determined that a line broadening less than 10 cm^{-1} did not provide smooth spectra.

Fragment-based NMR calculations were used to predict solid phase ^{13}C chemical shifts of neutral 1-methylcytosine, protonated 1-methylcytosine triflate salt, 1-methylcytosine PBD triflate salt (**1-Triflate**), and 1-methylcytosine PBD iodide salt (**1-Iodide)** at the PBE0/6-311+G(2d,p) level by Josh Hartman, a graduate student in Prof. G. O. Beran's research group. Initial geometries were supplied through X-ray diffraction and is refined using the Crystal09 suite at the

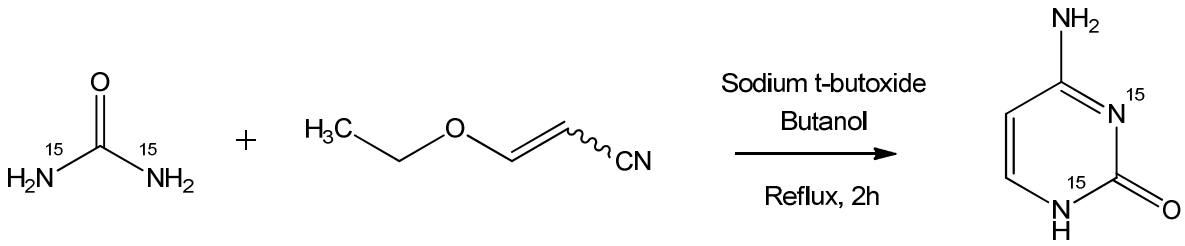
B3LYP/TZP level. Hydrogen-only optimizations using fixed unit cells were used to address the discrepancies in ^{13}C chemical shifts.

Synthesis and Characterization / Methods

Infrared powder spectra were obtained on a Perkin-Elmer One FT-ATR, Raman (1064 nm exciting line) on a Thermo Scientific Nicolet 6700 FT-IR with an NXR FT-Raman module, and the single crystal IR on a Bruker Equinox 55 equipped with an adjustable microscope. The techniques used to obtained gas phase IRs are described[32-34] elsewhere and in the previous chapter. To summarize, ions were generated using an ESI source from approximately 1 mM solution of 50/50 methanol:water mixture containing a trace of acetic acid. The ions were injected into a home-built 4.7 T FT-ion cyclotron resonance (FT-ICR) mass spectrometer using a quadrupole deflector and an octopole ion guide. Unwanted ions were then ejected by SWIFT ejection pulses applied to the excited plates of the FT-ICR cell[35] Gas phase IRMPD spectra were obtained by monitoring the expulsion of dissociated daughter ions as a function of IR wavelength. IR spectra in the fingerprint domain (600 – 1800 cm^{-1}) used a free-electron laser, while the CH/NH/OH stretching domain (2800 – 3800 cm^{-1}) utilized a LaserVision benchtop optical parametric oscillator (OPO) laser.

Synthesis

Two different published preparations of 15N 1-methylcytosine were tried, one of which was successful and the other unsuccessful. Both methods are outlined below.

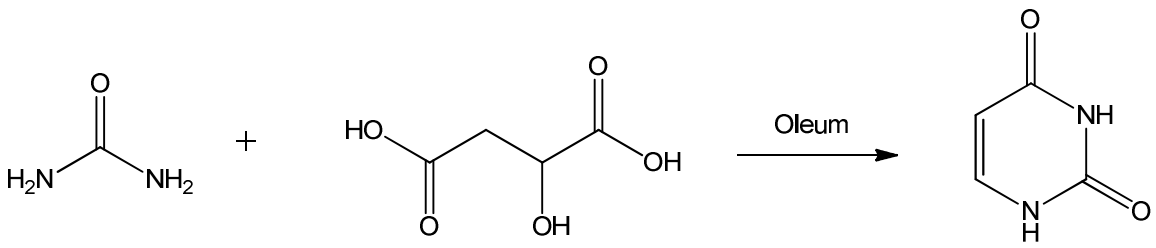


Scheme 2.1: Synthesis of cytosine by reacting urea with 3-ethoxyacrylonitrile in a solution containing sodium ethoxide in ethanol. 3-ethoxyacrylonitrile contained a mixture of both *cis* and *trans* isomers.[36] While the original procedure calls for unlabeled urea, the same reaction conditions were used to synthesize ^{15}N-labeled cytosine using ^{15}N-labeled urea.

$^{15}N_2$-urea and 3-ethoxyacrylonitrile were purchased and used without any further purification from Cambridge Isotope Laboratories, Inc. and Alfa-Aesar, respectively. The synthesis of $^{15}N_2$-cytosine was prepared as described in literature starting from unlabeled urea and 3-ethoxyacrylonitrile[36] as outlined in **Scheme 2.1**. To 338 mg sodium in 50 ml butanol in a 100 ml round bottom flask, equipped with a stir bar is added 1.0 grams $^{15}N_2$-urea (0.016 mol) and 1.7 mL (0.017 mol) 3-ethoxyacrylonitrile. The resulting solution was heated for 2 hours at 115^oC with a reflux condenser attached. After cooling to room temperature, 10 mL DI H_2O and 4 mL conc. H_2SO_4 were added and stirred at room temperature for 30 minutes. After stirring, 100 mL absolute ethanol was added and the

solution was heated at 90°C for 30 minutes. The resulting mixture was placed in the freezer overnight at 0°C. The solid was filtered and dried under vacuum. The filtrate was treated with a mixture containing 20 mL DI H_2O and 2 mL conc. ammonium hydroxide, and refiltered. A yield of 370 mg (21%) was obtained after recrystallization using hot absolute ethanol.

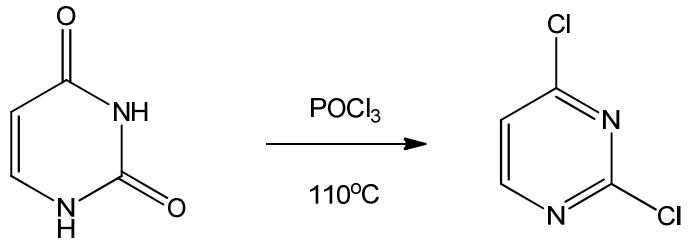
APCI-ESI MS: MH^+: 114.0449 (calc. 114.0446)



Scheme 2.2: Synthesis of uracil from urea and (DL)-malic acid.

Uracil was prepared as described in literature starting from urea and (DL)-malic acid[37] as outlined in **Scheme 2.2**. A 50 mL round bottom flask containing 4 mL sulfuric acid, 20% fuming, was cooled to 0°C. 1 gram (0.016) urea was added to the 50 mL round bottom flask equipped with a stir bar and stirred for 5 minutes. 1 Gram (DL)-malic acid (0.0074 mol) was added slowly over 20 minutes. The resulting solution was heated for 1 hour at 100°C with a reflux condenser attached. After cooling to room temperature, 15 mL DI H_2O was added slowly and cooled to 0°C. The white solid was filtered under vacuum and a yield of 456 mg (54%) obtained.

^{1}H NMR (300 MHz, D_2O): δ 5.81 – 5.83 (1H d) δ 7.53 – 7.56 (1H d)

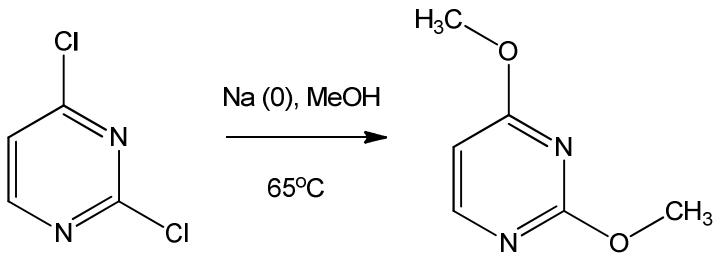


Scheme 2.3: Synthesis of 2,6-dichloropyrimidine.

456 mg (0.0041 mol) of uracil was added to 4 mL $POCl_3$ in a 25 ml round bottom flask equipped with a stir bar. The mixture was heated for 5 hours at 110°C with a reflux condenser attached. After 5 hours, excess $POCl_3$ was removed using short-path distillation. A brown goo remained after distillation and 10 grams of ice added. The brown solution was extracted three times with diethyl ether (20 mL diethyl ether each time). The organic layer was then dried using anhydrous sodium sulfate and the solution concentrated under reduced pressure. A yield of 191 mg (32%) was obtained.

APCI-ESI MS: MH^+: 148.9670 (calc. 148.9668)

^{1}H NMR (300 MHz, $CDCl_3$): δ 7.34 – 7.36 (1H d) δ 8.53 – 8.55 (1H d)

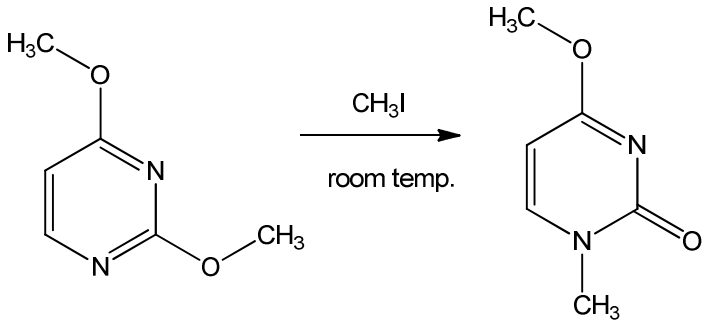


Scheme 2.4: Synthesis of 2,4-dimethoxypyrimidine.

To 1.2 g (0.052 mol) sodium in 10 mL methanol (MeOH) in a 100 ml round bottom flask equipped with a stir bar was added 500 mg (0.0034 mol) of 2,4-dichloropyrimidine. The reaction was heated and stirred at 65^{o}C for 12 hours, with a water reflux condenser attached. After 12 hours of heating and stirring, the solid (NaCl) was filtered under vacuum and excess methanol was distilled off using a short path distillation apparatus. 20 mL of 30% NaOH (aq) was then added and the solution extracted three times with diethyl ether (20 mL each time). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure. A yield of 37 mg (7.7%) was obtained.

APCI-ESI MS: MH^+: 141.0665 (calc. 141.0659)

^{1}H NMR (300 MHz, $CDCl_3$): δ 3.96 (3H s) δ 3.99 (3H s) 6.36 – 6.38 (1H d) δ 8.17 – 8.19 (1H d)

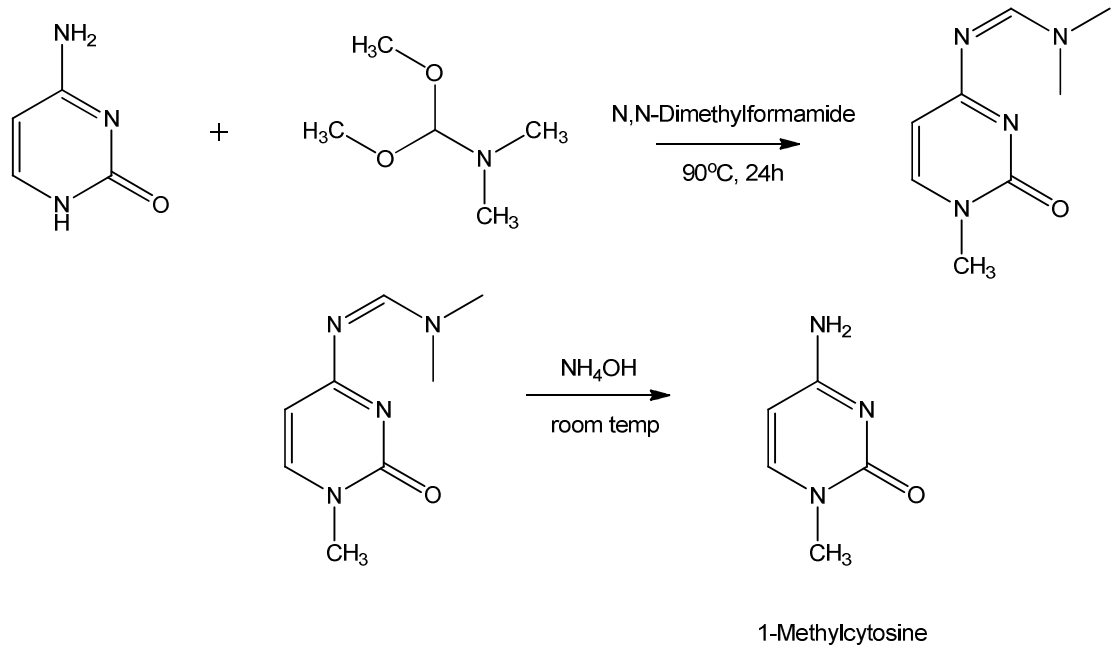


Scheme 2.5: Synthesis of 4-methoxy-1-methyl-2-pyrimidinone.

To 3 mL methyl iodide was added 37 mg (0.00026 mol) of 2,4-methoxyoropyrimidine. The reaction is stirred at room temperature for 12 hours

with an air reflux condenser attached. After 12 hours excess methyl iodide wasevaporated and a yellow solid remained. The yield was unattainable.

APCI-ESI MS: MH^+: 141.0684 (calc. 141.0658)



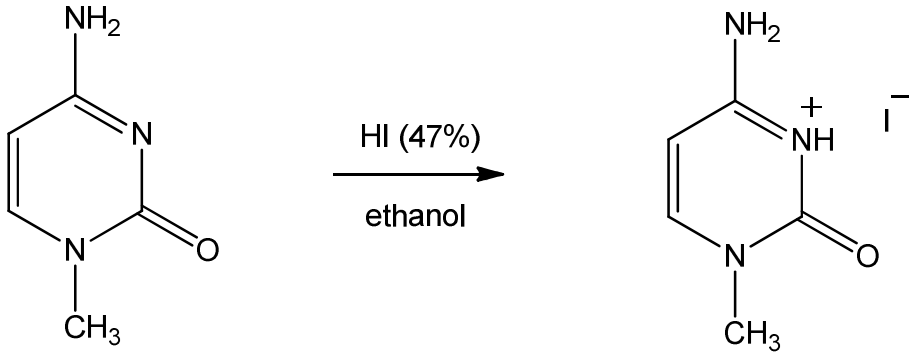
Scheme 2.6: Synthesis of 1-methylcytosine by treating cytosine with N,N-dimethylformamide dimethyl acetal. Deprotection was achieved upon treatment of concentrated ammonium hydroxide.

Methylation at the 1-position of cytosine followed a previously published procedure[38] as outlined in **Scheme 2.6**. To 2.3 grams (0.021 mol) of cytosine in 10 mL N,N-dimethylformamide (DMF) in a 100 mL round bottom flask equipped with a stir bar was added 19.4 mL (0.147 mol) of N,N- dimethylformamide dimethyl acetal. The resulting solution was heated for 12 hours at 90°C with a

reflux condenser attached. After cooling to room temperature, the solution was evaporated to dryness and 40 mL of conc. ammonium hydroxide is added. The solution was stirred at room temperature overnight. A yield of 980 mg (37%) was obtained after recrystallization using heated absolute ethanol. The same reaction procedure was employed to synthesize $^{15}N_2$-1-methylcytosine.

APCI-ESI MS: MH^+: 125.0668 (calc. 125.0661)

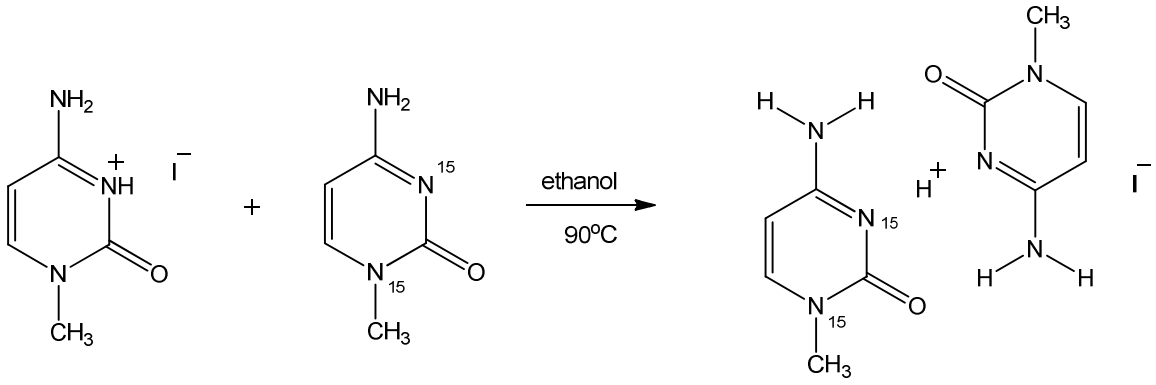
^{1}H NMR (300 MHz, D_2O): δ 3.39 (3H s) δ 5.96 – 5.98 (1H d) δ 7.56 – 7.59 (1H d)



Scheme 2.7: Synthesis of the monoprotonated iodide salt of 1-methylcytosine from 1-methylcytosine and hydroiodic acid.

To 350 mg (0.0028 mol) of 1-methylcytosine in 100 mL absolute ethanol in a 250 mL round bottom flask was added 1 mL HI (47%). The resulting solution was heated for at 90^{o}C until all solids dissolved. An additional heating time of 30 minutes with distillation of some of the solvent allowed for removal of any water that might be present. After cooling to room temperature, white crystals

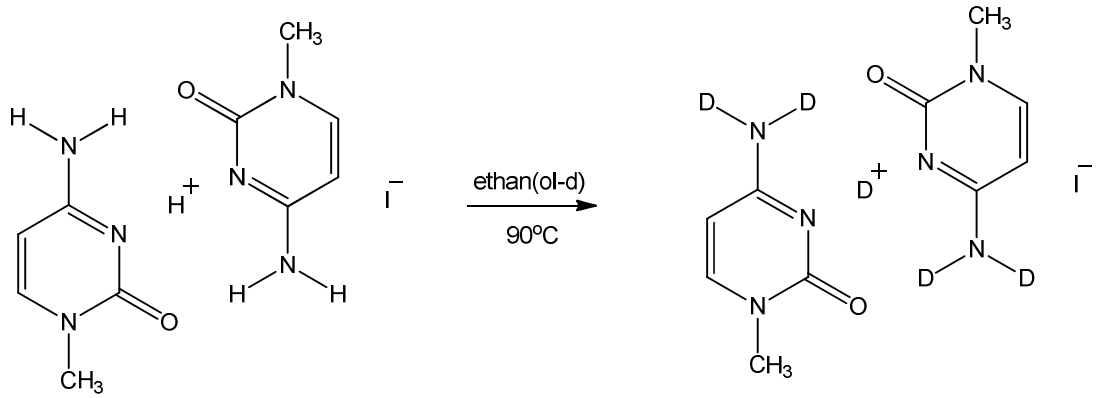
precipitated and the reaction mixture was cooled in an ice bath. Crystals were obtained via vacuum filtration. A yield of 700 mg (99%) was obtained.



Scheme 2.8: Synthesis of PBD iodide salt of $^{15}N_2$-1-methylcytosine starting from one equivalent of unlabeled monoprotonated 1-methylcytosine iodide salt and one equivalent of $^{15}N_2$-1-methylcytosine.

To 835 mg (0.0033 mol) of monoprotonated 1-methylcytosine iodide salt in 50 mL absolute ethanol in a 100 mL round bottom flask was added 420 mg (0.0033 mol) $^{15}N_2$-1-methylcytosine. The resulting solution was heated at 90°C until all solids dissolved. An additional heating time of 30 minutes with distillation of some of the solvent allowed for removal of any water that might be present. After cooling to room temperature, white needle-like crystals formed. The needle-like crystals were obtained via vacuum filtration. The filtrate was concentrated under reduced pressure and the solid refiltered. A yield of 450 mg (36%) was obtained after two recrystallizations using hot absolute ethanol.

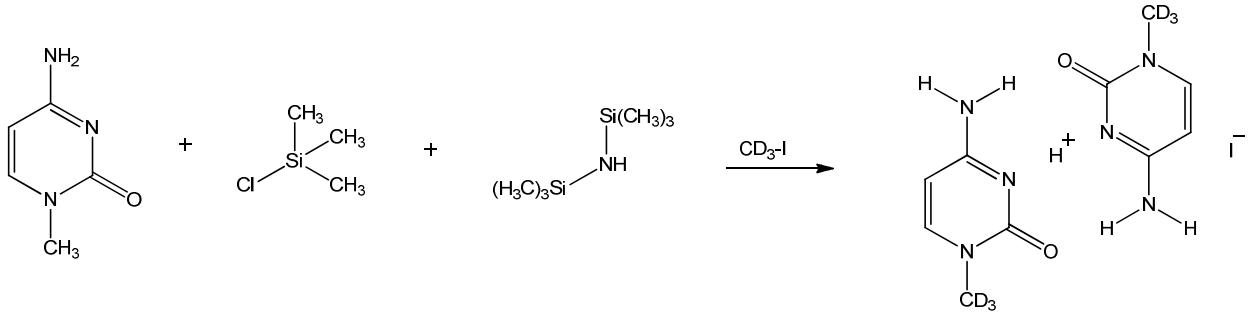
APCI-ESI MS: MH^+: 251.1265 (calc. 251.1251)



1-iodide **1**-iodide-d_5

Scheme 2.9: Synthesis of **1-Iodide-*d*$_5$** starting from **1-Iodide**.

350 mg (0.00094 mol) of 1-Iodide was dissolved in a minimum amount of heated ethanol(-od) (99%) in a 250 mL round bottom flask. The resulting solution was heated for at 90°C for an additional 30 minutes with distillation of some of the solvent allow for removal of any water that might be present. After cooling to room temperature, white needle-like crystals formed. The needle-like crystals were obtained via vacuum filtration. The filtrate was concentrated under reduced pressure and the solid refiltered. A yield of 150 mg (42%) was obtained after two recrystallizations.



Scheme 2.10: Synthesis of the **1-methylcytosine-*d*$_6$** PBD iodide salt.

1-Methylcytosine-d_6 PBD Iodide salt was synthesized following a published procedure and modified for use with methyl iodide-d_3 [38]. To 100 mg (0.0009 mol) of cytosine in 2 mL hexamethyldisilazane (HMDS) in a 25 ml round bottom flask was added 0.127 mL (0.00099 mol) trimethylchlorosilane (TMSCl). The reaction mixture was heated and stirred at 110^{o}C for 3 hours with a water reflux condenser attached. After 3 hours of stirring, the reaction was cooled to room temperature and 1 ml CD_3I added. The reaction pot was heated at 60^{o}C for an additional 18 hours with a water reflux condenser attached. After cooling to room temperature, the solution was concentrated under reduced pressure. The white solid was recrystallized using absolute ethanol and collected using vacuum filtration. A yield of 100 mg (29%) was obtained.

APCI-ESI MS: MH^+: 129.0871 (calc. 129.0850)

^{1}H NMR (300 MHz, D_2O): δ 6.17 (1H d) δ 7.85 (1H d)

Results

X-ray Diffraction Analysis

The X-ray crystal structure of 1-methylcytosine PBD iodide salt (**1-Iodide)** has been previously published [27,28] Our recrystallization method using slow evaporation of absolute ethanol yields two different crystal habits. The crystals corresponding to habit A looks to be plate-like, while habit B looks like needles. X-ray analysis of crystal habit A, plates, composed of the same unit cell

dimensions as previously published for **1-Iodide**[28]. Crystal habit A is monoclinic C; *a* = 7.1783 Å, *b* = 8.6135 Å, *c* = 11.4634 Å., *β* angle = 97.22 °. Two sets of proton-bound dimers occupy one unit cell. X-ray diffraction analysis of **1-Iodide** shows two proton-bound dimers within one unit cell. In order for **1-Iodide** to be symmetric, the two N-O distances must be equal. A complete X-ray analysis of **1-Iodide** habit B was unsuccessful due to racemic twinning within the crystal.

Like **1-Iodide**, two sets of proton-bound dimers occupy one until cell of **1-Iodide-*d*$_5$**. The synthesis of **1-Iodide-*d*$_5$** follows the same pathway as **1-Iodide**, except the recrystallization process calls for the use of ethanol(-od) in place of ethanol. The slow evaporation of ethanol(-od) produced colorless needles, identical to that of **1-Iodide** habit B. **1-Iodide-*d*$_5$** exhibits a monoclinic P, *a* = 7.1779 Å, *b* = 8.6271 Å, *c* = 11.4615 Å., *β* angle = 97.152°. Like **1-Iodide**, the crystal structure analysis of **1-Iodide-*d*$_5$** shows two separate proton-bound dimers per unit cell.

Although the crystal morphology of **1-Iodide** habit A (plates) and **1-Iodide-*d*$_5$** (needles) differ, they share the same crystal structure dimensions. 1-Methylcytosine PBD triflate salt (**1-Triflate)** exhibits a monoclinic P, *a* = 6.0732 Å, *b* = 20.2742 Å, *c* = 12.1555 Å., *β* angle = 99.6784°.

Gaussian09 NMR Predictions

The Gaussian09 program suite was also used to predict ^{13}C and ^{15}N chemical shifts of **1-Iodide** at the B3LYP/6-311++G** level in the gas phase and

corrected to give solution and solid phase estimates. If **1-Iodide** is a symmetrical salt, the ^{13}C experimental spectrum should display five carbon resonances and the ^{15}N spectrum should display three nitrogen resonances. Comparison between ^{13}C calculated and experimental solid phase NMR chemical shifts of **1-Iodide** is shown in **Figure 2.11**. Prediction does not match each ^{13}C resonances perfectly, but predicted and experimental bands are in the general vicinity. The two predicted methyl resonances (labeled A and A') are split experimentally (**Figure 2.11** top), but the splitting is predicted to be close enough to look like one band in the predicted spectrum (**Figure 2.11** bottom).

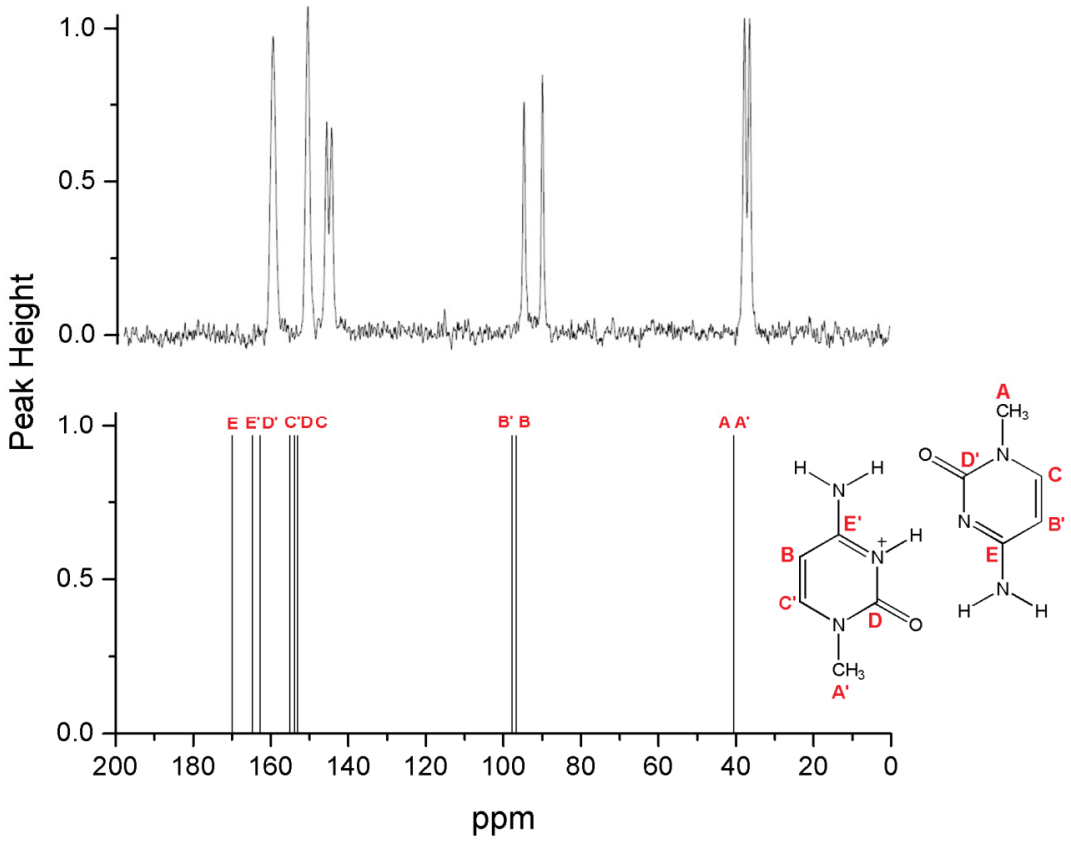


Figure 2.11: Comparison between experimental ^{13}C ssNMR spectrum (top spectrum) of **1-Iodide** and the predicted carbon spectrum computed from Gaussian09 calculated at the B3LYP/6-311++G** level (bottom spectrum).

The predicted gas phase ^{15}N spectrum does not correlate as well to experimental spectrum (**Figure 2.12**). Two predicted nitrogen resonances (labeled **f** and **f'**) are predicted to be separated by a greater distance than observed exerimentally. Calculations also predict nitrogen **h** to shift far downfield. The most deshielded nitrogen is observed experimentally at 170 ppm. Clearly, gas phase calculations do a poor job in this case in predicting ^{15}N chemical shifts.

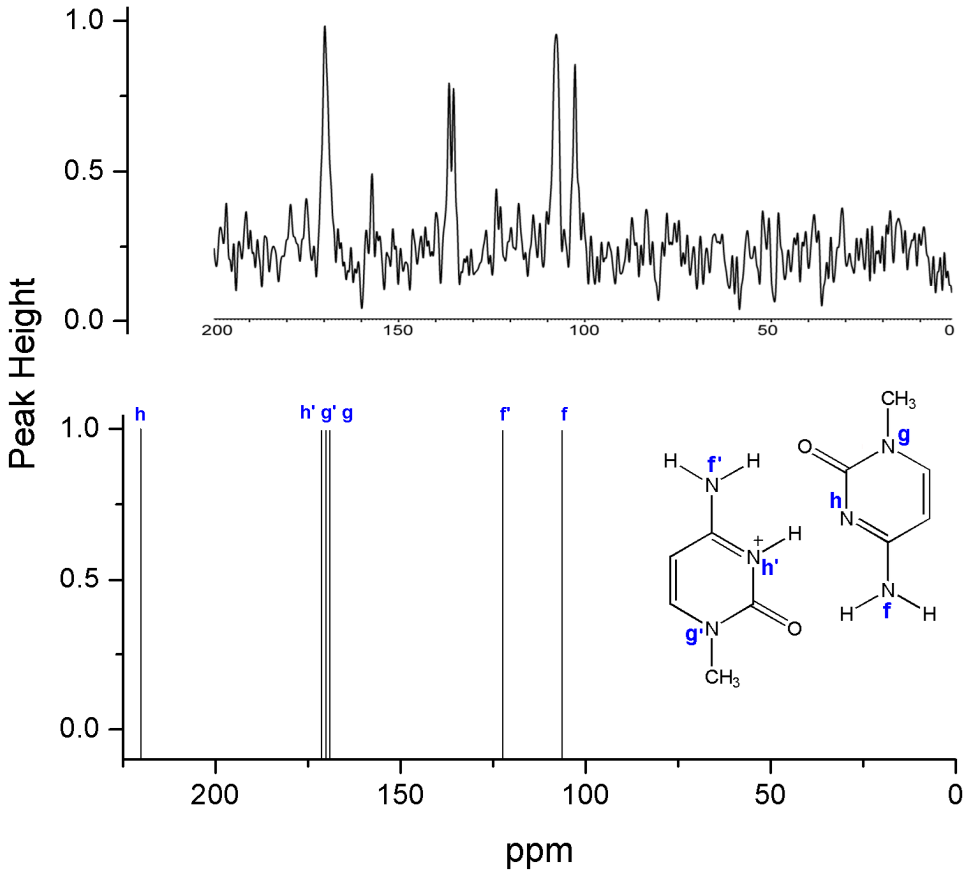


Figure 2.12: Comparison between experimental ^{15}N ssNMR spectrum (top panel) of **1-Iodide** and the predicted carbon spectrum computed from Gaussian09 calculated at the B3LYP/6-311++G** level (bottom panel).

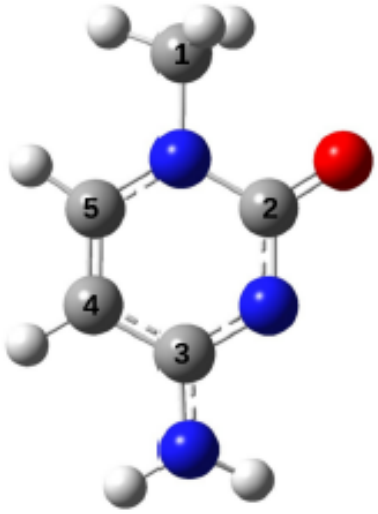
Fragment-Based Chemical Shift Predictions

Fragment-based NMR calculations were used to predict the ^{13}C chemical shifts of neutral **1-methylcytosine**, **protonated 1-methylcytosine triflate salt**, **1-Iodide**, and **1-Iodide-*d*$_5$**. These NMR predictions were performed by Joshua Hartman from Prof. Gregory Beran's group at UCR. NMR predictions were performed at PBE0/6-311+G(2d,p) and B3LYP/6-311+G(2d,p) with hydrogen-

only optimizations utilizing Crystal09 and all-atom fixed-cell optimization utilizing Quantum Espresso.

The published crystal structure of neutral 1-methylcytosine can be found in literature[39]. **Table 2.1** shows the comparison of ^{13}C chemical shifts between experimental and calculation. The predicted ^{13}C NMR chemical shifts of neutral 1-methylcytosine uses the published crystal structure geometry without further refinement and the experimental ^{13}C chemical shifts were obtained from solid state NMR experiments. The experimental values match that of predictions well with discrepancies within 5 ppm. With the comparison between calculations and experiment of 1-methylcytosine in good agreement, we turn our attention to protonated 1-methylcytosine triflate salt and proton-bound dimer salts of 1-methylcytosine. Because many of the peaks for the triflate salt of 1-methylcytosine are doubled, it seems reasonable to suspect that nonequivalent molecules within the unit cell give rise to distinct ssNMR signals.

Table 2.1: Comparison between experimental and calculated ssNMR ^{13}C chemical shifts of neutral 1-methylcytosine obtained from fragment-based NMR calculations.



Geom. Opt.: Functional	Expt	H-only PBE0	H-only B3LYP	All atom PBE0
1	39.7	39.44	39.01	39.40
2	157.2	153.41	153.42	153.97
3	168.4	162.46	162.42	163.79
4	94.7	90.30	89.58	92.71
5	147.2	144.92	144.26	147.97

Coordinates from protonated 1-methycytosine triflate salt were obtained from X-Ray diffraction and used as an initial starting point to predict its ^{13}C chemical shifts. The ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt shows two singlets in the domain where the methyl resonances are expected (**Figure 2.13**).

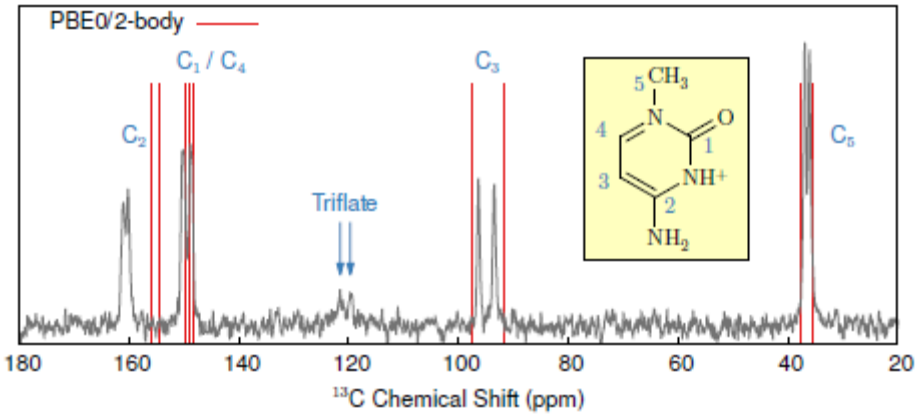
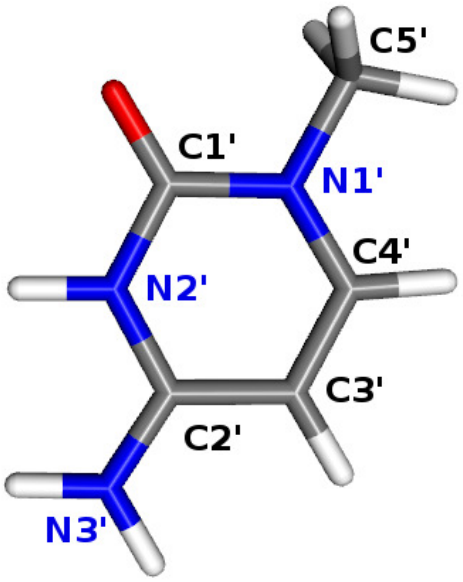


Figure 2.13: Experimental ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt plotted over calculated ^{13}C chemical shifts (red lines).

Initially, we thought this was due to the formation of PBDs since two methyl singlets are observed in the ^{13}C ssNMR data of **1-Iodide**. X-Ray diffraction, however, shows two molecules of protonated 1-methylcytosine triflate salt per unit cell. A methyl resonance from each protonated 1-methylcytosine triflate salt could explain why two methyl resonances are observed. **Table 2.2** compares the experimental and predicted ^{13}C chemical shifts in a table format. The largest discrepancy is about 5 ppm.

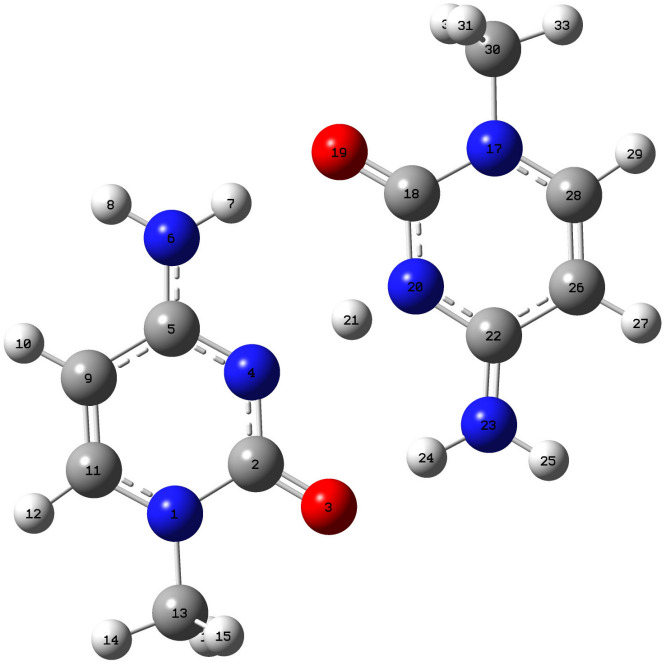
Table 2.2: Comparison between experimental and calculated ssNMR ^{13}C chemical shifts of protonated 1-methylcytosine triflate salt obtained from fragment-based NMR calculations.



Carbon num.	Exp.	PBE0 6-311+G(2d,p) 2-Body 6 Å	Diff.
C1a	149.95	148.40	1.55
C2a	160.92	155.85	5.07
C3a	96.31	97.66	1.36
C4a	148.73	148.83	1.13
C5a	36.02	35.53	0.49
C1b	149.95	148.22	0.51
C2b	160.12	154.47	5.65
C3b	93.43	91.74	1.70
C4b	148.73	149.70	0.97
C5b	36.89	37.81	0.93

X-ray diffraction of **1-Iodide** also shows two PBD iodide salts per unit cell. **Table 2.3** summarizes the predicted and experimental ssNMR ^{13}C chemical shifts. The ^{13}C chemical shifts were calculated based on the hydrogen placements and the largest of the discrepancy is approximately 8 ppm.

Table 2.3: Comparison between experimental and calculated ^{13}C chemical shifts of **1-Iodide** Habit B obtained from fragment-based NMR calculations.

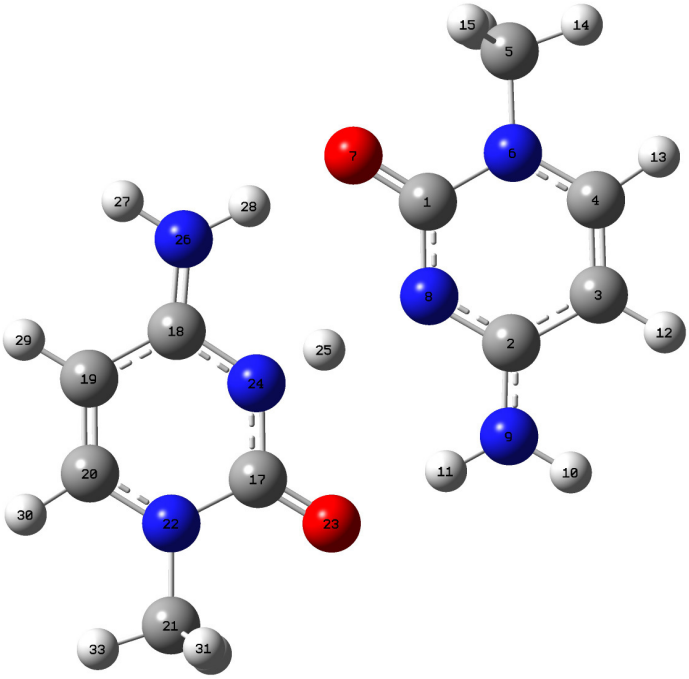


	Exp.	Avg. Scaled	Difference
C2	155	154.43	0.57
C5	164.2	159.89	4.31
C11	98.4	94.06	4.34
C9	148.7	150.13	-1.43
C13	40.8	45.68	-4.88
C18	155	154.17	0.83
C22	163.7	158.56	5.14
C28	93.6	98.43	-4.83
C26	150	149.64	0.36
C30	39.5	43.02	-3.52

The experimental chemical shifts are matched with predicted values based on the position of the chemical shifts.

Like **1-Iodide**, X-ray diffraction analysis of **1-Iodide-d_5** shows two PBDs occupying one unit cell. The predicted ^{13}C ssNMR spectrum of **1-Iodide-d_5** is reproduced in **Table 2.4**

Table 2.4: Comparison between experimental and calculated ^{13}C chemical shifts of **1-Iodide-d_5** obtained from fragment-based NMR calculations.



Carbon num.	Exp.	Dimer
C1	159.89	156.73
C2	160.00	161.15
C3	95.41	93.17
C4	148.37	149.45
C5	36.95	41.14
C17	151.34	151.19
C18	160.00	159.74
C19	95.41	92.08
C20	158.41	151.57
C21	39.49	47.91

A large discrepancy (approximately 7 ppm) belongs to the carbon (#20) at the 6-position. The values for the experimental are matched with their corresponding carbon resonances based on chemical shift.

The question is whether doubling of many of the ssNMR ^{13}C resonances of solid salts of hemiprotonated 1-methylcytosine are due to unsymmetrical placement of the proton or whether they represent different resonances of nonidentical molecules within the unit cell. Experimental results from the triflate salt of protonated 1-methylcytosine, which also shows doubling of ssNMR ^{13}C and ^{15}N resonances suggests that the latter alternative is the case.

Gas Phase Experiments

Chart 1 depicts the six ions investigated in the gas phase by Aaron Moehlig. Ions **1** – **3** are various proton-bound homodimers while ions **4** – **6** are proton-bound heterodimers. The IRMPD spectrum of ion **1** was previously investigated in the fingerprint domain from 400 – 1800 cm^{-1} [31], and the IRMPD spectrum showed the disappearance of a single band at 1570 cm^{-1} when all of the exchangeable hydrogens were replaced by deuterium. Ions **2** and **3** are expected to have a band shift around the same domain since ions **2** and **3** are also proton-bound homodimers. Ions **4** – **6** differ in the sense that the bridging proton is not shared equally between the two bases due to their difference in proton affinities. This results in the appearance of more than one band between 1500 – 1600 cm^{-1} domain which will be discussed in greater detail.

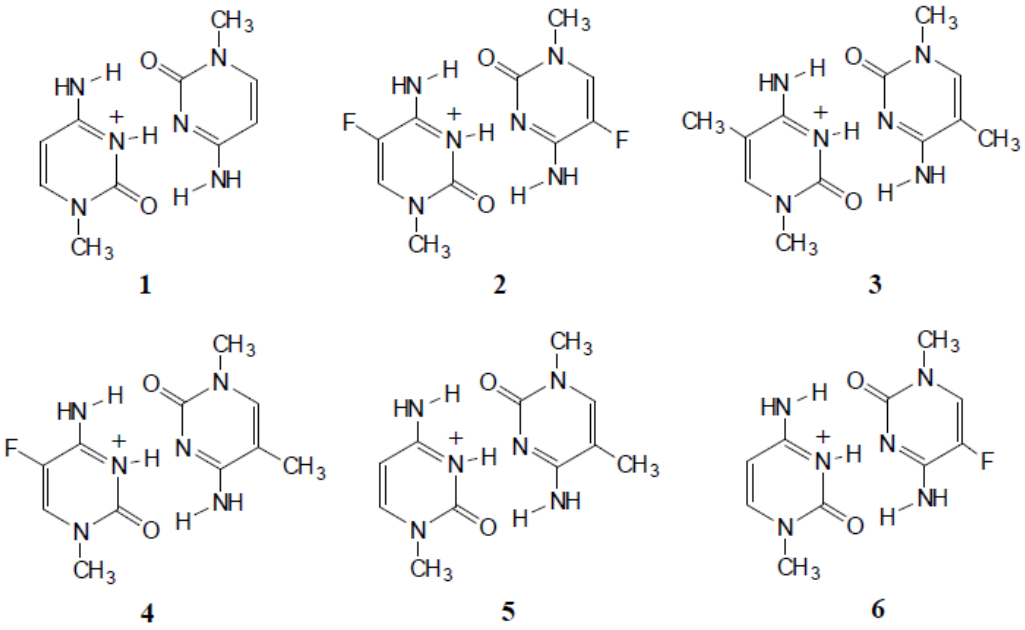


Chart 1: 1-Methylcytosine homo- and heterodimers investigated using IRMPD.

The use of IRMPD spectroscopy of ions in the gas phase was described earlier. Briefly, ions are generated via electrospray ionization of a 1 mM solution in water/methanol with a small amount of acetic acid. IRMPD spectra are obtained via action spectroscopy, in which an ion with a specific mass to charge ratio (*m/z*) is isolated using an FT-ICR mass spectrometer. The ion is then irradiated with either a fixed frequency or a tunable infrared laser for several seconds and the resulting fragments plotted against the scanned laser frequency, generating an experimental infrared spectrum of the gaseous ion. The isolated ions were isolated and irradiated in the CH/NH/OH domain from 2450 – 3800 cm^{-1} using used a LaserVision benchtop optical parametric oscillator (OPO) laser. IRMPD spectra were obtained by plotting formation of charged fragments

as well as the depletion of the parent ion as the parent ion is being irradiated as function of the scanned wavelength. DFT calculations can be utilized to predict the most stable conformation as well as predict infrared spectra. Comparison of predicted and experimental infrared spectra provides a clue to how ions interact in a solvent-free environment.

Before discussing results pertaining to ions **1** – **6**, the gas phase IR spectum of protonated 1-methylcytosine monomer in the fingerprint and CH/OH/NH stretching domain is revisited. The comparison between calculated anharmonic normal modes with experimental IRMPD spectra in the fingerprint and CH/NH/OH stretching domain of protonated 1-methylcytosine are reproduced in **Figure 2.14 and 2.15**. Anharmonic stretching modes and intensities for both *N3*-protonated and *O*-protonated 1-methylcytosine monomer were calculated at the B3LYP/aug-CC-PVTZ level. **Figure 2.14** shows the calculated anharmonic stretches in the fingerprint domain from 600 – 2000 cm^{-1}, while **Figure 2.15** depicts the calculated stretches in the CH/NH/OH stretching domain from 2700 – 3800 cm^{-1} of both *N3*-protonated and *O*-protonated 1-methylcytosine monomers. **Figure 2.14C** and **2.15C** were obtained at the FOM Institute for Plasma Physics by Dr. Giel Berden. A fixed-frequency CO_2 laser pulse following the scanned laser pulse was used for additional excitation.

Comparison of the calculated anharmonic stretching modes with experimental IRMPD spectra agrees with our previous assignments. The fingerprint domain looks to be predominately *N3*-protonated (comparison

between **Figure 2.14B** and **2.14C**), while the CH/NH/OH stretching domain looks to be *O*-protonated (comparison between **Figure 2.15A and 2.15**). In the fingerprint stretching domain, a band is observed at 1790 cm^{-1} in the experimental spectrum that is only predicted to be present in the *N3*-protonated monomer. The black trace and red trace looks to be almost identical in **Figure 2.14C**. The black trace represents the experimental IRMPD spectrum of protonated 1-methylcytosine obtained from ESI, while the red trace represents the IR spectrum obtained from dissociation of its PBD. The experimental IR spectra look to be identical whether it was obtained from direct electrospray or from dissociation of a PBD.

The CH/NH/OH stretching domain of protonated 1-methycytosine displays at least three bands at 3400, 3500, and 3550 cm^{-1}. The *O*-protonated monomer looks like it best fits this spectrum because the calculated spectrum of *N3*-protonated 1-methylcytosine shows only two bands in this domain. Overtones and combination bands are also observed in the CH/NH/OH stretching domain at 3210, 3240, and 3310 cm^{-1}. We infer that both tautomers of protonated 1-methylcytosine must be present in the gas phase.

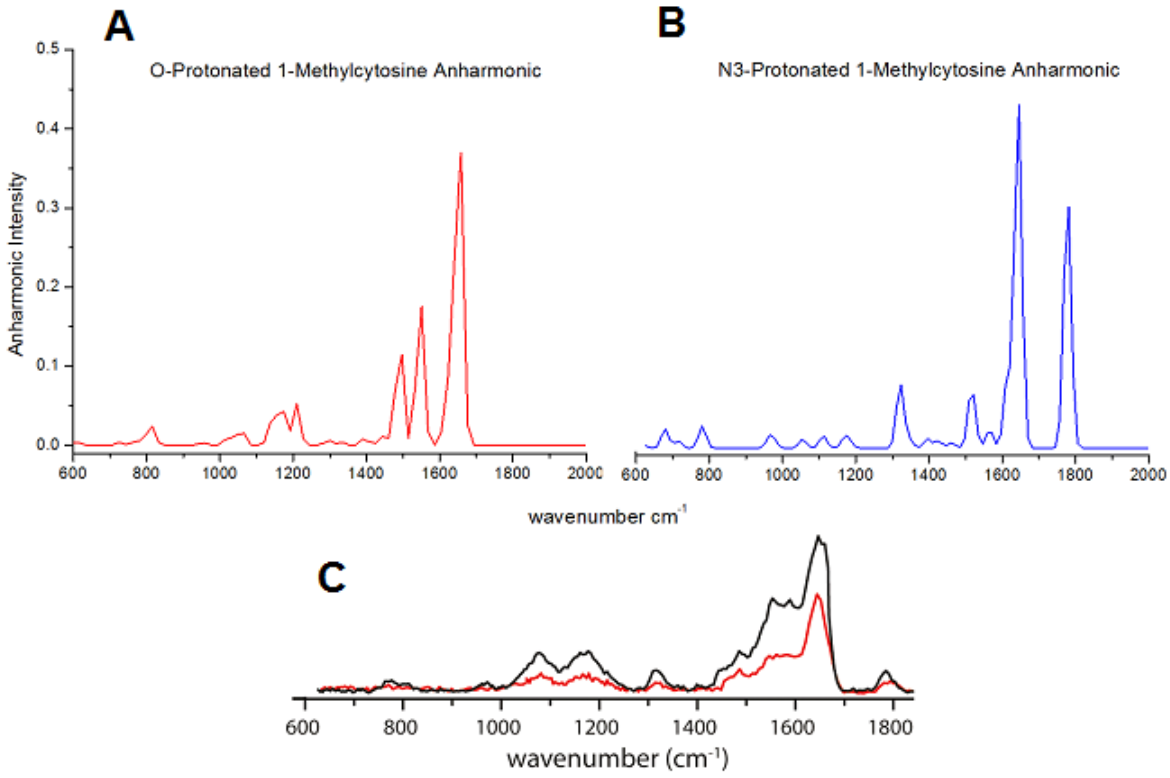


Figure 2.14: **A)** Calculated anharmonic stretching modes of *O*-protonated 1-methylcytosine at the B3LYP/aug-cc-pVTZ level in the fingerprint domain from 600 – 2000 cm^{-1}. **B)** Calculated anharmonic stretching modes of *N3*-protonated 1-methylcytosine at the same level in the fingerprint domain from 600 – 2000 cm^{-1}. **C)** Experimental IRMPD spectra of protonated 1-methylcytosine in the fingerprint stretching domain. The black trace depicts the generation of protonated 1-methylcytosine ion from dissociation of the proton-bound dimer and the red trace shows its generation using electrospray ionization.

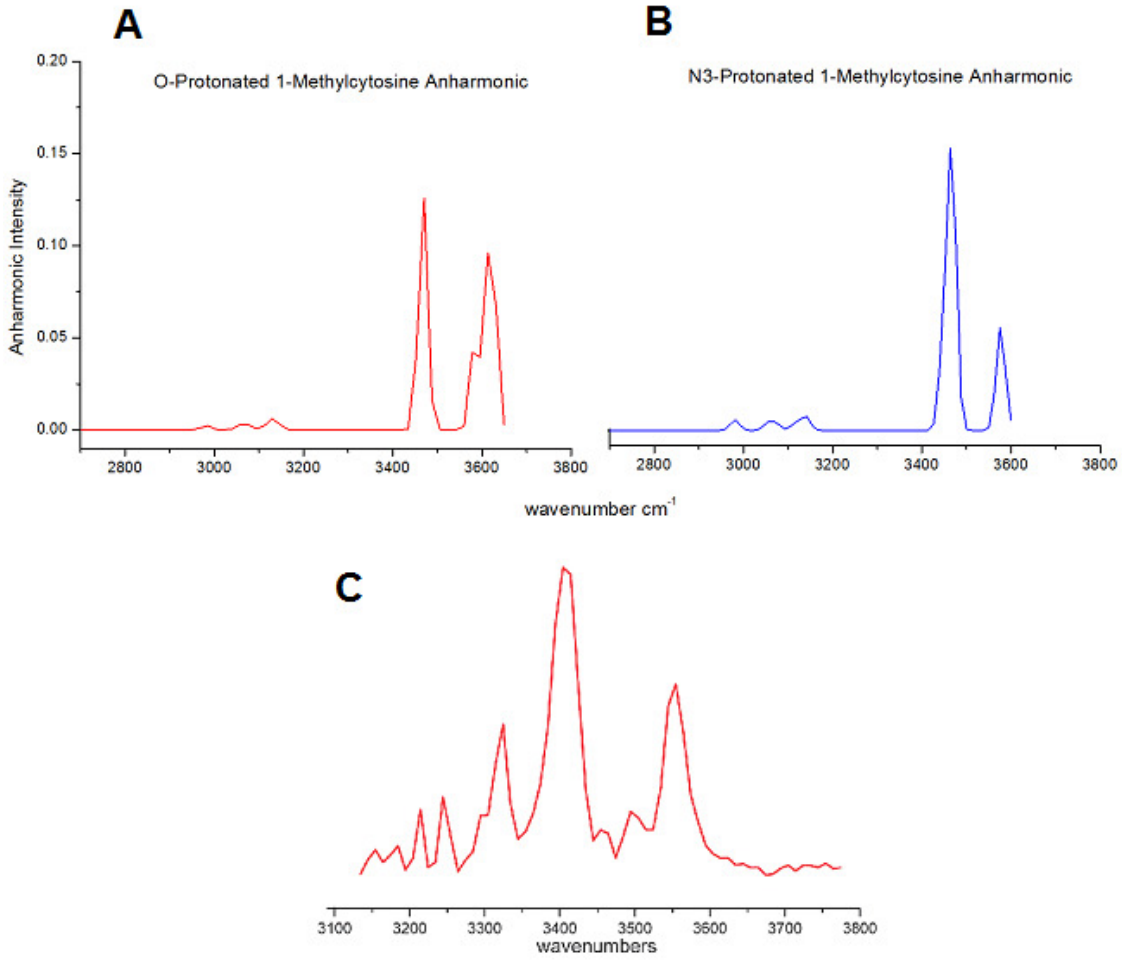


Figure 2.15: **A)** Calculated anharmonic stretching modes of *O*-protonated 1-methylcytosine at the B3LYP/aug-cc-pVTZ level in the CH/NH/OH stretching domain from 2700 – 3800 cm^{-1}. **B)** Calculated anharmonic stretching modes of *N3*-protonated 1-methylcytosine at the same level in the CH/NH/OH stretching domain from 2700 - 3800 cm^{-1}. **C)** Experimental IRMPD spectrum of protonated 1-methylcytosine in the CH/NH/OH stretching domain from 3100 – 3800 cm^{-1}.

A previous publication on the vibrational signature of the bridging proton of 1-methylcytosine PBD iodide salt assigned the band at 1890 cm^{-1} to the motion of the bridging proton, but that domain is inaccessible to IRMPD[30]. Our more recent

comparison between calculated harmonic normal modes with experimental IRMPD spectra and deuteration replacement experiments of 1-methylcytsine PBD revealed a band in the fingerprint domain at 1570 cm^{-1} neither predicted by theory nor in the deuterated spectrum[31]. Deuteration experiments, where all of the exchangeable protons are exchanged for deuterium, caused the 1570 cm^{-1} band to vanish, which leads to the assignment of that band to a motion associated with the bridging proton. Solid powder IR as well as single crystal IR of 1-methylcytosine PBD iodide salt show the same band at 1570 cm^{-1}, further endorsing our assignment.

The gas phase IRs of other proton-bound homo- and proton-bound heterodimers of 1-methylcytosine and 1-methylcytosine derivatives have been investigated. **Figure 2.16** depicts the experimental IRMPD spectra of the PBD of 1-methylcytosine in the CH/NH/OH domain. Panel A reproduces the IR spectrum without the use of partial deuteration experiments and Panel B shows the results from employing partial deuteration experiments. Partial deuteration experiments help deconvolve the spectrum. They also help eliminate overtones and combination bands. The result displays five distinct bands that can be attributed to be N-H stretches and the C-H stretch of the methyl group.

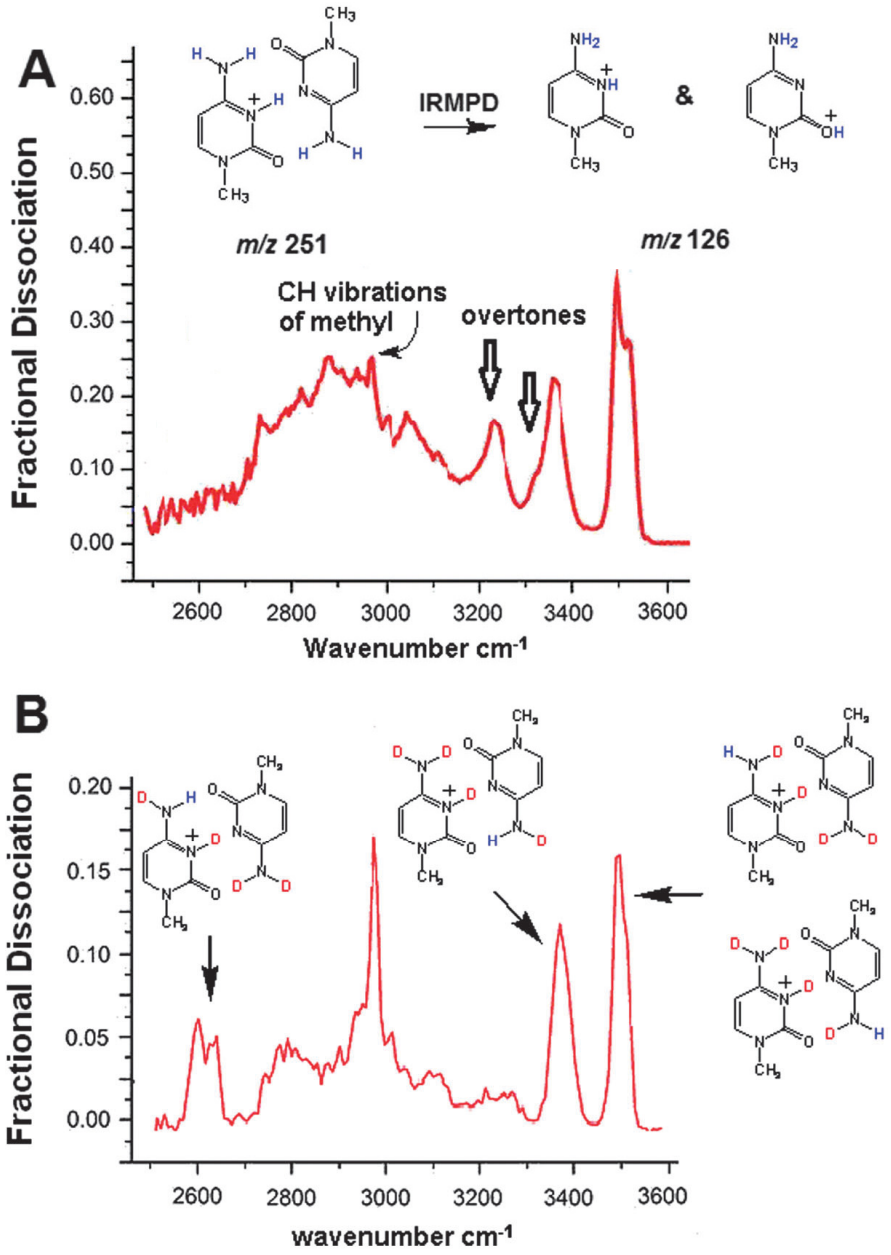


Figure 2.16: **A)** Experimental IRMPD spectra of the PBD of 1-methylcytosine in the functional group domain from 2500 – 3600 cm^{-1}. Two overtone bands are observed around 3250 and 3370 cm^{-1}. **B)** Partial deuteration experiments was performed where four of the five exchangeable hydrogens are exchanged for deuterium. Doing so helps deconvolve the spectrum and eliminate overtones and combination bands.

A broad cluster of bands is observed in the CH/NH/OH stretching domain between 2700 – 3000 cm^{-1} in **Figure 2.16A**. The methyl resonance appears clearly after conducting partial deuteration experiments where four of the five exchangeable hydrogens are exchanged for deuterium. To ensure the band at 2950 cm^{-1} is indeed a methyl C-H stretch, a separate experiment was done where CH_3 group is replaced with a CD_3 group. The IRMPD spectrum of the incompleted deuterated ion **1-d_{10}** is shown in **Figure 2.17B**, and the disappearance of the methyl C-H stretches confirms our assignments.

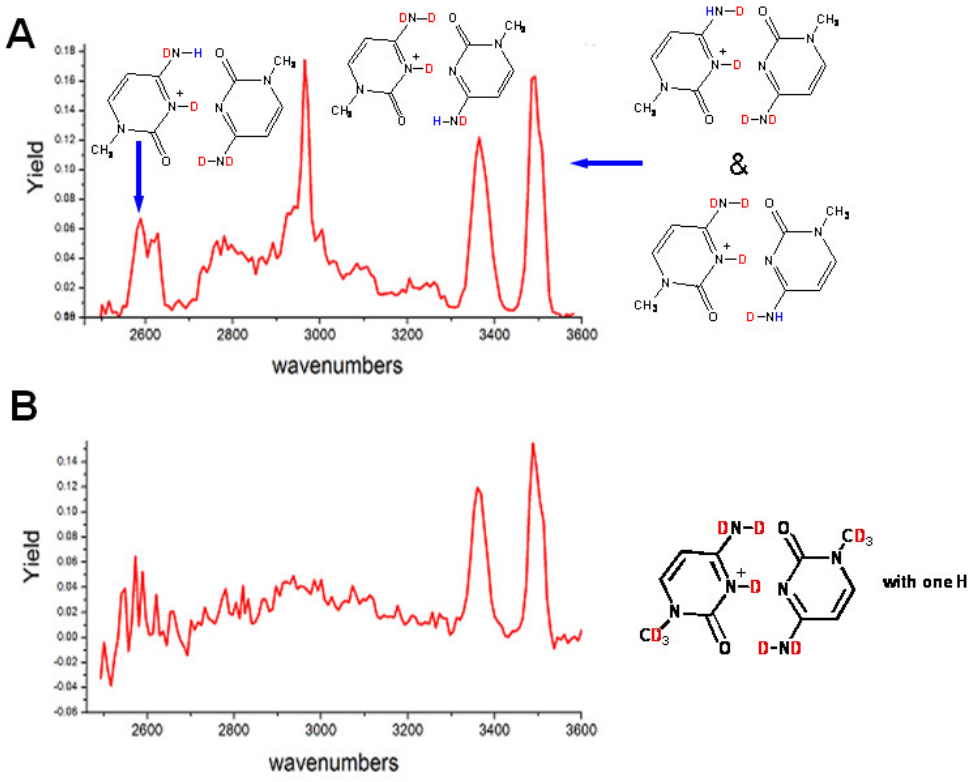


Figure 2.17: **A**) Experimental IRMPD spectrum of ion **1-d_4** in the CH/NH/OH domain from 2600 – 3600 cm^{-1}. **B**) Experimental IRMPD spectrum of the incomplete deuterated ion **1-d_{10}** in the CH/NH/OH domain from 2600 – 3800 cm^{-1}.

Figure 2.18 compares the IRMPD spectra of ion **2** with DFT predictions calculated at the B3LYP/6-31G** level. A band present at 1580 cm^{-1} does not coincide with theory. Like that of ion **1**, this band appears around the same domain. A deuteration experiment was performed on ion **2** where all five of the exchangeable hydrogens are exchanged for deuterium. **Figure 2.19** shows the comparison between experimental IRMPD and DFT predicted spectra of ion **2-d_5**. The experimental spectrum does not show a band at 1580 cm^{-1}. That band is therefore assigned to be the asymmetric stretch associated with the bridging proton. Several other bands are observed between 3200 – 3300 cm^{-1}. Those bands are assigned to be overtones of fundamental bands, just like that observed in **Figure 2.16A**.

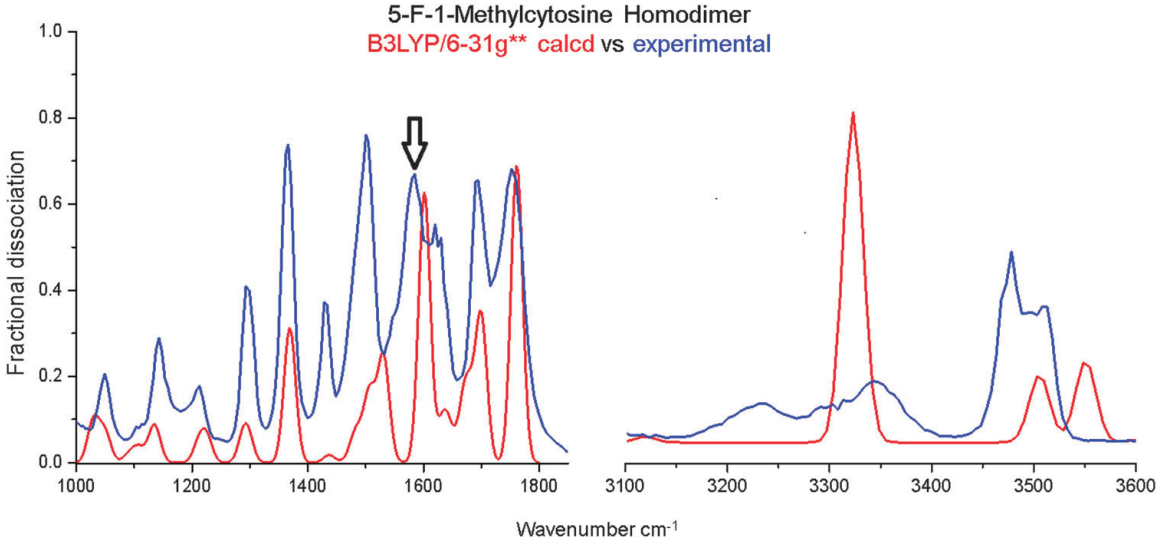


Figure 2.18: Experimental IRMPD spectra of **2** in the fingerprint domain (blue trace in the left hand panel) using a free electron laser and in the CH/NH/OH stretching domain (blue trace in the right hand panel) using an OPO laser. The red traces, in both panels, corresponds to the calculated scaled IR spectra at the B3LYP/6-31G** level. The arrow indicates the motion of the bridging proton that is not predicted by theory.

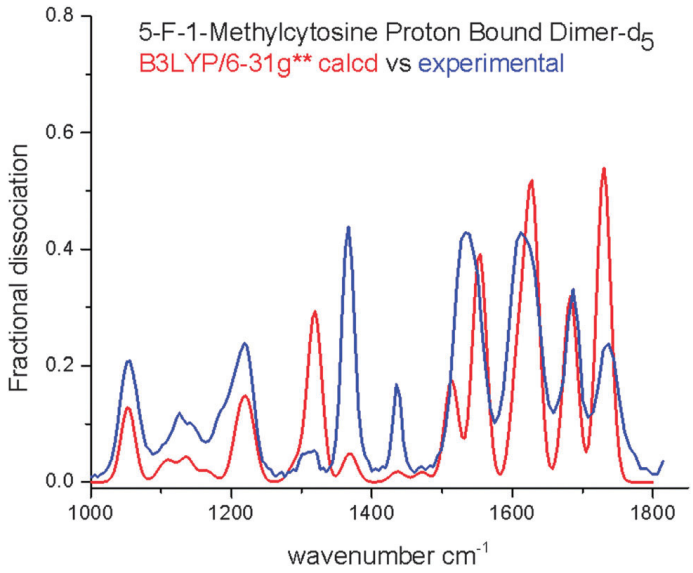


Figure 2.19: Experimental IRMPD spectrum plotted against theory calculated at the B3LYP/6-31G** level for ion **2-*d*$_5$**. The blue trace corresponds to the experimental IRMPD spectrum while the red trace corresponds to the scaled DFT predicted spectrum.

Figure 2.20 shows IR comparison between theory and experiment of ion **3** and the **3-*d*$_5$**. The comparison between experimental and theory of ion **3** and **3-*d*$_5$** show a good fit of peaks. It is apparent that all three protonated homodimers (ions **1**, **2** and **3**) all show an extra band not predicted by theory in the 1500 – 1600 cm^{-1} domain. Just like protonated ions **1** and **2**, the extra band around 1580 cm^{-1} in **Figure 2.20A** is assigned to be the N-H$^{+\cdots}$N stretch associated with the bridging proton. All three proton-bound homodimers exhibit a band in a similar domain. Would we expect proton-bound heterodimers to behave in the same fashion?

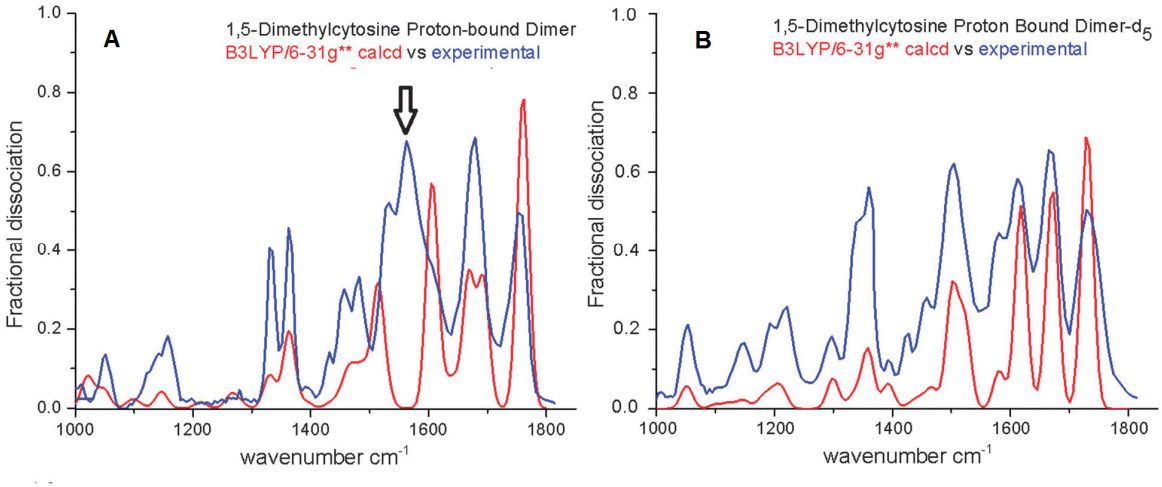


Figure 2.20: **A)** Experimental IRMPD spectrum of ion **3** (blue) compared with scaled DFT calculated at the B3LYP/6-31G** level (red) in the fingerprint domain from 1000 – 1800 cm^{-1}. **B)** Experimental IRMPD spectrum of ion **3-d_5** (blue) compared with scaled DFT calculated at the B3LYP/6-31G** level (red) in the fingerprint domain from 1000 – 1800 cm^{-1}.

Figure 2.21 shows the gas phase IR between 1-methylcytosine and 5-fluoro-1-methylcytosine (ion **4**). The comparison between experimental and predicted IR spectra show a good fit of peaks, except for two bands around 1550 and 1560 cm^{-1} (black arrows). As previously seen in the case of proton-bound homodimers, one band was assigned to be the motion of the bridging proton. For proton-bound heterodimers, more than are observed due to differences in proton affinity. The proton either resides on one nucleobase or the other, depending on their proton affinities, which results in two different N-H stretching resonances. The IRMPD spectrum of ion **4-*d_5*** was not obtained, but the IRMPD spectrum of the **d_5** analog would be expected to show the disappearance of the two aforementioned bands at 1550 and 1560 cm^{-1}.

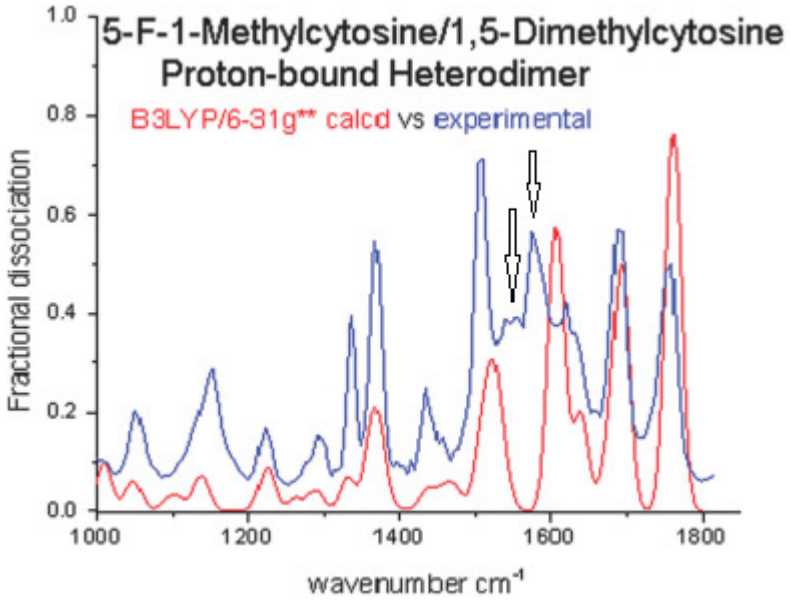


Figure 2.21: Comparison between experimental and calculated IRMPD spectra of ion **4** (5-fluoro-1-methylcytosine and 1,5-dimethylcytosine heterodimer) in the fingerprint domain from 1000 – 1800 cm^{-1}. The experimental spectrum is represent by the blue trace while scaled calculation is shown using a red trace. Two bands are present in the experimental spectrum around 1550 and 1560 cm^{-1} (black arrows) that is not predicted by theory.

The comparison between experimental and predicted IRMPD spectra of ion **5** (**Figure 2.22**) again shows two bands not predicted by theory around 1570 and 1580 cm^{-1} (indicated by arrows). Just like ion **4**, those two bands are motions associated with the bridging proton. Looking at the CH/NH/OH stretching domain, there is also an extra band between 3200 – 3300 cm^{-1} that is not predicted by calculations. This band is assigned to be an overtone, similar to that observed previously with ion **1**. Like ion **4**, the two bands at 1570 and 1580 cm^{-1} are assigned to be motions associated with the bridging proton of the two tautomers. Complete deuteration experiments were not performed on this ion, but if

deuteration experiments were to be conducted, the three bands around 1570, 1580, and between 3200 - 3300 cm^{-1} would be expected to disappear.

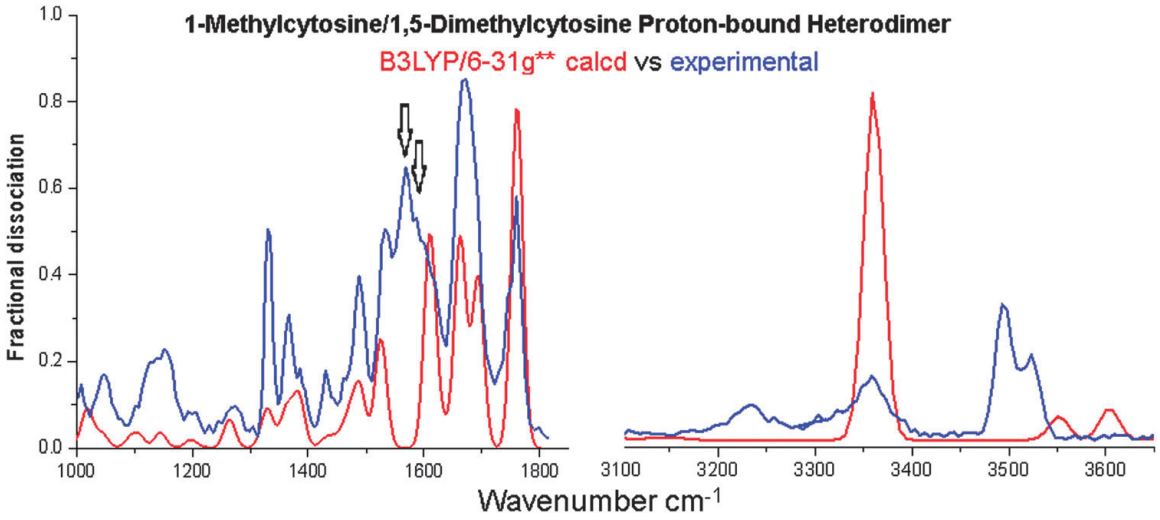


Figure 2.22: Comparison of experimental (blue trace) vs calculated (red trace) IRMPD spectra of ion **5** (1-methylcytosine and 1,5-dimethylcytosine heterodimer) in the fingerprint and CH/NH/OH domain. Two bands at 1570 and 1580 cm^{-1} are observed experimentally, but not predicted by calculations. An overtone band is also present between 3200 – 3300 cm^{-1}.

The comparison between experimental and DFT predicted IRMPD spectra of ion **6** is reproduced in **Figure 2.23**. The fingerprint domain shows a good fit of peaks except for two bands around 1570 and 1580 cm^{-1} (black arrows). The difference in proton affinity between 1-methylcytosine and 5-fluoro-1-methylcytosine is responsible for the two bands that are not present in the calculated spectrum, but which are present in the experimental spectrum. The difference in proton affinity also causes the proton to prefer one nucleobase over the other, and the two different N-H stretches give rise to two different bands. Complete deuteraton experiments were not performed on ion **6**.

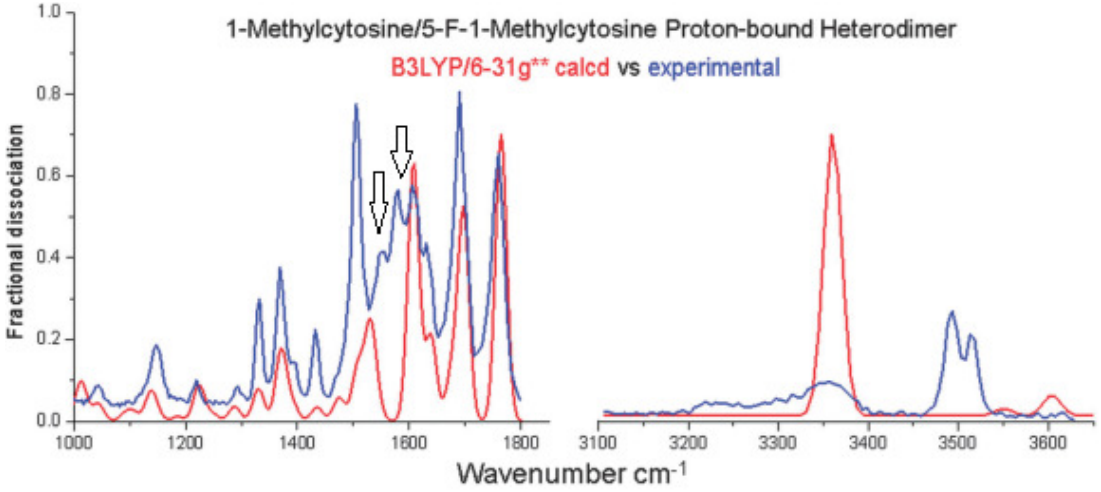


Figure 2.23: Comparison between experimental (blue trace) and calculated (red trace) IRMPD spectra of ion 6 in the fingerprint and CH/NH/OH domain. Two bands are observed around 1560 and 1580 cm-1 are observed experimentally, but not predicted by theory. An overtone band is also observed experimentally between 3200 – 3300 cm^{-1}.

The experimental IRMPD spectrum of 1-methylcytosine protonated monomer was plotted against scaled harmonic theory calculated at B3LYP/cc-pVTZ level, while the proton-bound homo- and heterodimers were calculated at the scaled harmonic B3LYP/6-31G** level. Anharmonic vibrations and intensities were calculated for 1-methylcytosine protonated monomer while the proton-bound homo-and heterodimers looked at harmonic vibrations and intensities. In general, the fingerprint domain shows a good fit of peaks for all spectra, while the functional group domain show a slight mismatch of the N-H stretching resonances.

Tautomer assignment of 1-methylcytosine protonated monomer agrees with that previously published[40]. Two tautomers of protonated 1-methylcytosine

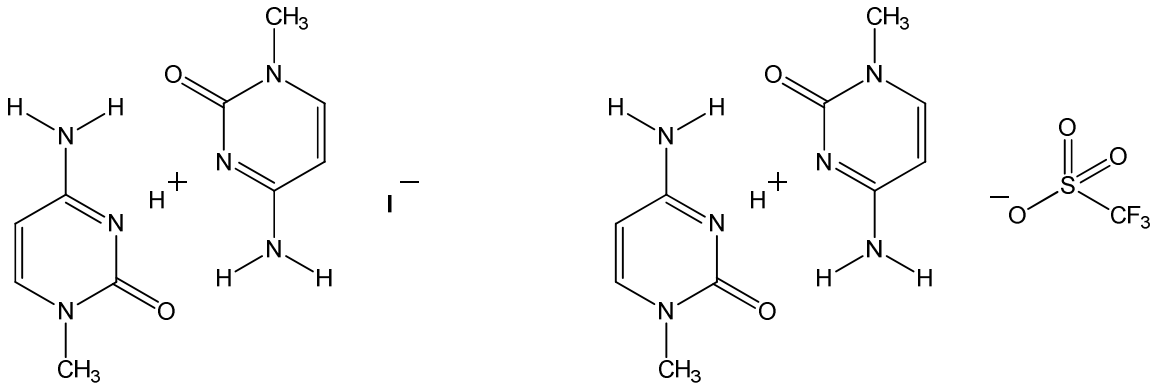
exist in the gas phase as the fingerprint domain of protonated 1-methylcytosine looks to be *N3*-protonated, while the CH/NH/OH stretching domain looks to be *O*-protonated. Our anharmonic calculations agree with previously published harmonic calculations for 1-methylcytosine protonated monomer.

All three proton-bound homodimers (ions **1**, **2**, and **3**) show a band between 1500 – 1600 cm^{-1} that disappears upon deuteration. This band is not predicted by theory and is assigned to be the motion associated with the bridging proton. Complete deuteration experiments, where all of the five exchangeable hydrogens are exchanged for deuterium, help deconvolve the spectra by eliminating bands associating with the exchangeable hydrogens. After doing so, the band between 1500 – 1600 cm^{-1} disappeared, suggesting it is associated to be the motion of the bridging proton. The CH/NH/OH stretching domain (between 3200 – 3300 cm^{-1}) shows an extra band not predicted by theory. The CH/NH/OH domain of **1** showed an extra band that was determined to be an overtone of a fundamental band. Since ions **2** and **3** both have a band in the same domain, that extra band is also assigned to be an overtone of a fundamental band.

Comparison of the experimental IRMPD spectra of proton-bound heterodimers (ions **4**, **5**, and **6**) with theory also shows a good fit of peaks in the fingerprint domain. The CH/NH/OH domain also showed the same extra band between 3200 – 3300 cm^{-1} and mismatch of the N-H resonances. Like the proton-bound homodimers, bands are present between 1500 – 1600 cm^{-1} that are assigned to be motions associated with the bridging proton. Unlike the

proton-bound homodimers, there is more than one band observed. This phenomenon is due to the different proton affinity of the two nucleobases. This allows the proton to be shared unequally, and requires a higher barrier for proton transit. Although complete deuteration experiments were not performed on the proton-bound heterodimers, the two bands between 1500 – 1600 cm^{-1} are expected to disappear if the aforementioned experiments were conducted.

Solid State Experiments



1-Iodide **1-Triflate**

Figure 2.24: Two 1-methylcytosine proton-bound dimer salts of interest.

1-Iodide

Solid state NMR experiments will focus on two PBD salts shown in **Figure 2.24**. **1-Iodide** was prepared as described in the experimental section of this chapter. **1-Triflate** was prepared in the same manner except triflic acid is used in

place of hydroiodic acid. Both methods require several rounds of recrystallization using absolute ethanol to obtain crystals pure enough for ssNMR experiments. Recrystallization of **1-Iodide** from slow evaporation of absolute ethanol overnight yield crystals exhibiting crystal habit **B**. Cooling the hot ethanolic solution too quickly (using an ice bath) produces colorless flake-like crystals which exhibits crystal habit **A**.

The ^{1}H ssNMR spectrum of **1-Iodide** shows five resonances and is reproduced in **Figure 2.25**. The largest peak near 3.5 ppm corresponds to the methyl resonance and the two peaks around 7.5 and 8.5 ppm corresponds to the hydrogens at the 5- and 6-position of the rings. The broader peak near 10 ppm is assigned to be the hydrogens on the external amino group. The peak around 16 ppm corresponds to the *N3*-bridging hydrogen, which matches well with DFT prediction and previously reported *i*-motif forming oligomers[19,20].

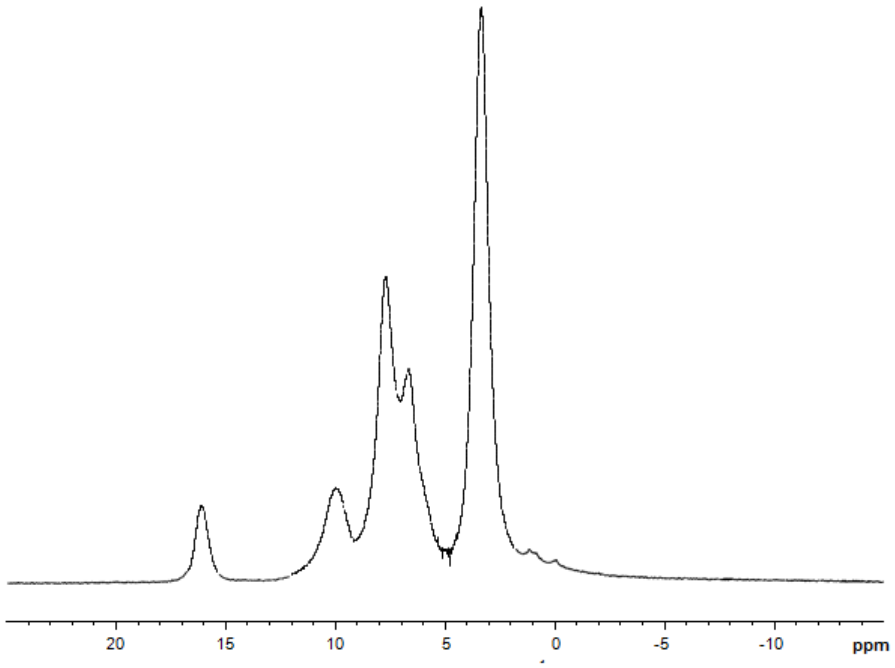


Figure 2.25: Experimental ^{1}H ssNMR spectrum of **1-Iodide**.

The five observed bands in the ^{1}H ssNMR spectrum suggests **1-Iodide** is symmetrical, but results from ^{15}N and ^{13}C ssNMR suggest otherwise. If the bridging proton was to reside directly in the middle between two cytosine residues, it should render all of the carbons and nitrogen equivalent. Five bands are observed in the ^{15}N ssNMR spectrum and eight bands are observed in the ^{13}C ssNMR spectrum, suggesting the bridging proton either preferentially associates with one of the cytosine ring or the other or that two PBDs per unit cell exhibit difference resonances. **Table 2.5** shows the comparison of chemical shifts between neutral, protonated iodide salt, and **1-Iodide**. Assignments of each carbon and nitrogen resonances of **1-Iodide** crystal habit **B** are reproduced in **Figure 2.26**.

	1-MeCytosine	1-MeCytosineH^+ I^-	(1-MeCytosine)$_2H^+$ I^-
Carbon			
Methyl	40.0	42.8	39.5, 40.8
C5 vinyl	94.9	92.6	93.6, 98.4
C6 vinyl	147.5	148.3	148.7, 150.0
Carbonyl	157.5	152.8	155.0
C–NH_2	168.7	159.8	163.7, 164.2
Nitrogen			
N–Me	139.7	144.0	135.2, 136.4
H_2N	93.6	101.2	102.6, 107.7
Ring N	207.8	147.3	169.7

Table 2.5: ^{13}C and ^{15}N chemical shifts (ppm) of neutral 1-methylcytosine, protonated 1-methylcytosine iodide salt, and 1-methylcytosine PBD iodide salt (**1-Iodide**). All peaks are referenced using external standards. Adamantane referenced to TMS was used for carbon, and ammonium chloride referenced to liquid ammonia was used for nitrogen.

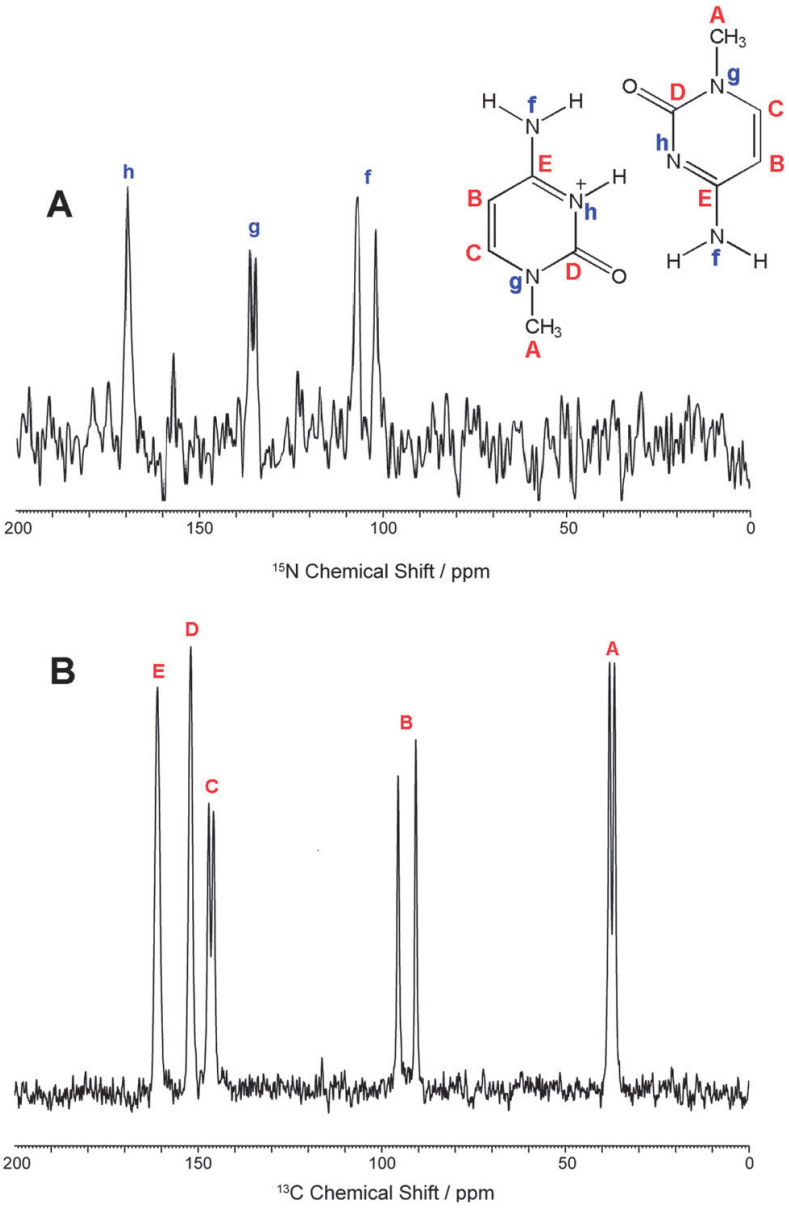


Figure 2.26: Experimental ^{15}N (panel **A**) and ^{13}C (panel **B**) ssNMR spectra of **1-Iodide**.

Our assignment of the carbon and nitrogen spectra was confirmed by a carbon-nitrogen dipolar coupling performed by Ryan Kudla. The ring nitrogens of 1-methylcytosine were partially labeled starting from $^{15}N_2$-urea. $^{15}N_2$-1-

methylcytosine was synthesized as described in the experimental section along with its partially labeled PBD iodide salt. Like that of unlabeled **1-Iodide**, the ^{13}C spectrum of **$^{15}N_2$-1-Iodide** shows eight carbon resonances. Only two resonances are observed in the ^{15}N spectrum because the external nitrogen is not ^{15}N labeled. Similar to that of the unlabeled **1-Iodide**, the number of carbon and nitrogen resonances observed suggests the bridging proton does not reside in the middle between two cytosine residues, but on one cytosine or the other.

Figure 2.27 depicts the outcome of a nitrogen-carbon dipolar coupling experiment performed on **$^{15}N_2$-1-Iodide**. Nitrogens g and h are both ^{15}N labeled in this experiment. Nitrogen g couples to carbons A, C, and D while nitrogen h couples to carbons D and E. This experiment confirms our previous ^{13}C and ^{15}N spectra assignments. Calculations put the external N-H distances at 1.057 and 1.784 Å. The experimental N-H distances match almost perfectly at 1.056 and 1.784 Å.

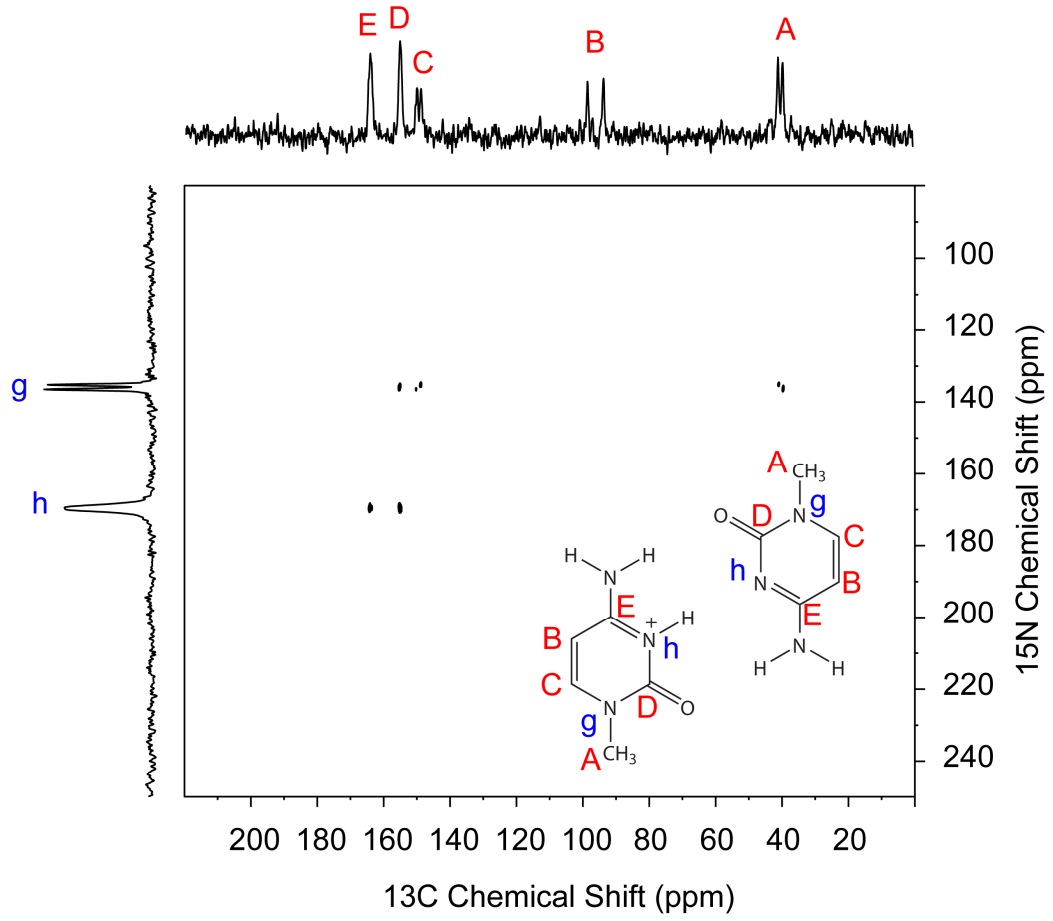


Figure 2.27: A 2-Dimensional ^{15}N - ^{13}C dipolar coupling experiment depicting the coupling between adjacent carbon and nitrogen atoms of **$^{15}N_2$-1-Iodide**. The ^{13}C ssNMR spectrum is on the x-axis, while the ^{15}N ssNMR spectrum is plotted on the y-axis. Coupling between nitrogens and carbons are depicted with black spots.

Deuteration of the exchangeable hydrogens of **1-Iodide** has a profound effect on its ^{13}C chemical shifts. **Figure 2.28** compares the ^{13}C ssNMR spectra of **1-Iodide** (red trace) with **1-Iodide-*d*$_5$** (blue trace). The smaller bands with stars indicate first and second order spinning side bands. These side bands result from spinning the solid sample at a certain speed when the spectrum is being obtained. The methyl resonances of **1-Iodide** and **1-Iodide-*d*$_5$** do not directly

overlap and some bands that are split in the **1-Iodide** spectrum are overlapping in the **1-Iodide-*d*$_5$** spectrum.

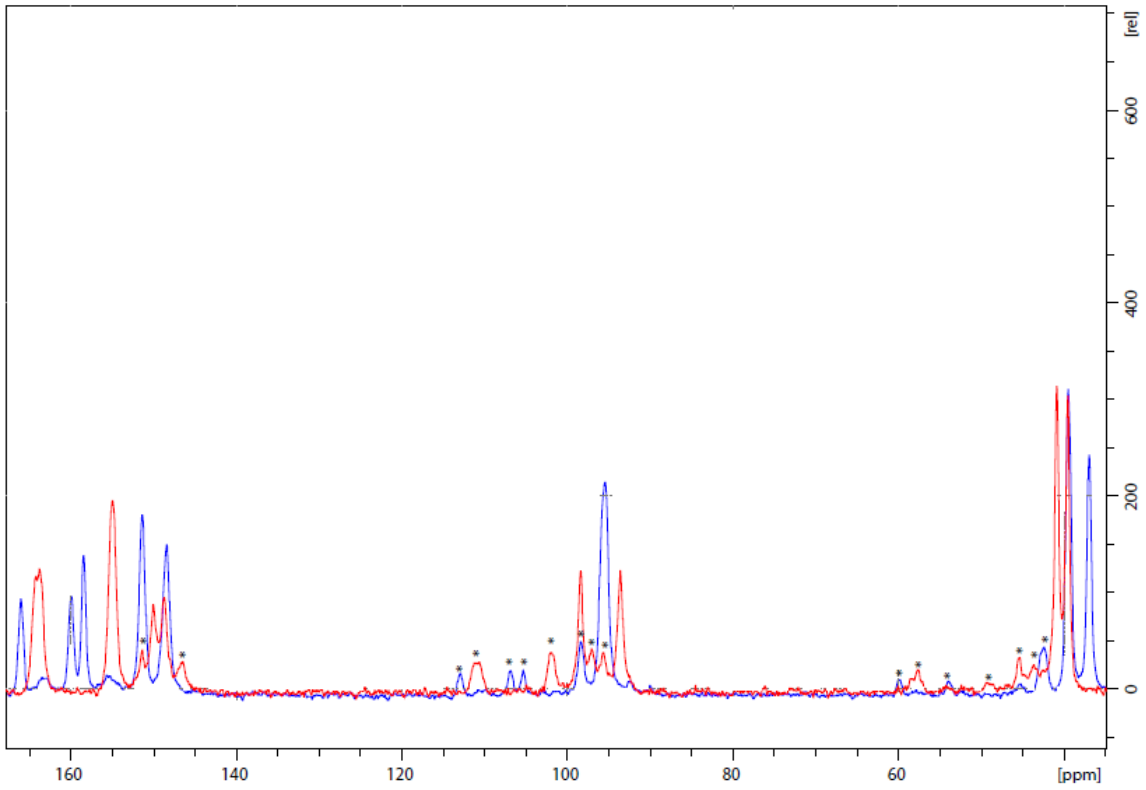


Figure 2.28: ^{13}C ssNMR comparison between **1-Iodide** (red trace) with **1-Iodide-*d*$_5$** (blue trace).

Due to the occupation of two sets PBD iodide salt per unit cell, the carbon resonances associated with methyl carbons are drawn into question. Do these two bands arise from differences within the PBD, or between the two sets of PBD within the unit cell? In attempt to answer this question, we turn to fragment-based chemical shift calculations (performed by Joshua Hartman) and the experimental ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt.

As shown in **Table 2.3**, two different sets of chemical shifts are calculated for **1-Iodide**. Comparison between prediction and experimental ^{13}C chemical shifts matches well except for the methyl carbons. Dimer A's predicted methyl shifts are off by approximately 6 ppm, while Dimer B's predicted methyl shifts are off by approximately 11 ppm. The distance between the two methyl resonances is 1.3 ppm experimentally and calculated to be 1.7 and 1.3 ppm for Dimers A and B, respectively.

Like **1-Iodide**, X-ray diffraction of protonated 1-methylcytosine triflate salt also shows two salt units occupying one unit cell. **Figure 2.29** shows the experimental ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt. The two methyl resonances in **Figure 2.29** are separated by 0.8 ppm. Since protonated 1-methylcytosine triflate salt only contains one methyl carbon per monomer unit, the two resonances observed experimentally must come from the two individual methyl groups. The separation between methyl resonances in the ^{13}C experimental spectrum of **1-Iodide** is larger than protonated 1-methylcytosine triflate salt at 1.3 ppm, suggesting the two methyl resonances arise from the methyl groups within the PBDs rather than between PBDs.

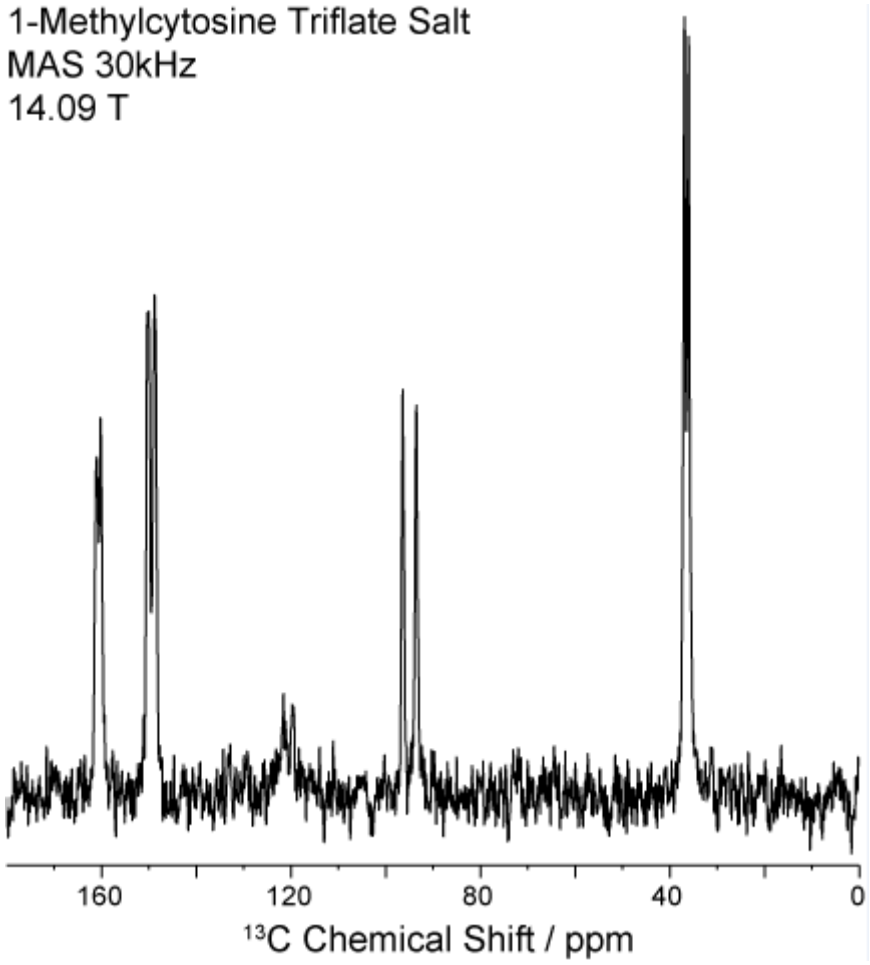


Figure 2.29: Experimental ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt.

1-Triflate

The ^{13}C ssNMR spectrum of **1-Triflate** is reproduced in **Figure 2.30**. The two methyl resonances of **1-Triflate** are clustered together and appear to be one broad peak, whereas **1-Iodide** showed two separate peaks. Another notable difference between **1-Triflate** and **1-Iodide** is the peak associated with the carbonyl carbon at 163 – 164 ppm. That band seems to be split in the case of **1-Triflate**, but is overlapping in the case of **1-Iodide**. The difference in chemical

shifts and splitting between **1-Triflate** and **1-Iodide** can be attributed to the packing of the crystal structure.

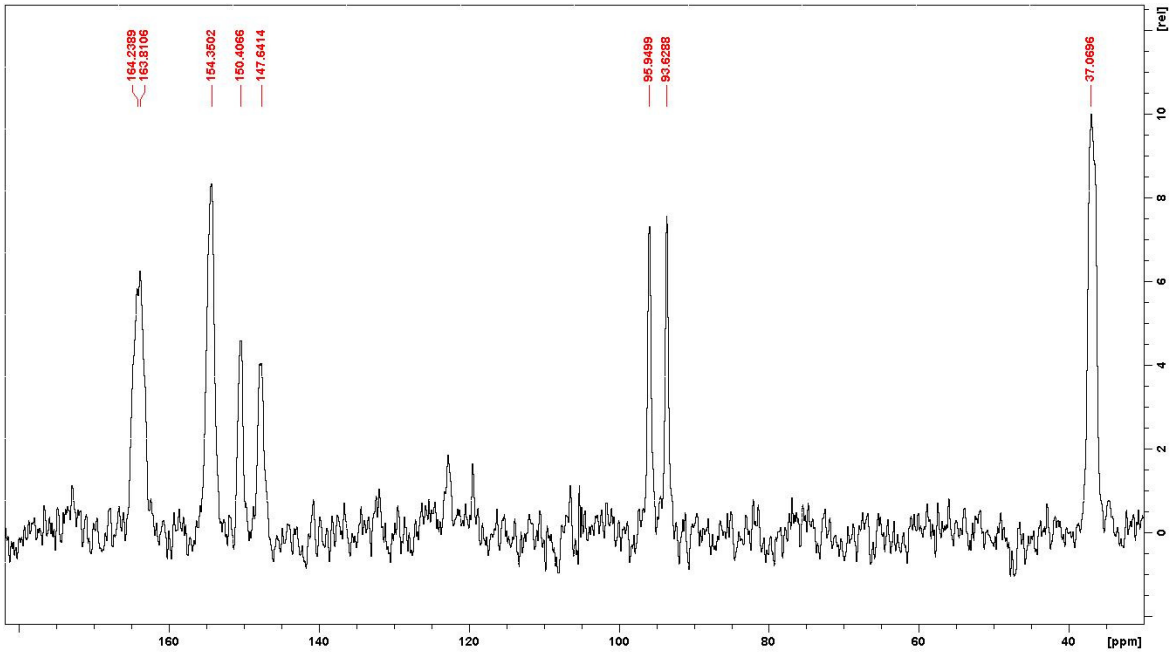


Figure 2.30: Experimental ^{13}C ssNMR spectrum of **1-Triflate**.

Solid-State Infrared Spectroscopy

The IR signature corresponding to the bridging proton of ion **1** appears at 1570 cm^{-1} in the gas phase[31]. The IR signatures of other proton-bound homodimers (ion **2** and **3**) also show a band around the same domain. When the proton affinities of the bases differ, there are two bands of unequal intensities that appear between 1550 and 1600 cm^{-1} in the gas phase (ions **3**, **4**, and **5**). Solid proton-bound heterodimer salts were not prepared due to the inability to control salt formation. A portion of the sample will contain proton-bound homodimers and another portion will have proton-bound heterodimers.

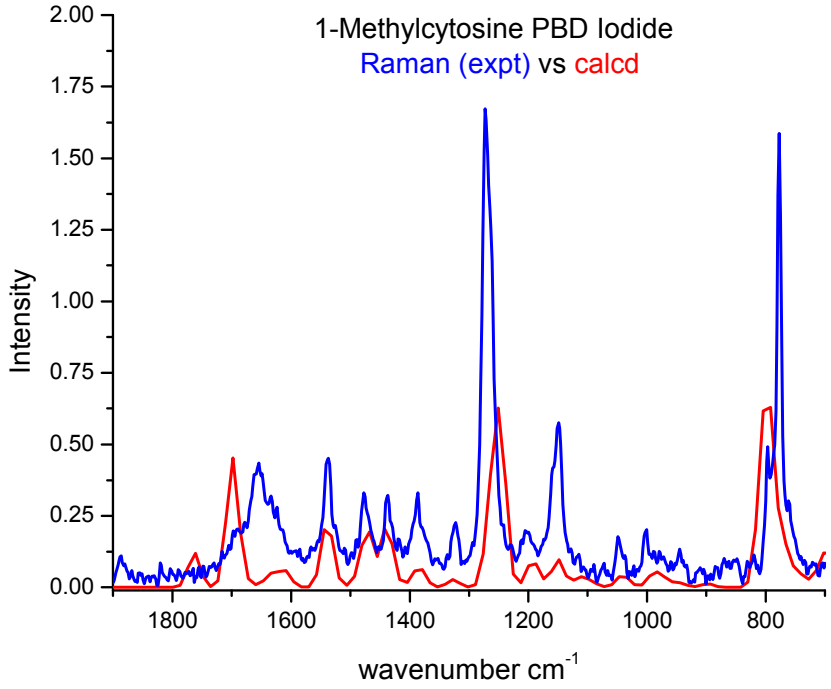


Figure 2.31: Comparison of experimental (1064 nm exciting line, blue) Raman spectrum plotted against a scaled gas phase prediction calculated at B3LYP/6-311++G** (red) of **1-Iodide**.

Figure 2.31 compares the experimental Raman spectrum with prediction of **1-Iodide**. The experimental spectrum matches the predicted spectrum, except for a slight mismatch between 1600 – 1700 cm^{-1}. **Figure 2.32** depicts two separate IR (orange and red lines) and a raman (blue line) spectra of **1-Iodide**. All three spectra in **Figure 2.32** is of **1-Iodide habit B**. The scale is cut off at 2200 cm^{-1} in the yellow line spectrum because it looked identical to the red line. A

close inspection of the yellow and red line shows a band present at 1570 cm^{-1}, just like that of the gas phase spectrum.

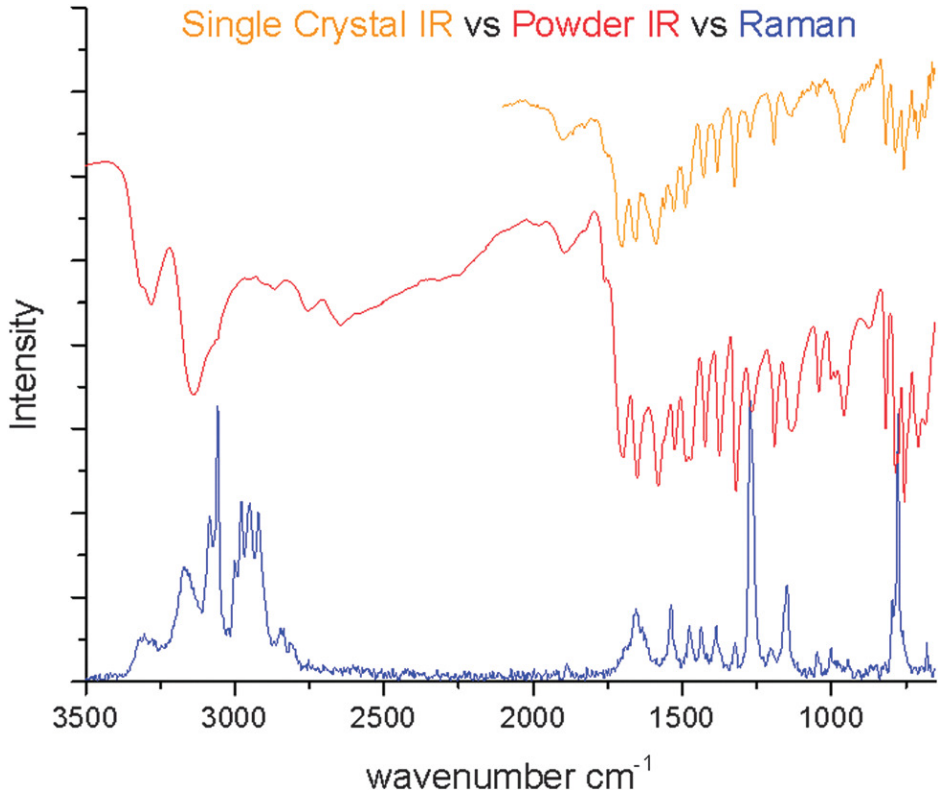


Figure 2.32: Single crystal IR of **1-Iodide-B** (orange) plotted against powder IR spectrum of a mixture of crystals containing forms A and B (red). FT-Raman of the same powder is shown in blue (1064 nm exciting line).

The previously reported IR on **1-Iodide** places the N-H stretching of the bridging proton at 1890 cm^{-1}. **Figure 2.33** shows the comparison between **1-Iodide** and **1-Iodide-d_5**. All of the exchangeable hydrogens are exchanged with deuterium in the d_5 analog, and the blowup of both spectra show the band at 1570 cm^{-1} disappearing (red arrow). The two black arrows show the presence of the band at 1890 cm^{-1} that was previously reported to be the motion of the bridging proton. Since that band did not shift upon deuteration, it therefore cannot

be the bridging N-H stretch. Deuteration also causes a cluster of bands originally around 3000 cm^{-1} to shift to 2300 cm^{-1}.

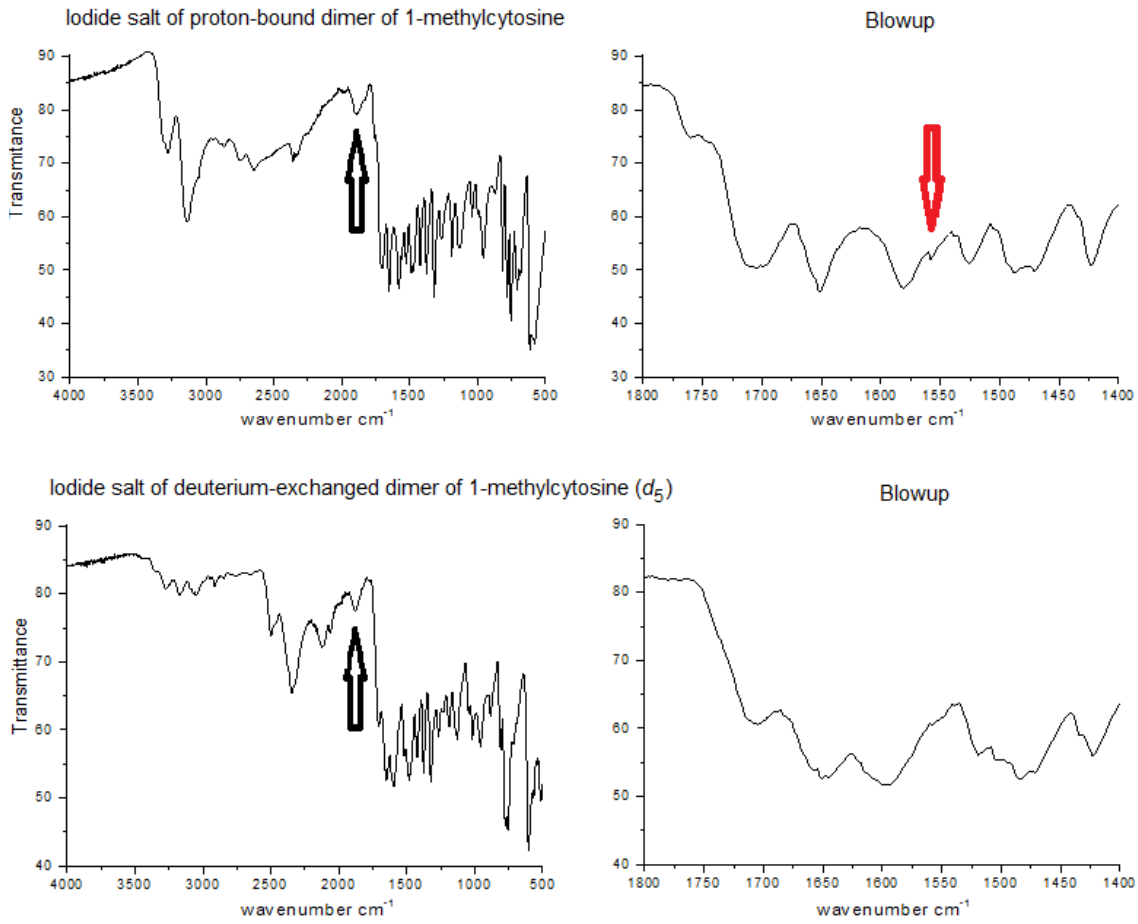


Figure 2.33: FT-IR spectra of **1-Iodide** and **1-Iodide-*d*$_5$** deuterium-exchanged isotopomer. Black arrows indicate the 1890 cm^{-1} band present in both spectra. Red arrow in the top blowup indicates the band that coincides with the 1570 cm^{-1} band observed by IRMPD for gaseous ion **1**.

1-Triflate was synthesized by another group member, Jay-Ar Bendo, following the same procedure for synthesizing **1-Iodide**. The powder IR spectrum of **1-Triflate** is shown in **Figure 2.34** and **1-Triflate-*d*$_5$** in **Figure 2.35**. Similar to

the IR spectrum of **1-Iodide**, a band is observed around 1890 cm^{-1}. This band is also observed in the spectrum of **1-Triflate-*d*$_5$** so it is unlikely to be the motion of the briding proton. Looking at the fingerprint domain, there is a band present at 1581 cm^{-1} in the spectrum of **1-Triflate**. That band is no longer present in the spectrum of **1-Triflate-*d*$_5$**.

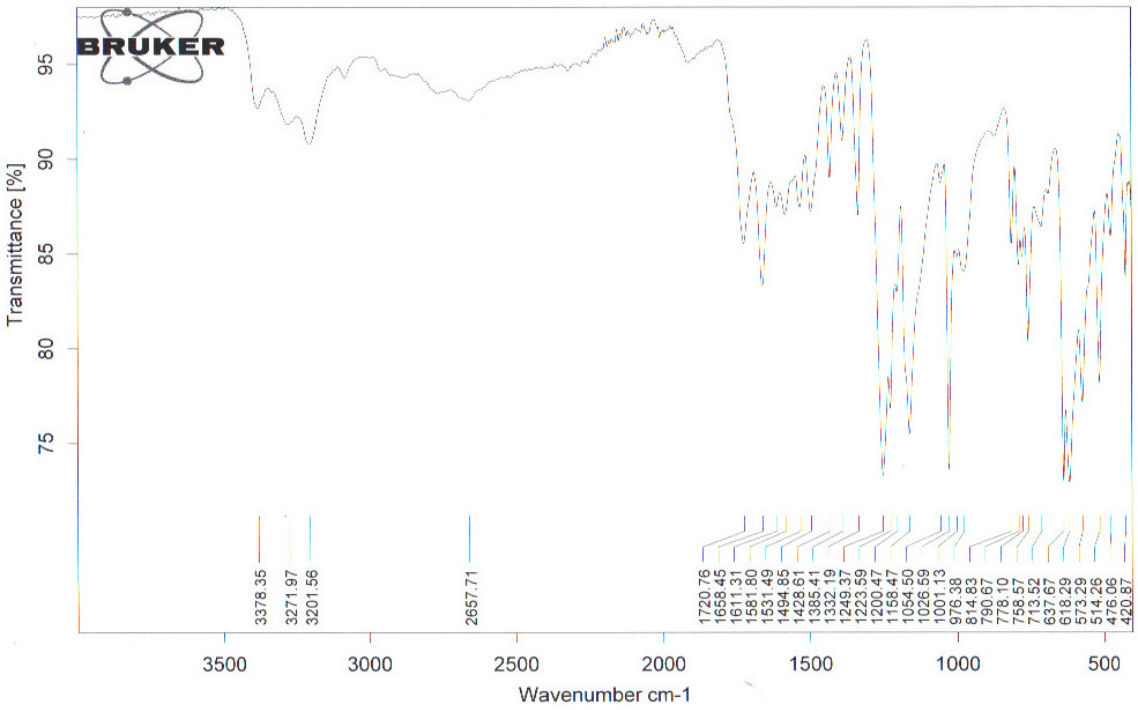


Figure 2.34: Powder IR spectrum of **1-Triflate**.

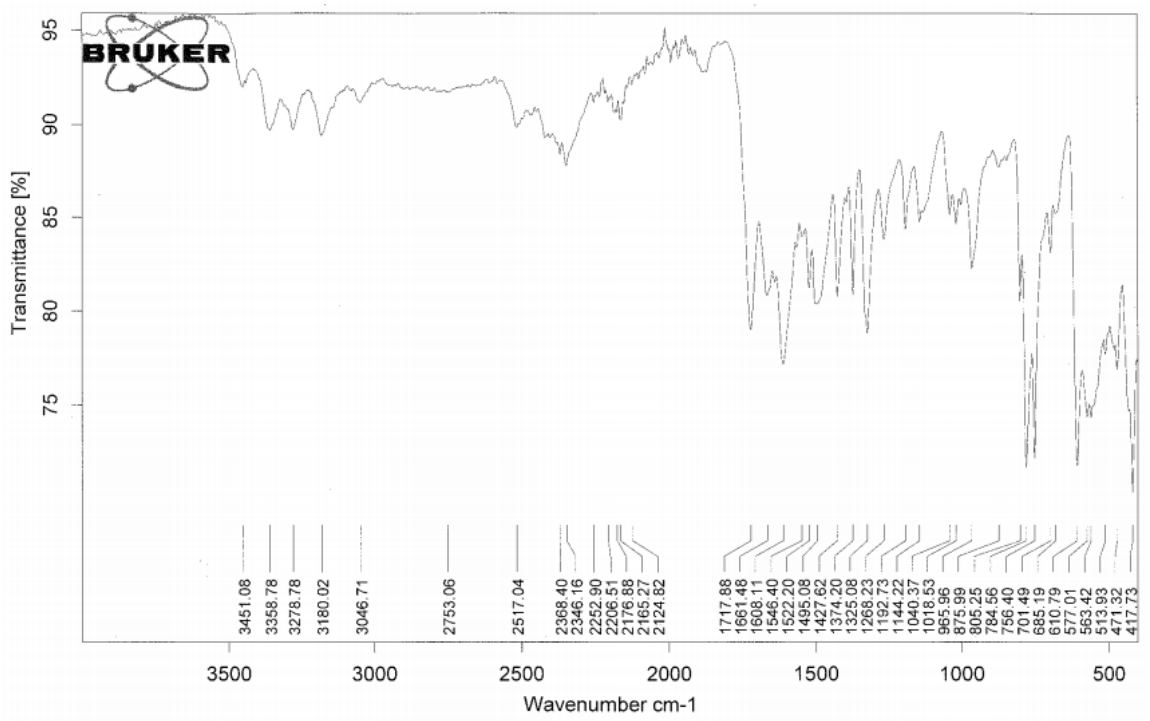


Figure 2.35: Powder IR spectrum of **1-Triflate-d_5**.

Inelastic Neutron Scattering

INS spectra were obtained at the Los Alamos Neutron Scattering Center (LANSCE) and at the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory with the assistance of Dr. Luke Daemen. Neutrons are generated by the collision of accelerated protons against a heavy-metal target, such as tungsten or uranium. At LANSCE, neutrons enter an instrument called the Filter Difference Spectrometer (FDS), where it collides with the sample, and pass through filters as they exit. Neutrons that interact with the sample are scattered, and are sorted by their time of flight (TOF) from when they enter to when they reach the detector. As the incoming neutrons interact with the sample, energy is

transferred in the form of molecular vibrations. The time it takes for the neutron to reach the detector is related to the energy lost by the neutron. A schematic representation of the FDS is shown in **Figure 2.36**.

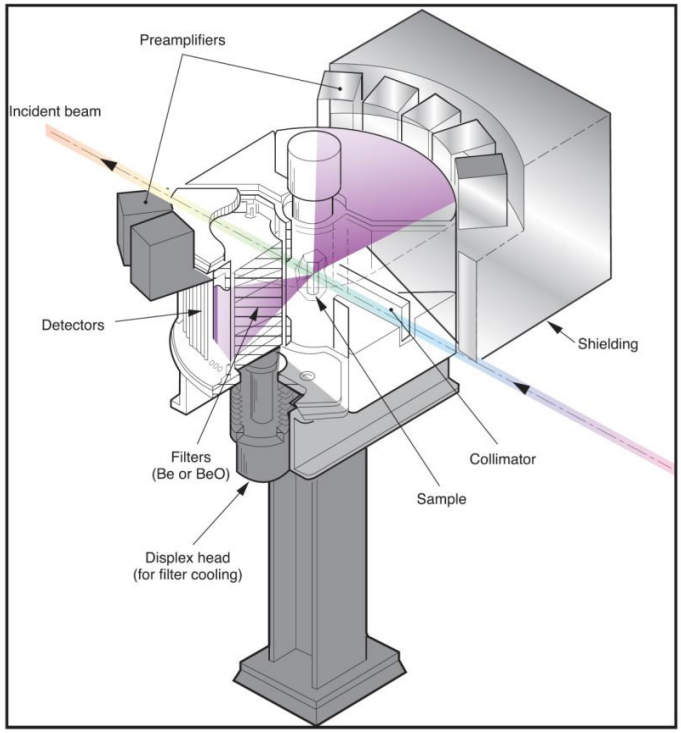


Figure 2.36: Schematic representation of FDS at LANSCE. Schematic was obtained from Los Alamos National Laboratory website.

Molecules with high numbers of hydrogens are ideal when performing INS experiments. Hydrogen has a scattering cross section approximately 80 times that of deuterium and more than 160 times greater than that of nitrogen. A drawback associated INS is its sensitivity to impurities. Any impurity containing hydrogen will distort and produce false bands in the experimental spectrum. All salts designated for INS experiments were recrystallized many times to remove

any impurities. A chart depicting the compounds selected for INS experiments is shown in **Chart 2**.

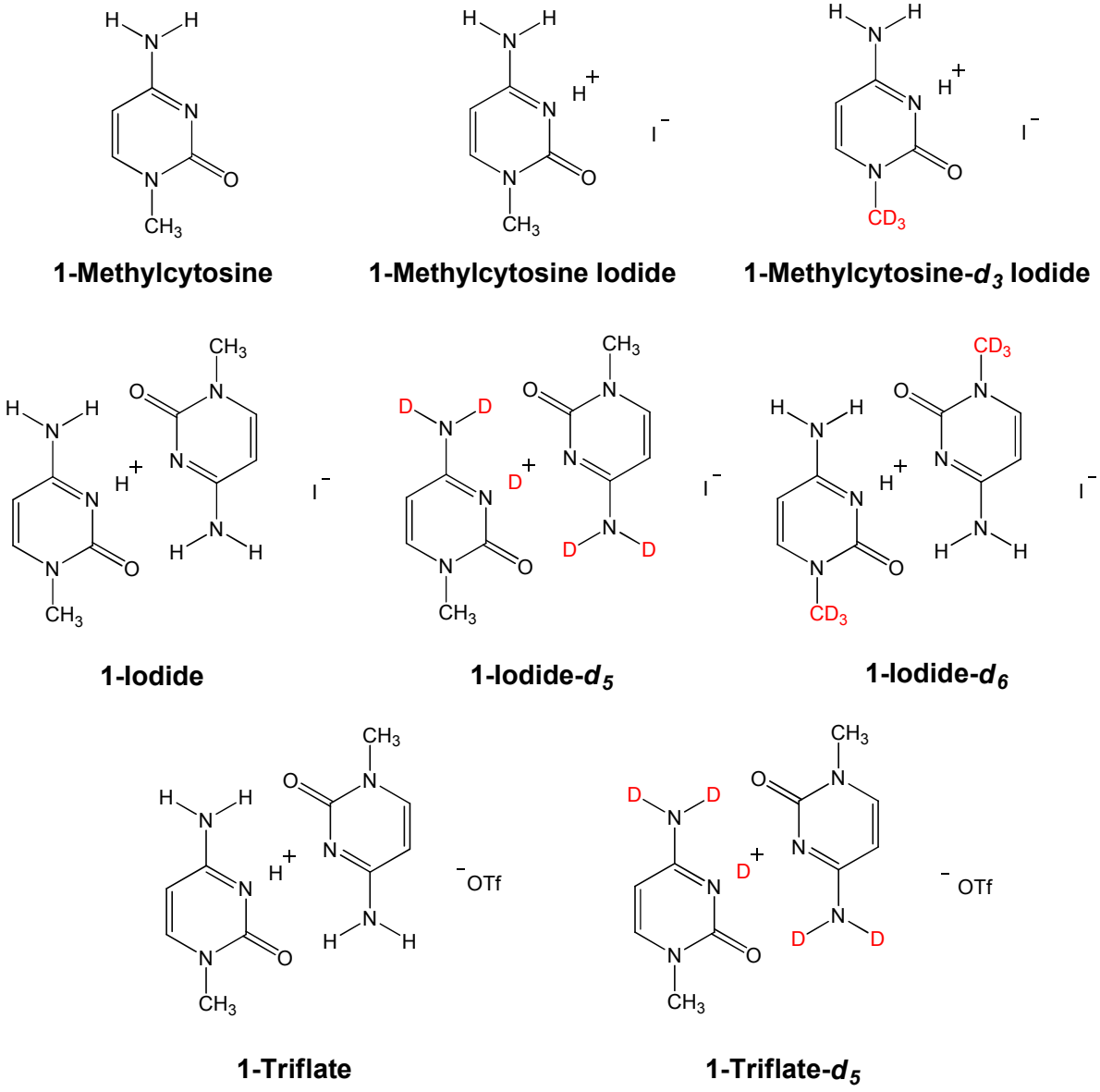


Chart 2: 1-Methylcytosine and its analogs chosen for INS experiments.

Comparison between experimental and predicted INS spectra of 1-methylcytosine is reproduced in **Figure 2.36**. The black trace represents the

predicted INS spectrum, while the red and blue traces correspond to the experimental spectrum. The blue trace utilizes a maximum entropy algorithim to help resolve and smooth out the raw spectrum (red). Comparison between the predicted and experimental spectra do not match very well. There are bands present in the experimental spectra not predicted by theory. Applying the maximum entropy algorithm smooths out the spectrum, but it also shifts and eliminates some bands. The intense band originally observed near 240 cm^{-1} (red trace) shifts to 250 cm^{-1} (blue trace). A band in the red trace at 100 cm^{-1} disappears almost completely in the blue trace. An intense band is observed in both the red and blue traces around 240 cm^{-1}. This band could be the external amino N-H bending motion since it is predicted appear around 290 cm^{-1}. The motion associated with the methyl torsion is predicted at 94 cm^{-1}. Due to the shifting of bands when applying the maximum enthropy algorithm, comparison between theory and experiment will make use of the raw INS spectrum.

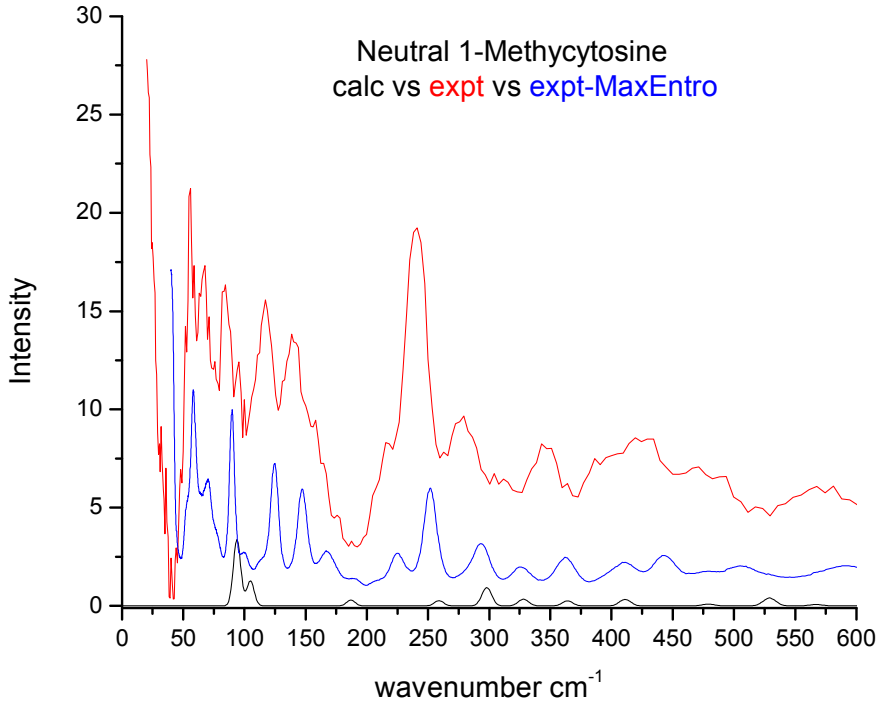


Figure 2.37: Comparison between experimental and predicted INS spectrum of neutral 1-methylcytosine in the domain between 50 – 600 cm^{-1}.

1-Methylcytosine-d_2 (the hydrogens on the amino group are exchanged for deuterium using recrystallization from *O*-deuterated ethanol) was synthesized, but due to the allotted beam time at LANSCE, the INS spectrum of this molecule was not obtained. The INS spectrum of this molecule would have provided clues as to where to expect the N-H bending motions when plotted against neutral 1-methylcytosine.

The comparison between experimental INS spectra of protonated 1-methycytosine iodide salt and protonated 1-methylcytosine-d_3 Iodide salt is

reproduced in **Figure 2.38**. The addition of the CD_3 group should suppress the methyl torsion while leaving the N-H bending motions intact. If the protonated salt of 1-methylcytosine behaves like its neutral counterpart, we should expect a band around 240 cm^{-1} to shift when compared to the CD_3 analog. A small band is observed around 240 cm^{-1} but it is not as intense as in the INS spectrum of neutral 1-methylcytosine. Because of overlap with a larger peak in **Figure 2.38**, it is not clear whether this peak disappears upon deuteration of the methyl.

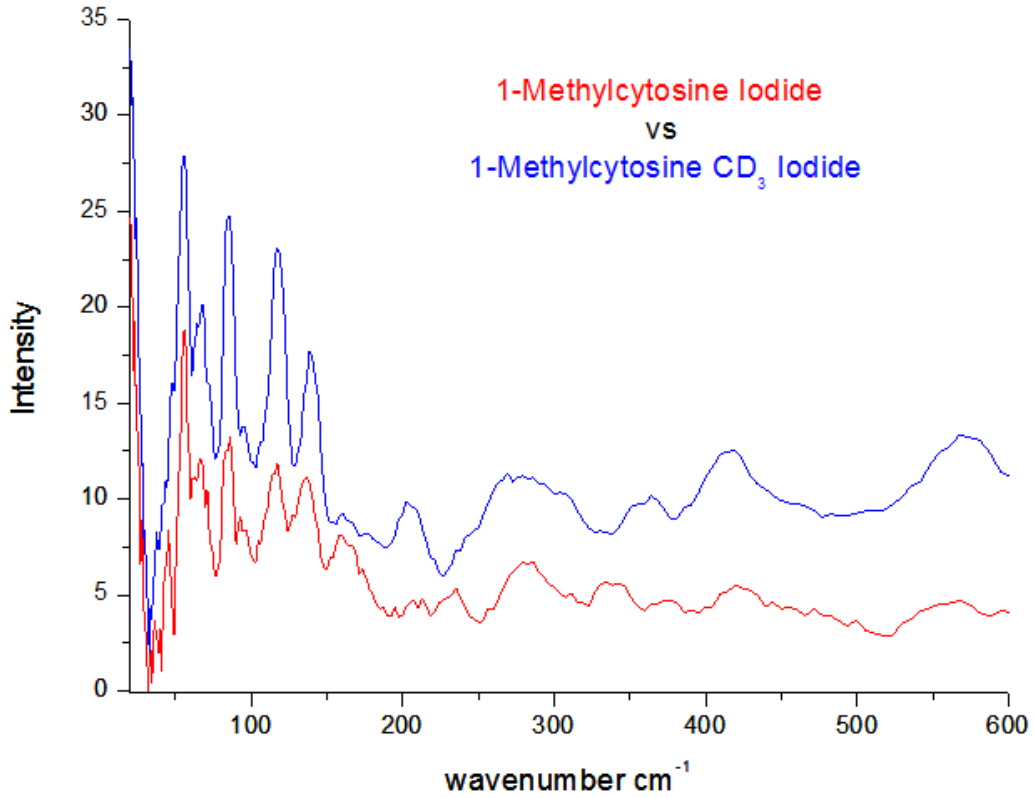


Figure 2.38: Comparison of INS spectra of protonated 1-methylcytosine iodide salt and 1-methylcytosine-d_3 iodide salt between 50 – 600 cm^{-1}.

Figure 2.39 depicts the comparison between raw experimental and predicted INS spectra of **1-Iodide**. The raw INS data is represented by the blue

trace, while the predicted INS spectrum is in red. Like neutral 1-methylcytosine, more bands are observed experimentally than predicted. The amino N-H bending motions of **1-Iodide** are predicted to be at 443 and 540 cm^{-1}. There is only noise above 240 cm^{-1} in the experimental spectrum so the domain is cut off at 240 cm^{-1}. The methyl torsions in the gas phase are predicted at 79 and 106 cm^{-1}, and the "butterfly" flapping motion of the two bases is predicted at 38 cm^{-1}.

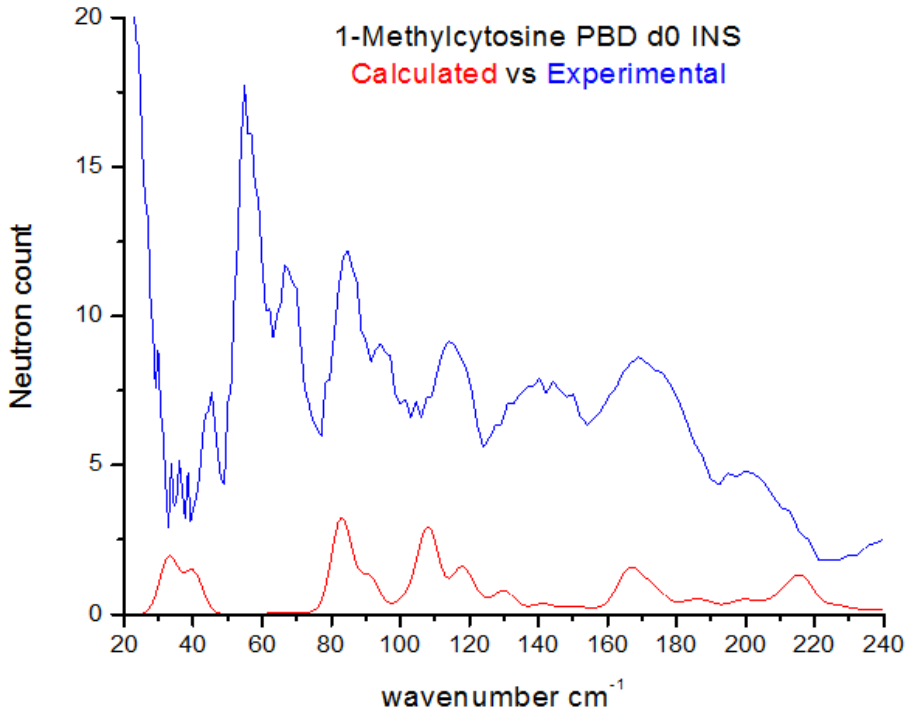


Figure 2.39: Comparison between experimental and predicted INS spectra of **1-Iodide** in the domain between 50 – 600 cm^{-1} recorded at LANSCE.

Figure 2.40 depicts the comparison between predicted and experimental INS spectra of **1-Iodide-*d*$_5$** and **1-Iodide-*d*$_6$**. Like before, more bands are observed than predicted. Since the external amino hydrogens ought to be exchanged in **1-Iodide-*d*$_5$**, there should not be a motion associated with the N-H

bending modes. The methyl groups are still intact, so there should be a motion associated with the methyl torsions. Conversely, **1-Iodide-d_6** should show the external amino N-H bending motion and not the methyl torsions.

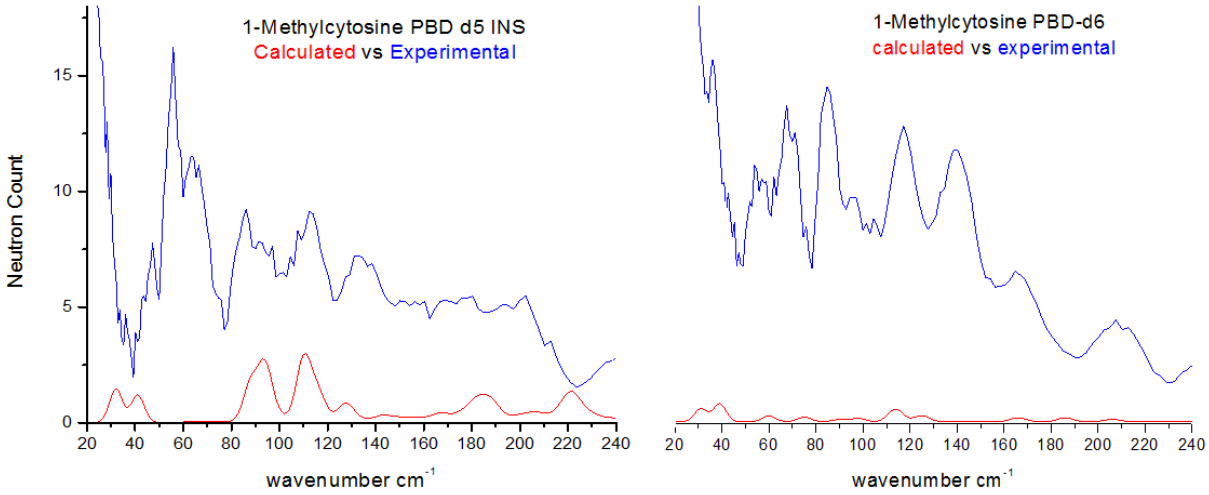


Figure 2.40: Comparison between experimental and predicted INS spectra of **1-Iodide-d_5** and **1-Iodide-d_6** in the domain between 50 – 600 cm^{-1} recorded at LANSCE.

A comparison between experimental INS spectra of **1-Iodide** and **1-Iodide-d_5** should show a band in the **1-Iodide** spectrum that is not present in the **1-Iodide-d_5** spectrum. **Figure 2.41** depicts the comparison and a band around 160 cm^{-1} seems to vanish when the exchangeable hydrogens are exchanged for deuteria. As previously stated, calculations predict two methyl torsions at 79 and 106 cm^{-1}. If the band near 160 cm^{-1} is in fact the motion of two methyl torsions, they seem to cluster together as one broad band and appear at higher frequency than predicted by calculations.

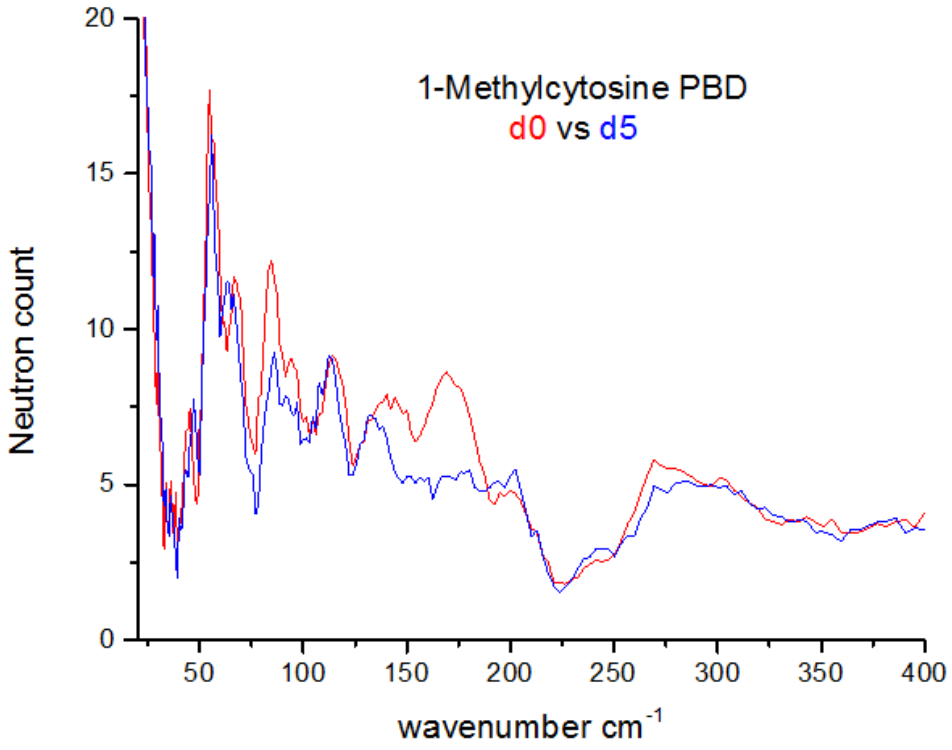


Figure 2.41: Comparison between experimental INS spectra of **1-Iodide** (red) and **1-Iodide-*d*$_5$** (blue) in the domain between 50 – 400 cm^{-1} recorded at LANSCE.

The motion of the external amino N-H bends should be detectable by comparing experimental INS spectra of **1-Iodide** and **1-Iodide-*d*$_6$**. A band around 420 cm^{-1} is much more intense in the INS spectrum of **1-Iodide** than in **1-Iodide-*d*$_6$** spectrum (see **Figure 2.42**). The N-H bends of the external amino group are predicted at 443 and 540 cm^{-1} for **1-Iodide**. Although the band at 420 cm^{-1} did not disappear completely, it did decrease in intensity when plotted against the **d$_6$** analog. It is conceivable this band could be the N-H bending modes clustered into one broad band.

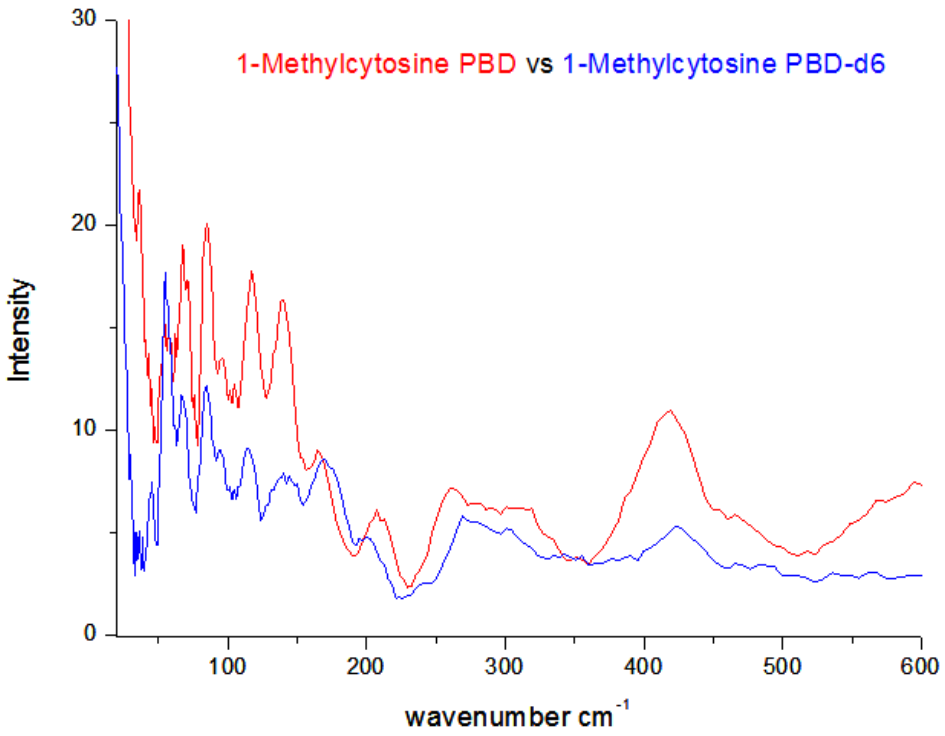


Figure 2.42: Comparison between experimental INS spectra of **1-Iodide** (red) and **1-Iodide-*d*$_6$** (blue) in the domain between 50 – 600 cm^{-1} recorded at LANSCE.

Due to the shut down of the neutron scattering facility at LANSCE, additional neutron scattering spectra of neutral **1-methylcytosine** (black trace), **1-Iodide** (red trace), **1-Iodide-*d*$_5$** (green trace), **1-Triflate** (dark blue trace), and **1-Triflate-*d*$_5$** (light blue trace), from **Figure 2.43**, were obtained at Oak Ridge National Laboratory (ORNL). A band stands out in the spectrum of **1-Iodide** near 2 meV, which translate to approximately 16 cm^{-1}. The top panel of **Figure 2.44** reproduces the neutron scattering spectrum of **1-Iodide** from -2 to 10 meV. There looks to two bands present that is not in the d_5 spectrum (**Figure 2.44** bottom panel). The tunneling of the bridging proton from one nucleobase to

another is of interest and neutron scattering allows us to observe hydrogen motions.

Unfortuately, X-ray diffraction analysis of **1-Iodide** showed the localizationof the bridging proton on one side of 1-methylcytosine. Although the proton can't be onserved by X-ray, the two N-O distances were shown to have unequal distances. The displacement of the bridging proton from one 1-methylcytosine to the other causes the two N-O distances to shorten and lengthen with respect to the other. Since the bridging proton is isolated on one side of the PBD, it is unlikely the band observed near 2meV is unlikely to be a signal of proton tunneling.

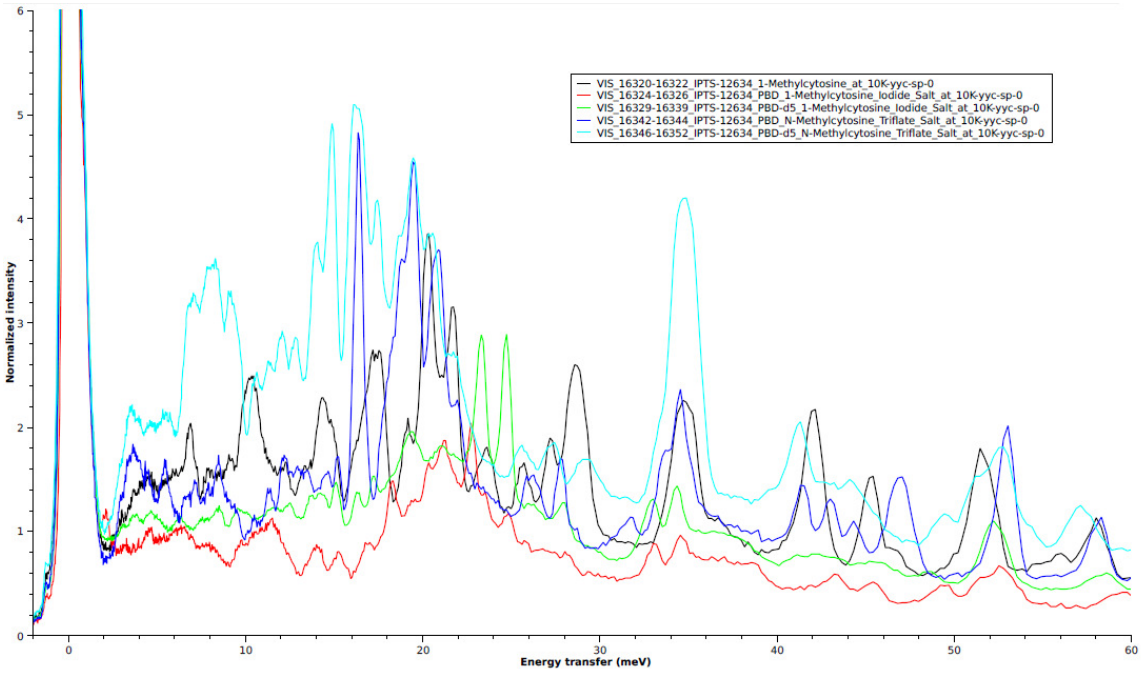


Figure 2.43: Neutron scattering spectra of **1-methylcytosine** (black trace), **1-Iodide** (red trace), **1-Iodide-*d*$_5$** (green trace), **1-Triflate** (dark blue trace), and **1-Triflate-*d*$_5$** (light blue trace) recorded at SNS by Dr. Luke Daemen.

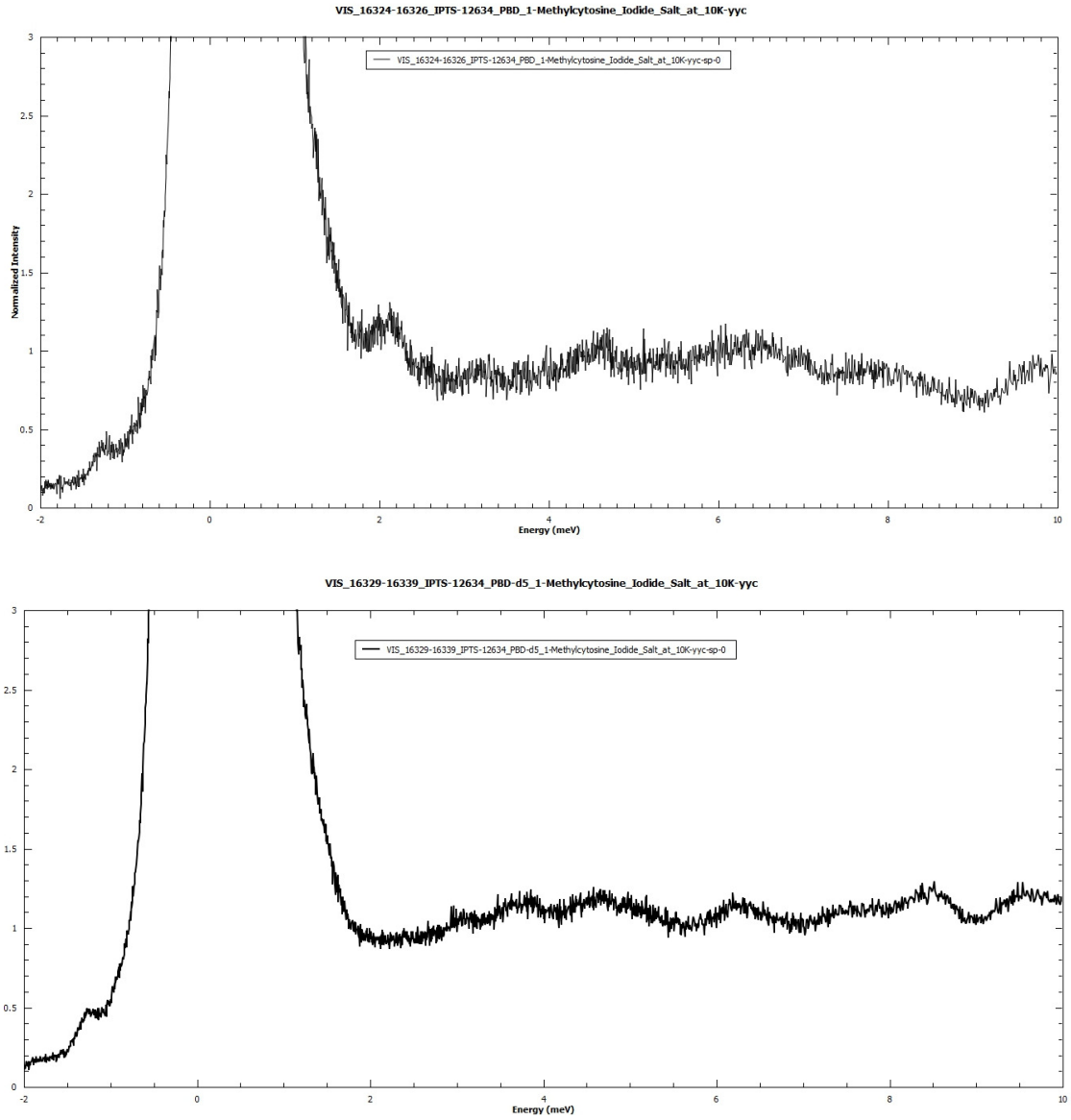


Figure 2.44: Top Panel) Zoomed in view of **1-Iodide** from -2 to 10 meV. Bottom Panel) Zoomed in view of **1-Iodide-*d*$_5$** from -2 to 10 meV recorded at SNS by Dr. Luke Daemen.

Conclusions

In summary, the IRMPD spectra of protonated 1-methylcytosine have been plotted against the predicted IR spectra calculated using anharmonic

normal modes. The comparison reaffirms the previous assignment of protonated 1-methylcytosine. The fingerprint domain of 1-methylcytosine between 600 – 1800 cm^{-1} looks to correspond to the *N3*-protonated tautomer while the CH/NH/OH stretching domain looks to correspond to the *O*-protonated tautomer. The experimental IRMPD spectrum of protonated 1-methylcytosine was generated in two ways. One way is through direct ionization of the monomer and the second method fragmented the PBD to produce protonated monomers. Both methods produced the same spectra.

The fingerprint and CH/NH/OH stretching domain of six proton-bound homo- and hetero- dimers were investigated in the gas phase. The IRMPD spectrum of 1-Iodide in the CH/NH/OH stretching domain showed fundamental along with possible overtones and combination bands. Partial deuteration experiments, where four out of the five exchangeable hydrogens are exchanged for deuterium, help deconvolute the spectrum by suppressing overtones and combination bands. All three proton-bound homodimers exhibit a single band shift upon complete deuteration of the exchangeable hydrogens around 1570 cm^{-1}. This band was previously assigned to be the $N\text{-}H^{+\cdots}N$ stretch of the bridging proton. The bases of the proton-bound heterodimers share the bridging proton unequally. This unequal sharing leads to two band shift upon complete deuteration (between 1550 – 1600 cm^{-1}) of the exchangeable hydrogens.

1-Iodide was synthesized and recrystallized from slow evaporation of absolute ethanol. Two crystal habits of 1-Iodide are observed. Crystal habit A

looks flakey-cubic, while crystal habit B looks like needles. X-ray diffraction of **1-Iodide** crystal habit A and **1-Iodide-*d*$_5$** crystal habit B shows the same crystal dimensions (two sets of PBD salt per unit cell) but different compared to **1-Iodide** crystal habit B. Results from X-ray diffraction suggest the two bases are not identical with the bridging proton residing on one side of the nucleobase or the other. ^{13}C ssNMR spectra of **1-Iodide** and **1-Triflate** both show atleast 8 resoanaces. The resonance associated with bridging proton is observed around 16 ppm. ^{13}C –^{15}N dipolar coupling experiment places the N-H distances at 1.057 and 1.784 Å, matching prediction.

Fragment-based chemical shift calculations were used to predict the ^{13}C chemical shifts of **1-Iodide** and **1-Iodide-*d*$_5$**. The 13C ssNMR spectra of **1-Iodide** habit B and **1-Iodide-*d*$_5$** are vastly different. The difference in carbon resonances from the experimental ^{13}C ssNMR arise from differences between PBDs within a unit cell rather than difference within the PBD. This conclusion was brought to light when the X-ray structure of protonated 1-methylcytosine triflate salt also shows the occupation of two sets of protonated salt per unit cell. ^{13}C ssNMR spectrum of protonated 1-methylcytosine triflate salt also shows doubling of several carbon resonances, suggesting the doubling of carbon resonance arise from the difference between molecules.

INS spectra of **1-Iodide**, **1-Triflate**, and their deuterated analogs were obtained at LANSCE and ORNL. INS spectra comparison between completely protonated and various deuterated analogs of **1-Iodide** show band shifts in the

deuterated spectrum that could correspond to methyl torsions (160 cm^{-1}) and external amino N-H bends (420 cm^{-1}). INS spectra obtained at ORNL show two bands for **1-Iodide**, around 2 meV, which is not present in the deuterated spectrum. This phenomenon will be revisited in the future by conducting more experiments at lower temperatures to see if these two bands increase as an inverse function of temperature.

References

1. J. D. Watson, F. H. Crick “Molecular Structure of Nucleic Acids; A Structure for Deoxyribose Nucleic Acid” *Nature* **1953**, 171, 737 – 738.

2. G. Biffi,D. Tannahill, J. McCafferty, S. Balasubramanian “Quantitative Visualization of DNA G-quadruplex Structures in Human Cells” *Nat Chem.* **2013**, 5, 182 – 186.

3. H. R. Nasiri, N. M. Bell, K. I. E. McLuckie, J. Husby, C. Abell, S. Neidle, S. Balasubramanian “Targeting a c-MYC G-quadruplex DNA with a Fragment Library” *Chem. Commun.* **2014**, 50, 1704 – 1707.

4. H. Fernando, S. Sewitz, J. Darot, S. Tavare, J. L. Huppert, S. Balasubramanian “Genome-wide Analysis of a G-quadruplex-Specific Single-Chain Antibody that Regulates Gene Expression” *Nucl. Acids. Res.* **2009**, 37, 6716 – 6722.

5. J. L. Huppert, S. Balasubramanian “G-quadruplexes in Promoters Throughout the Human Genome” *Nucl. Acids. Res.* **2007**, 35, 406 – 413.

6. K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan “Structure of the Human Telomere in K+ Solution: A Stable Basket-Type G-Quadruplex with Only Two G-Tetrad Layers” *J. Am. Chem. Soc.***2009**, 131, 4301 – 4309.

7. D. Hanahan, R. A. Weinberg “The Hallmarks of Cancer” *Cell* **2000**, 100, 57 – 70.

8. T. A. Brooks, S. Kendrick, L. Hurley “Making Sense of G-quadruplex and i-Motif Functions in Oncogene Promoters” *FEBS J.* **2010**, 277, 3459 – 3469.

9. Y. Qin, L. H. Hurley “Structures, Folding Patterns, and Functions of Intramolecular DNA G-quadruplexes Found in Eukaryotic Promoter Regions” *Biochimie* **2009**, 90, 1149 – 1171.

10. S. L. Palumbo, R. M. Memmott, D. J. Uribe, Y. Krotova-Khan, L. H. Hurley, S. W. Ebbinghaus “A Novel G-quadruplex-Forming GGA Repeat Region in the c-myc Promoter is a Critical Regulator of Promoter Activity. *Nucl. Acids. Res.* **2008**, 36, 1755 – 1769.

11. S. L. Palumbo, S. W. Ebbinghaus, L. H. Hurley “Formation of a Unique End-to-End Stacked Pair of G-Quadruplexes in the hTERT Core Promoter

with Implications for Inhibition of Telomerase by G-Quadruplex-Interactive Ligands" *J. Am. Chem. Soc.* **2009**, 131, 10878 – 10891.

12. T. A. Brooks, L. H. Hurley "The Role of Supercoiling in Transcriptional Control of MYC and its Importance in Molecular Therapeutics" *Nature Reviews Cancer*, **2009**, 9, 849 – 861.

13. B. Yang, R. R. Wu, G. Berden, J. Oomens, M. T. Rodgers "Infrared Multiple Photon Dissociation Action Spectroscopy of Proton-Bound Dimers of Cytosine and Modified Cytosines: Effects of Modifications on Gas-Phase Conformations" *J. Phys. Chem. B* **2013**, 117, 14191 – 14201.

14. B. Yang, M. T. Rodgers "Base-Paring Energies of Proton-Bound Heterodimers of Cytosine and Modified Cytosines: Implications for the Stability of DNA *i*-motif Conformations" *J. Am. Chem. Soc.* **2014**, 136, 282 – 290.

15. S. Kendrick, H. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht, L. H. Hurley "The Dynamic Character of the BCL2 Promoter i-Motif Provides a mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure" *J. Am. Chem. Soc.* **2014**, 136, 4161 – 4171.

16. A. G. Tsai, A. E. Engelhart, M. M. Hatmal, S. I. Houston, N. V. Hud, I. S. Haworth, M. R. Lieber "Conformational Variants of Duplex DNA Correlated with Cytosine-Rich Chromosomal Fragile Sites" *J. Biol. Chem.***2009**, 284, 7157 – 7164.

17. M. M. Mautner "Strong Ionic Hydrogen Bonds" *Chem. Rev.* **2012**, 112, PR22 – PR103.

18. W. C. Hahn, S. A. Stewart, M. W. Brooks, S. G.York, E. Eaton, A. Kurachi, R. L. Beijersbergen, J. H. M. Knoll, M. Meyerson, R. A. Weinberg "Inhibition of Telomerase Limits the Growth of Human Cancer Cells" *Nature Medicine* **1999**, 5, 1164 – 1170.

19. R. J. Haslam, H. B. Koide, B. A. Hemmings "Pleckstrin Domain Homology" *Nature* **1993**, 363, 309 – 310.

20. J-L. Leroy, M. Gueron "Solution Structures of the i-Motif Tetramers of d(TCC), d(5methylCCT) and d(d5methylCC): Novel NOE Connections Between Amino Protons and Sugar Protons" *Structure* **1995**, 3 (1), 101 – 120.

21. J. L. Mergny, L. Lacroix “Kinetics and Thermodynamics of i-DNA Formation: Phosphodiester Versus Modified Oligodeoxynucleotides” *Nucl. Acids. Res.* **1998**, 26, 4797 – 4803.

22. D. Collin, K. Gehring “Stability of Chimeric DNA/RNA Cytosine Tetrads: Implications for *i*-Motif Formation by RNA” *J. Am. Chem. Soc.* **1998**, 120, 4069 – 4072.

23. S. Robidoux, M. J. Damha “D-2-Deoxyribose and D-Arabinose, but not D-Ribose, Stabilize the Cytosine Tetrad (*i*-DNA) Structure” *Journal of Biomolecular Structure and Dynamics* **1997**, 15, 529 – 535.

24. R. E. Marsh, R. Bierstedt, E. L. Eichhorn “The Crystal Structure of Cytosine-5-Acetic Acid” *Acta Cryst.* **1962**, 15, 310 – 317.

25. T. J. Kistenmacher, M. Rossi, L. G. Marzilli “A Model for the Interrelationship Between Asymmetric Interbase Hydrogen Bonding and Base-Base Stacking in Hemiprotonated Polyribocytidylic Acid: Crystal Structure of 1-methylcytosine Hemihydroiodide Hemihydrate” *Biopolymers* **1978**, 17, 2581 – 2585.

26. D. Armentano, G. De Munno, L. D. Donna, G. Sindona, G. Giorgi, L. Salvini “ Self-Assembling of Cytosine Nucleoside into Triply-Bound Dimers in Acid Media. A Comprehensive Evaluation of Proton-Bound Pyrimidine Nucleosides by Electrospray Tandem Mass Spectrometry, X-rays Diffractometry, and Theoretical Calculations” *Journal of the American Society for Mass Spectrometry* **2004**, 15, 268 – 279.

27. T. Kruger, C. Bruhn, D. Steinborn “Synthesis of $[(MeCyt)_2H]I$-Structure and Stability of a Dimeric Threefold Hydrogen-Bonded 1-Methylcytosinium 1-Methylcytosine Cation” *Org. Biomol. Chem.* **2004**, 2, 2513 – 2516.

28. J. Muller, E. Freisinger “$[(1\text{-Methylcytosine})_2H]I$, An Asymmetric Base Pair“ *Acta Crystallographica Section E* **2005**, E61, o320 – o322.

29. N. Bosnjakovic-Pavlovic, A. S-de. Bire “Cytosine-Cytosinium Dimer Behavior in a Cocrystal with a Decavanadate Anion as a Function of the Temperature” *J. Phys. Chem. A.* **2010**, 114, 10664 – 10675.

30. B. Borah, J. L. Wood “The Cytidinium – Cytidine Complex: Infrared and Raman Spectroscopic Studies” *Journal of Molecular Structure* **1976**, 30, 13 – 30.

31. J. Oomens, A. R. Moehlig, T. H. Morton “Infrared Multiple Photon Dissociation (IRMPD) Spectroscopy of the Proton-Bound Dimer of 1-Methylcytosine in the Gas Phase” *J. Phys. Chem. Lett.* **2010**, 1, 2891 – 2897.

32. D. Oepts, A. F. G. van der Meer, P. W. van Amersfoort “The Free-Electron-Laser user facility FELIX” *Infrared Physics & Technology* **1995**, 36, 297 – 308.

33. N. C. Polfer, J. Oomens “Reaction Products in Mass Spectrometry Elucidated with Infrared Spectroscopy” *Phys. Chem. Chem. Phys.* **2007**, 9, 3804 – 3817.

34. R. C. Polfer “Infrared Multiple Photon Dissociation Spectroscopy of Trapped Ions” *Chem. Soc. Rev.* **2011**, 40, 211 – 2221.

35. S. Guan, A. G. Marshall “Stored Waveform Inverse Fourier Transform (SWIFT) Ion Excitation in Trapped-Ion Mass Spectrometry: Theory and Application” *International Journal of Mass Spectrometry and Ion Process*, **1996**, 157 – 158, 5 – 37.

36. P. J. Tarsio, L. Nicholl “Preparation of Cytosine” *J. Org. Chem.* **1957**, 22, 192 – 193.

37. D. Davidson, O. Baudisch “The Preparation of Uracil From Urea” *J. Am. Chem. Soc.* **1926**, 48, 2379 – 2383.

38. D. L. Helfer II, R. S. Hosmane, N. J. Leonard “Selective Alkylation and Aralkylation of Cytosine at the 1-Position” *J. Org. Chem. I* **1981**, 46, 4803 – 4804.

39. M. Rossi, T. J. Kistenmacher “1-Methylcytosine: A Refinement” *Acta Cryst.* **1977**, B33, 3962 – 3695.

40. J. M. Bakker, J-Y. Salpin, P. Maitre “Tautomerism of Cytosine Probed by Gas Phase IR Spectroscopy” *International Journal of Mass Spectrometry* **2009**, 283, 214 – 221.

Chapter III:

Nucleosides

Introduction

DNA and RNA nucleobases utilize hydrogen bonding, regardless of whether neutral nucleobases associate to form traditional Watson-Crick double-helix or alternative secondary structures are created. As discussed previously, alternative secondary structures can arise due to supercoiling of single-stranded nucleic acids. DNA, for instance, can form non-traditional Watson-Crick bonds when guanine-rich and the complementary cytosine-rich strands separate from one another and self-associate. Strands that are primarily guanine-rich may self-associate to form the G-quadruplex, which utilizes hydrogen bonding at both the Watson-Crick and Hoogsteen faces, while strands that are cytosine-rich can form the *i*-motif. One G-quadruplex unit consists of four guanine residues bound via hydrogen bonding at the Watson-Crick and the Hoogsteen face surrounding a potassium ion. The *i*-motif consists of six (or more) proton-bound dimers (PBDs) of cytosine, which are also called hemiprotonated cytosine dimers, intercalated with one another. The PBDs are oriented 90^{o} from one another with respect to the PBDs above and below. **Figure 3.1** breaks down the *i*-motif into its individual components.

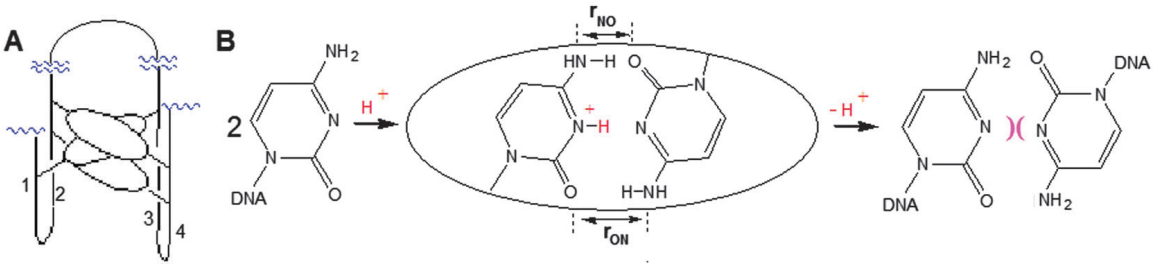


Figure 3.1[4]: **A**) Schematic representation of proton-bound dimers within the *i*-motif. **B**) Each oval from A is a representation of 1-methycytosine proton-bound dimer. The loss of the bridging proton causes the lone pairs on nitrogen to repel one another.

Double-stranded association of cytosine PBDs require a parallel orientation, but if a single strand bends three times into a hairpin structure, the intercalation within the strand can be incorporated into an anti-parallel orientation, as depicted in **Figure 3.2**. The driving force holding these PBDs together is hydrogen bonding. The *i*-motif utilizes two ordinary hydrogen bonds (at the top and bottom), and one ionic hydrogen bond (middle) to hold this secondary structure together. The dissociation enthalpy between cytosine and its conjugate acid has been determined experimentally to be from 160 – 173 kJ/mol, which agrees with DFT calculations[1].

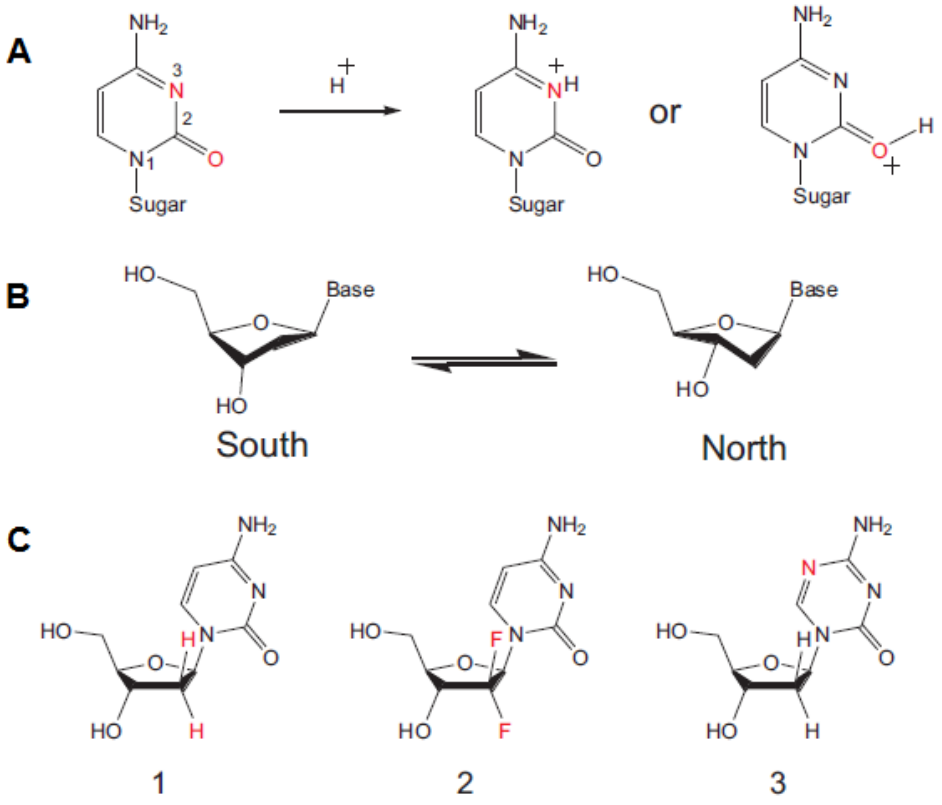


Figure 3.2: **A**) Depiction of the two most basic sites on the cytosine base (shown in red). **B**) Deoxyribose sugar ring puckering from south to north orientation. The designated South position has the carbon at the 3'-position downward and the North orientation has the carbon at the 3'-position upward. **C**) The three nucleosides under investigation: 1) 2'-deoxycytidine, 2) gemcitabine, and 3) decitabine. Gemcitabine and decitabine resemble 2'-deoxycytidine, but with two fewer hydrogens on the furanose sugar (gemcitabine) and one fewer on the base (decitabine).

More attention has been devoted to the stability of the G-quadruplex than to the stability of the *i*-motif, even though the two secondary structures should exist complementarily to one another. Promoter regions of several oncogenes are believed to adopt *i*-motif conformations when underwound. If this hypothesis is correct, stabilizing the *i*-motif with small organic molecules via hydrogen bond

formation may suppress the transcription of harmful mRNA, thus providing another avenue to combat cancer.

Background

NMR and gas phase vibrational studies of cytosine PBDs and their derivatives have previously been looked at, but nucleosides containing the cytosine base have not received as much attention. Rodgers *et al.* examined the IRMPD spectra of various protonated cytosine monomers and PBDs that are modified at the 5-position (with CH_3, F, Br, and I) as well as the base-paring energies of proton-bound heterodimers in the gas phase using threshold collision-induced dissociation[1]. Results from their findings determined that modification at the 5-position of cytosine affected the base-pairing energy of the proton-bound heterodimers.

Solution phase NMR studies of cytosine PBDs, modified at the 1-postion with either a methyl or an octyl group and at the 5-position with fluorine, have been investigated by Hooley *et al.*[2] Variable temperature NMR experiments were performed with acid titration experiments on 1-octylcytosine, 5-fluoro-1-methylcytosine, and 5-fluoro-1-octylcytosine homodimers. The slow addition of acid favors PBD formation, but an excess of acid causes the PBDs to revert to protonated monomers. This is observed by following proton resonances before and after addition of acid. The proton resonance would shift downfield upon addition of acid and would stop shifting downfield when 1 equivalent of acid is

added. The addition of more acid after 1 equivalent would shift the proton resonances back upfield.

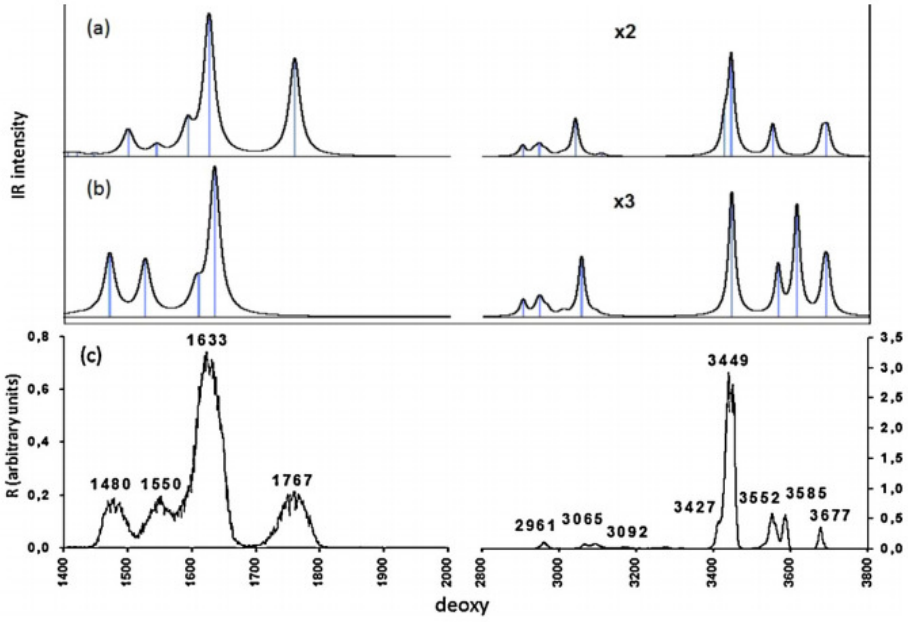


Figure 3.3[3]: Experimental and predicted (B3LYP/6-311++G**) IRMPD spectra of protonated 2'-deoxycytidine. A) Predicted scaled normal modes of *N3*-protonated 2'-deoxycytidine. B) Predicted scaled normal modes of *O*-protonated 2'-deoxycytidine. C) Experimental IRMPD spectrum of 2'-deoxycytidine in the fingerprint and CH/NH/OH stretching domain.

The IRMPD spectra of several protonated nucleosides have previously been investigated in the gas phase by Speranza *et al.*[3]. They reported the IRMPD spectra of 2'-deoxycytidine (**Figure 3.3**), cytidine, cytarabine, and gemcitabine (**Figure 3.4**) in the fingerprint domain (900 – 2000 cm^{-1}), and in the CH/NH/OH domain (2800 - 3800 cm^{-1}). It was concluded from the comparison between experimental and calculated IRMPD spectra that all of the protonated monomers exist in two tautomers, the *N3*-protonated and the *O*-protonated tautomer. The IRMPD spectra in the fingerprint domain from 900 – 2000 cm^{-1}

show good fit of peaks to match the *N3*-protonated tautomer, while the CH/NH/OH domain looks to be matching the *O*-protonated tautomer for both 2'-deoxycytidine and gemcitabine. Speranza *et al.* did not account for any overtone or combination bands that might be present in their spectra.

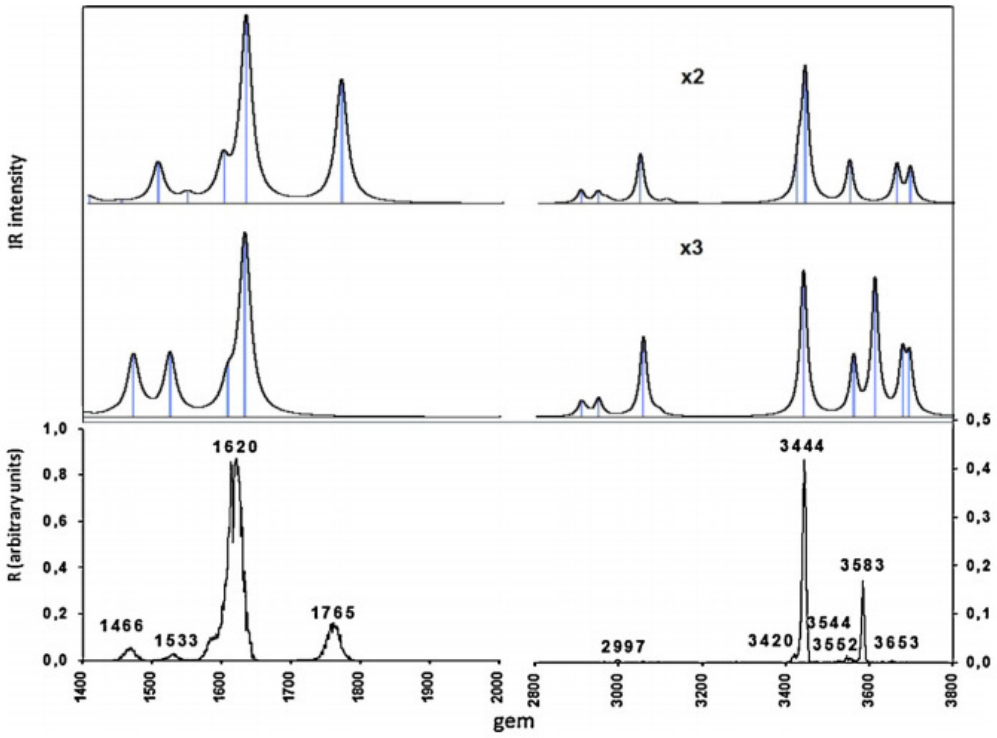


Figure 3.4[3]: Experimental and predicted (B3LYP/6-311++G**) IRMPD spectra of protonated gemcitabine. **A**) Predicted IR spectrum of *N3*-protonated gemcitabine. **B**) Predicted IR spectrum of *O*-protonated gemcitabine. **C**) Experimental IRMPD spectrum of gemcitabine in the fingerprint and CH/NH/OH stretching domain.

Partial deuteration experiments will be described in more detail below, but suggest that extra peaks reported by Speranza *et al.* in the CH/NH/OH stretching domain may arise due to overtones and combination bands. Isotopic substitution can give rise to two consequences: one being the shifting of band position of

fundamental bands and the other being disruption of anharmonic interactions. Partial deuteration has previously been employed to suppress overtones and combination bands of the PBD of 1-methylcytosine (**Figure 2.15**)[4] and have previously been shown in acetone/acetone-d_3[5].

Theory/calculations

The three nucleosides under investigation are 2'-deoxycytidine, gemcitabine, and decitabine. Gemcitabine and decitabine were designated for study because of their resemblance to 2'-deoxycytidine, minus a few hydrogens. Since the CH/NH/OH stretching domain is the primary focus of this study, using compounds with fewer hydrogens than 2'-deoxycytdine should have fewer bands in that domain. All three nucleosides were purchased from commercial vendors and were used without any further purification.

Since our results will be compared to that previously obtained by Speranza *et al.,* the same basis set used for DFT calculations is maintained along with the same scaling factor. All geometry optimization and harmonic vibrational frequencies were performed using the Gaussian 09 suite at B3LYP/6-311++G** level for both protonated monomers (*N3*- and *O*-protonated analogs) and nucleoside PBDs. Normal mode frequencies above 800 cm^{-1} were scaled by 0.961. A Gaussian line broadening of 10 cm^{-1} is applied to all theoretical spectra. This line broadening was chosen because a larger line broadening did not give a

good resolution of peaks, and a line broadening less than 10 cm^{-1} did not provide smooth spectra. All spectra were plotted using OriginPro 10.1 Student Edition.

It is possible for gemcitabine to be protonated at several different sites, but only the two most stable (*N3*- and *O*-protonation) protonation sites will be discussed. **Figure 3.5** outlines the predicted stabilities of the two aforementioned protonation sites while taking into account North and South ring puckering (**Figure 3.2B**). The most favorable conformation for gemcitabine is *N3*-protonated South followed by *N3*-protonated North. DFT predicts a 0.4 kcal/mol energy difference between from *N3*-protonated South to *N3*-protonated North. That energy gap is smaller when the less stable *O*-protonated North ring flips to *O*-protonated South. DFT predicts the least favorable orientation is *O*-protonated forming an intramolecular hydrogen bond with the hydroxyl group at the 5'-position.

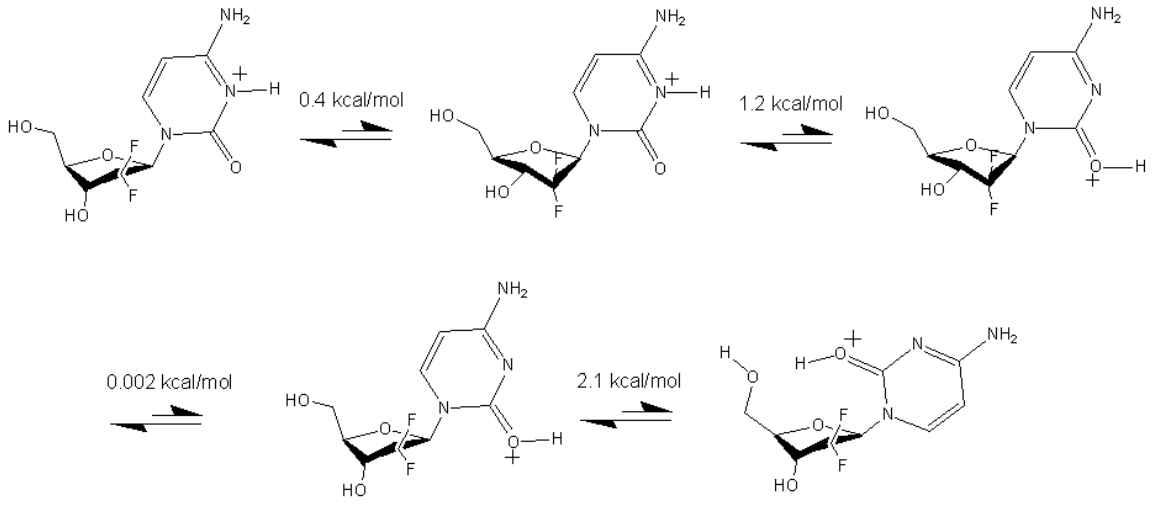


Figure 3.5: Comparison of predicted stabilities of *N3*- and *O*-protonated gemcitabine with varying ring puckering orientation, calculated at the B3LYP/6-311++G** level.

Similar to gemcitabine, decitabine has many possible protonation sites. Like gemcitabine, the two protonation sites that are focused on are the *N3*- and the *O*-protonated sites. **Figure 3.6** outlines the DFT calculated stabilities of the two protonation sites taking into account ring puckering. Unlike that of gemcitabine, having the furanose sugar ring in the South orientation was calculated to be slightly less stable when compared to North with a calculated barrier of transition from North to South sugar ring puckering requiring <0.1 kJ/mol. Comparison between experimental IRMPD spectra with DFT calculated IR spectra show a preference for the South orientation because having the furanose sugar ring in that orientation produced a better fit of peaks.

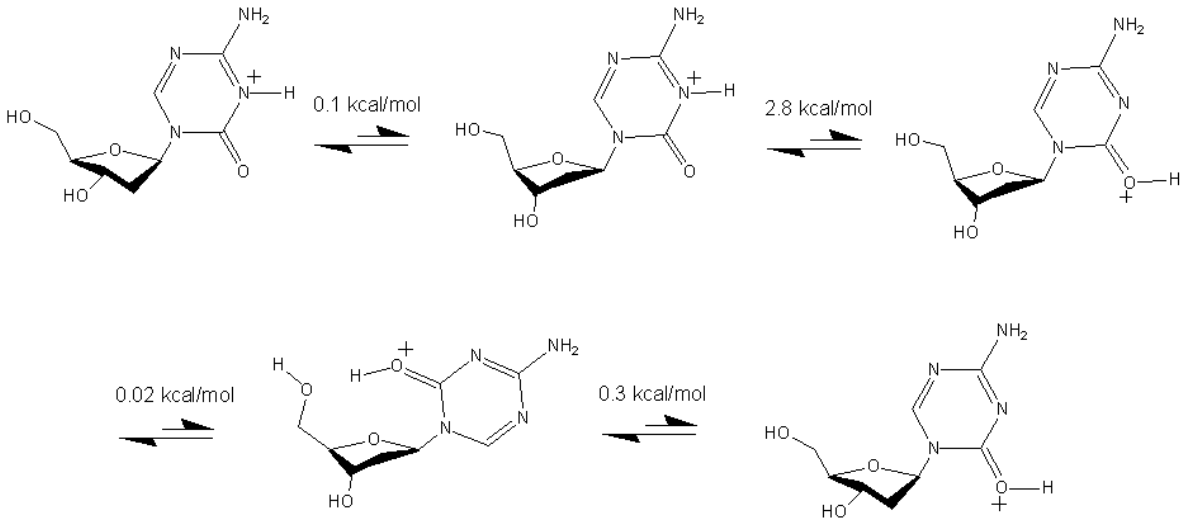


Figure 3.6: Comparison of predicted stabilities of *N3*- and *O*- protonated decitabine with different ring puckering orientation calculated at the B3LYP/6-311++G** level.

Results

Dimer equilibria of 1-methylcytosine with various nucleobases and nucleosides in the gas phase have been determined by Aaron Moehlig and Everado Arias. **Table 3.1** summarizes the proton-bound homo- and heterodimer equilibrium constants of selected nucleobases and nucleosides. PBD equilibria were determined by looking at mixtures of nucleobases and nucleosides at neutral pH with monomer concentration ratios ranging from 70/30 – 30/70. The peak intensity ratios of the observed dimers can be used to determine the equilibrium constants[6].

A	B	$K_{eq}(A_2H^+ \rightleftharpoons AHB^+)$	$K_{eq}(AHB^+ \rightleftharpoons B_2H^+)$
1-Methylcytosine	Cytidine	0.5	1.2
1-Methylcytosine	2'-deoxycytidine	3.2	1.14
1-Methylcytosine	Gemcitabine	1.0	0.07
1-Methylcytosine	1-Methylthymine	0.018	0.45
1-Methylcytosine	Thymidine	0.003	0.18
1-Methyluracil	1-Methylthymine	1.7	1.1
Thymidine	Cytidine	2.8	68
Decitabine	Gemcitabine	11.1	1.9

Table 3.1[7]: Experimental equilibrium constants for PBDs of various bases and nucleosides at 300K.

The statistical equilibrium constants should be K_{eq} ($A_2H^+ \rightleftharpoons AHB^+$) = 2 and K_{eq} ($AHB^+ \rightleftharpoons B_2H^+$) = 0.5 when ΔG values are zero. Heterodimer formation between 1-methylcytosine and cytidine are, for example, determined to be unfavorable: *i.e.* the formation of $(1\text{-methylcytosine})_2H^+$ and $(\text{cytidine})_2H^+$ homodimers are favored over the $(1\text{-methylcytosine/cytidine})H^+$ heterodimers. By

contrast, the (1-methylcytosine/2'-deoxycytidine)H^+ heterodimer is favored over the (1-methylcytosine)$_2H^+$ homodimer[7].

The proton affinities of 1-methylcytosine, 5-fluoro-1-methylcytosine, 2'-deoxycytidine, gemcitabine, decitabine, and cytidine were determined using the kinetic method in the laboratory of Prof. J.C. Poutsma at The College of William and Mary[8-12]. The results are reproduced in **Table 3.2**. 2'-deoxycytidne had the highest proton affinity and 5-fluoro-1-methylcytosine had the lowest proton affinity of the nucleobases and nucleosides selected for study.

Nucleobase	**Proton Affinity (kJ/mol)**	+/-
1-Methylcytosine	971	8
5-Fluoro-1-Methylcytosine	961	8

Nucleoside	**Proton Affinity (kJ/mol)**	+/-
2'-deoxycytidine	1011	9
Cytidine	1001	16
Gemcitabine	989	21
Decitabine	969	8

Table 3.2: Proton affinity values of modifed nucleobases and nucleosides from kinetic method expeiments.

Predicted IR spectra are obtained from calculations using the Gaussian 09 suite. The predicted IR spectra coupled with experimental IRMPD spectra provide a good representation of the ions in the gas phase. There is a variety of protonation sites available on the nucleosides of interest, either on nitrogen or oxygen. Our comparison will primarily focus on protonation at the *N3*-position

and the carbonyl oxygen of the base. Furanose ring puckering North and South are also explored. In the case of the nucleosides chosen for study, having the ring pucker in the South orientation was calculated to be the more favorable conformation. Therefore, comparison between *N3*- and *O*-protonated nucleosides will be with respect to the South ring puckering conformation.

Before evaluating the results from IRMPD experiments, overtones and combinations must first be discussed. Overtones represent multiple quantum transitions that are close to an integer multiple of a fundamental band. Second and third overtones of a fundamental band are possible, but with diminished intensity. Combination bands result from simultaneous excitation of two fundamental bands. Overtones and combination bands represent a problem when investigating the CH/NH/OH domain. To combat this problem, we employ the same partial deuteration experiment where four of the five exchangeable hydrogens are exchanged for deuterium. If all of the exchangeable hydrogens were to be replaced with deuterium, it would completely eliminate the NH/OH stretches. Using an FT-ICR makes this method possible. FT-ICR can isolate an ion of a specific *m/z*, which then can be isolated and irradiated to obtain its IRMPD spectra with suppressed overtones and combination bands.

Comparison of our experimental IRMPD spectrum of 2'-deoxycytidine in the CH/NH/OH domain from 2800 – 3800 cm^{-1} with the spectrum previously published by Speranza *et al.* looks to be very similar. As previously stated, partial deuteration will help resolve the spectra by suppressing overtones and

combination bands. Partial deuteration of 2'-deoxycytidine produces five isotopomers, as shown in **Figure 3.7A**.

Upon employing partial deuteration experiments, two of the five bands disappear leaving what seem to look like three fundamental bands, as Figure HB depicts. The IRMPD spectrum of the fingerprint domain of 2'-deoxycytidine was previously published by Speranza et al. An overtone of the fundamental band at 1767 cm^{-1} should give rise to a band at 3534 cm^{-1}. Our experimental IRMPD spectrum shows a band that is suppressed due to partial deuteration near 3540 cm^{-1}. Another band that is suppressed from partial deuteration looks to be the shoulder at 3400 cm-1. That shoulder could potentially be due to a combination band of the two fundamental bands at 1633 and 1767 cm^{-1}.

An initial IRMPD comparison of **Figure 3.7C** may look as though the experimental spectrum best fits that of the *O*-protonated tautomer, as previously reported by Speranza *et al.* However, the d_4-analog of 2'-deoxycytidine shows only three fundamental bands. This result best explains the predicted spectrum predicted by DFT, since only three bands are present. **Figure 3.7D** depicts the experimental calculated IR spectra of *N3*-protonated and *O*-protonated 2'-deoxycytidine-d_4. As previously discussed, partial deuteration suppresses an overtone at 3540 cm^{-1} and a combination band at 3400 cm^{-1} leaving behind three fundamental bands. The calculated *N3*-protonated tautomer predicts three bands, while the calculated *O*-protonated tautomer predicts four bands. The calculated N3-protonated tautomer best fits the experimental results, suggesting

protonated 2'-deoxycytidine predominately exist as the *N3*-protonated tautomer, which contradicts previously published results.

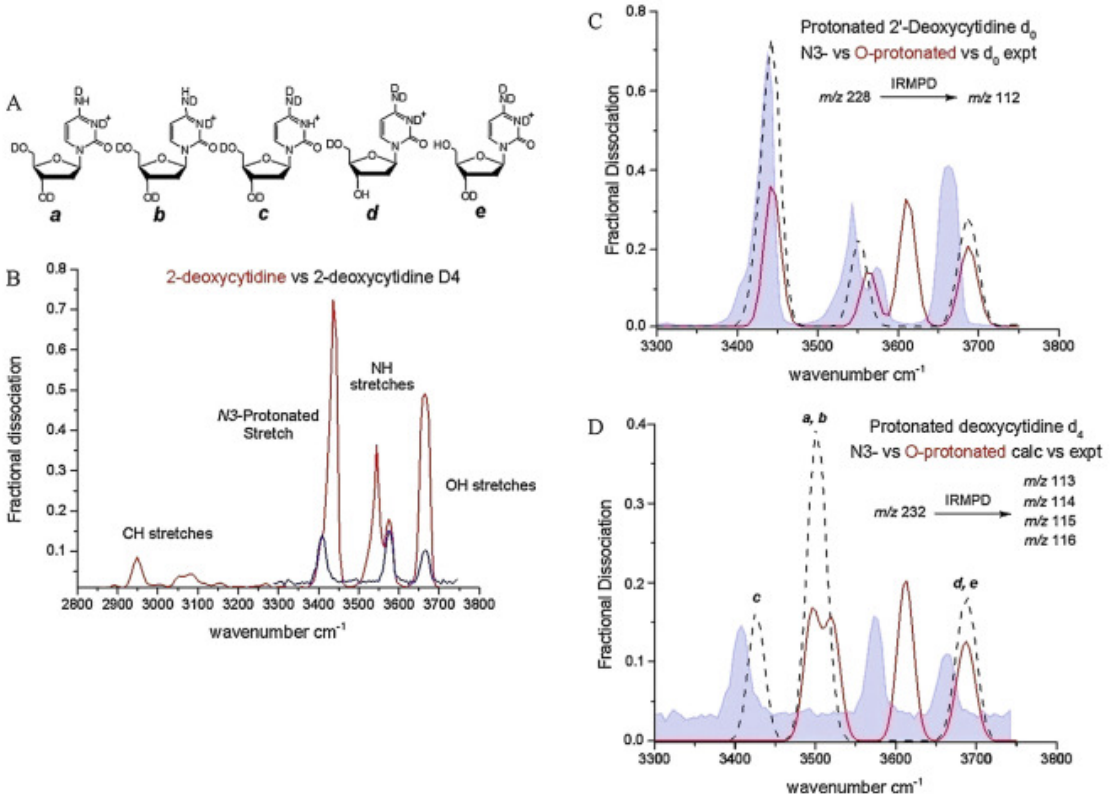


Figure 3.7: **A**) Five isotopomers of 2'-deoxycytidine are shown with the hydrogen residing at five different positions. **B**) Experimental IRMPD spectrum of protonated 2'-deoxycytidine in the CH/NH/OH domain from 2900 – 3700 cm^{-1}. The red trace depicts the protonated spectrum while the blue trace represents the protonated d_4 spectrum. **C**) Comparison of scaled normal modes obtained from DFT calculations plotted against experimental spectra of protonated 2'-deoxycytidine. The black dashed lines correspond to the calculated N3-protonated tautomer, the red trace correlates to the calculated *O*-protonated tautomer, and the blue silhouette depicts the experimental spectrum. **D**) Comparison of scaled normal modes obtained from DFT calculations plotted against experimental spectra of protonated 2'-deoxycytidine-d_4. The black dashed lines correspond to the various calculated *N3*-protonated d_4 tautomer, the red trace correlates to the calculated *O*-protonated d_4 tautomer, and the blue silhouette depicts the experimental protonated d_4 spectrum.

After evaluating the results obtained from partial deuteration of 2'-deoxycytidine, it is possible for the other nucleosides, gemcitabine and decitabine, to exhibit the same behavior. Like 2'-deoxycytidine, gemcitabine could potentially be protonated at various basic sites. The two sites that will primarily be focused on are the *N3*-protonated and *O*-protonated sites. Protonation at the *N3*-position of the base in the South orientation was calculated to be the most stable conformation. The next most stable conformation is also the N3-protonated tautomer, but in the North orientation. Protonation at the O-position was calculated to be the least stable of the three by approximately 7 kJ/mol.

Ring flipping from South to North was calculated to have a barrier of 0.4 kcal/mol, but produced significantly different IR spectra. **Figure 3.8** shows the comparison between two IRMPD spectra of *N3*-protonated gemcitabine with different ring puckering orientations. The predicted O-H stretches are predicted to be closer together so they are clustered in the North spectrum. DFT calculation predicts the O-H stretches to be farther apart and is split in the South spectrum. All predicted IRMPD spectra comparison from here on out will be in the South orientation.

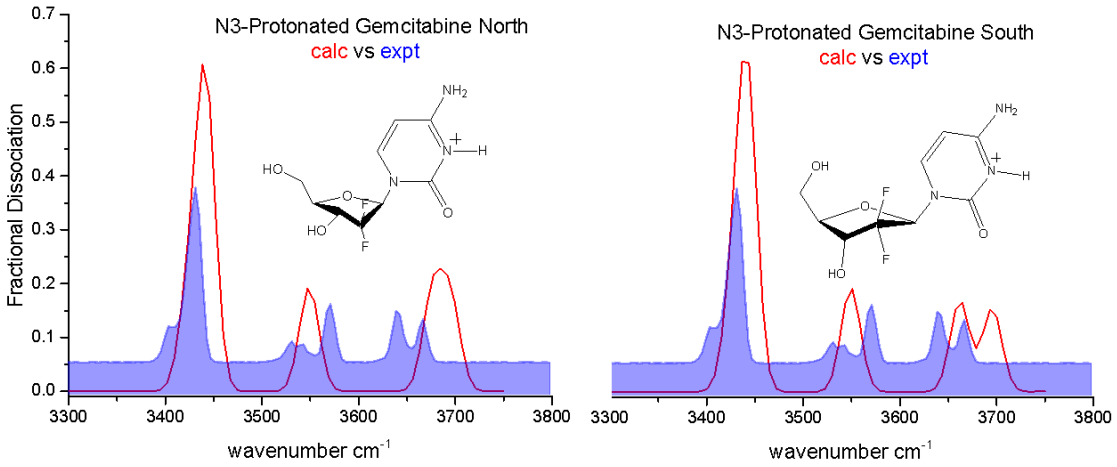


Figure 3.8: Comparison of the experimental IRMPD spectrum (blue silhouette) with predicted *N3*-protonated gemcitabine North and *N3*-protonated gemcitabine South. Scaled normal modes were calculated at the B3LYP/6-311++G** level (red traces).

Partial deuteration experiments were not performed on either gemcitabine or decitabine, but one can anticipate seeing an overtone and combination band in the CH/NH/OH stretching domain. The IRMPD spectrum in the fingerprint domain, published by Speranza *et al.*, shows the same two fundamental bands at 1620 and 1765 cm^{-1}. Like that of 2'-deoxycytidine, the fingerprint domain of gemicitabine can be used to determine any overtones or combination bands that might be present in the CH/NH/OH stretching domain.

Figure 3.9 depicts the experimental IRMPD spectra of protonated gemcitabine plotted against *N3*-protonated and *O*-protonated gemcitabine. A combination of the two fundamental bands at 1620 and 1765 cm^{-1} could give rise to the small shoulder around 3400 cm^{-1}. There are two smaller additional bands observed near 3540 and 3560 cm^{-1}. If our assignment using partial deuteration experiments is correct, then one of these bands would correspond to an overtone

of the fundamental band at 1765 cm^{-1}. An overtone of the fundamental band at 1765 cm^{-1} would give rise to a band at 3530 cm^{-1}. The band at 3540 cm^{-1} could be an overtone, but the assignment cannot be said with 100% certainty without using partial deuteration experiments to suppress any overtones or combination bands.

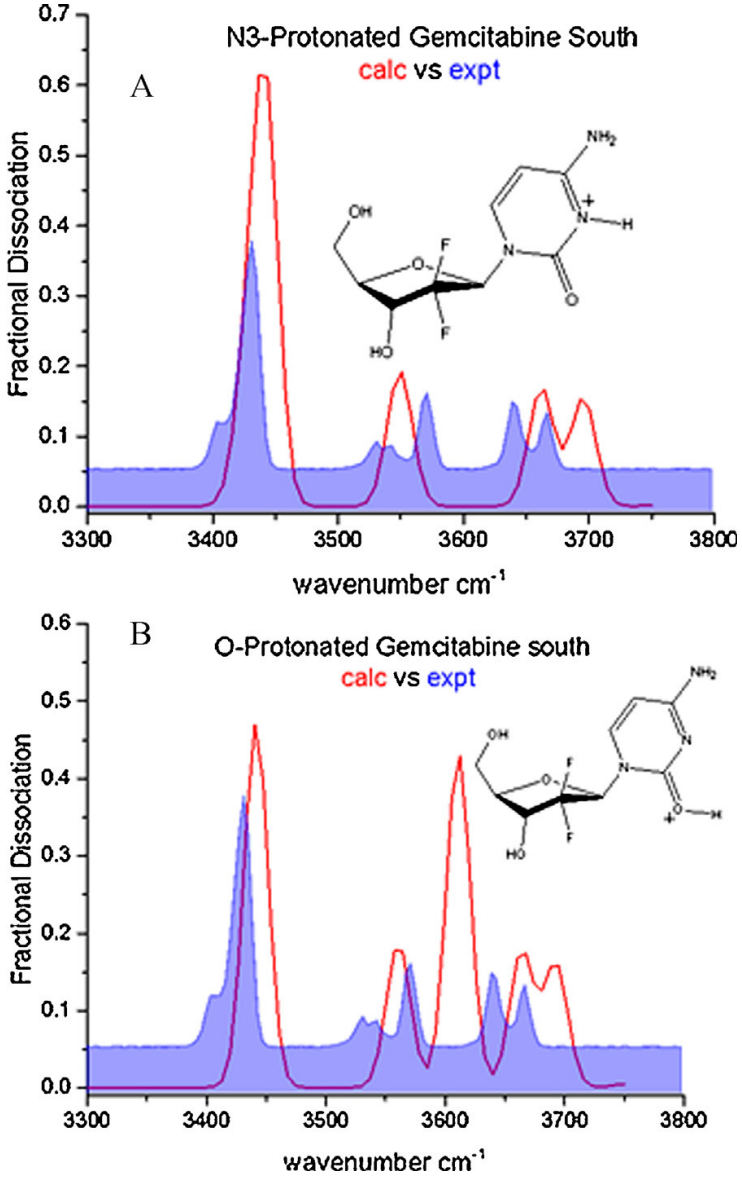


Figure 3.9: Scaled normal modes of **A**) *N3*-protonated gemcitabine and **B**) *O*-protonated gemcitabine calculated at the B3LYP/6-311++G** level (red) plotted against the experimental IRMPD spectrum (blue). The sugar conformation was calculated in the South orientation because it is more stable than the North orientation.

Figure 3.9 shows too many bands (even if we assume the bands at 3400 cm^{-1} and 3540 cm^{-1} are an overtone and combination band) for this experimental

spectrum to represent the *N3*-protonated tautomer. The N-H symmetric, asymmetric stretch, and OH stretches were predicted using DFT calculations at the B3LYP/6-311++G** level. Without the use of diffuse functions, the OH stretches were broadened because the bands were predicted to appear too close to one another to be observed as separate bands. The N-H symmetric and asymmetric stretch are calculated to be at 3440 and 3550 cm^{-1} and are observed at 3430 and 3570 cm^{-1}, respectively. The OH, from the furanose sugar ring and protonated carbonyl stretch were predicted to appear at 3655 and 3690 cm^{-1}, respectively, and they are observed experimentally at 3640 and 3670 cm^{-1}, respectively.

Comparison of the calculated IRMPD spectra between North and South ring puckering doesn't show much difference. **Figure 3.10** shows the comparison between calculated *N3*-protonated decitabine (South) and *N3*-protonated decitabine (North). Since the predicted IR of these two ions don't show any noticeable difference, the South orientation will be used to compare against the experimental spectrum.

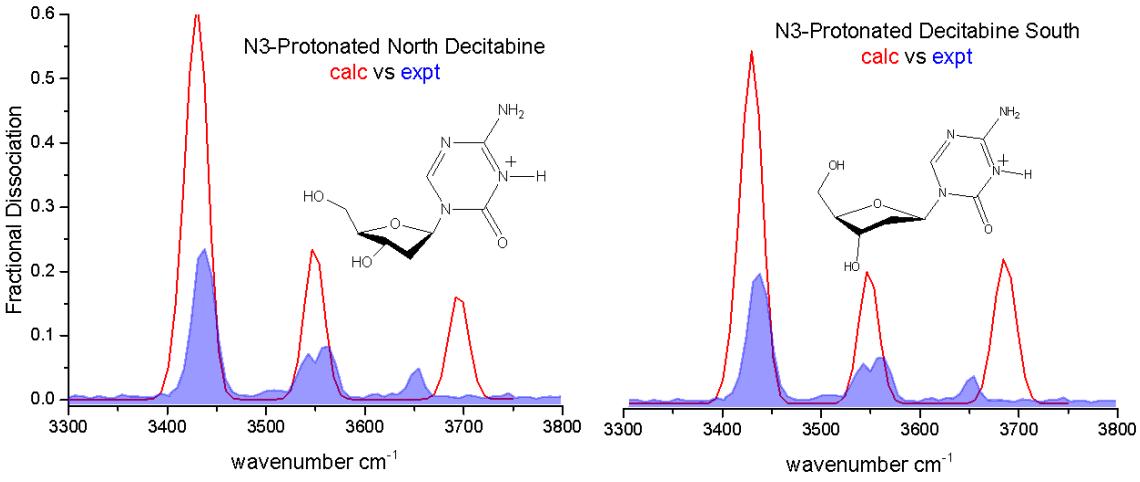


Figure 3.10: Comparison of scaled normal modes between *N3*-protonated decitabine North and *N3*-protonated decitabine South.

An initial look at **Figure 3.11** (which shows the comparison between calculated IR spectra with experimental) seems best to fit the *O*-protonated tautomer of decitabine. The experimental spectrum shows four bands and the calculation of the *O*-protonated tautomer predicts four bands. If this modified nucleoside follows the same trend as the previous two molecules we investigated, one would expect to find an overtone around the same domain. With this in mind, it is possible for the band at 3540 cm^{-1} in the experimental spectrum to correspond to an overtone of a fundamental band. It cannot be said with 100% certainty that the band at 3540 cm^{-1} is indeed an overtone without any IR data in the fingerprint domain. We can however turn to DFT predictions to determine if a fundamental band is present around the 1770 cm^{-1} domain.

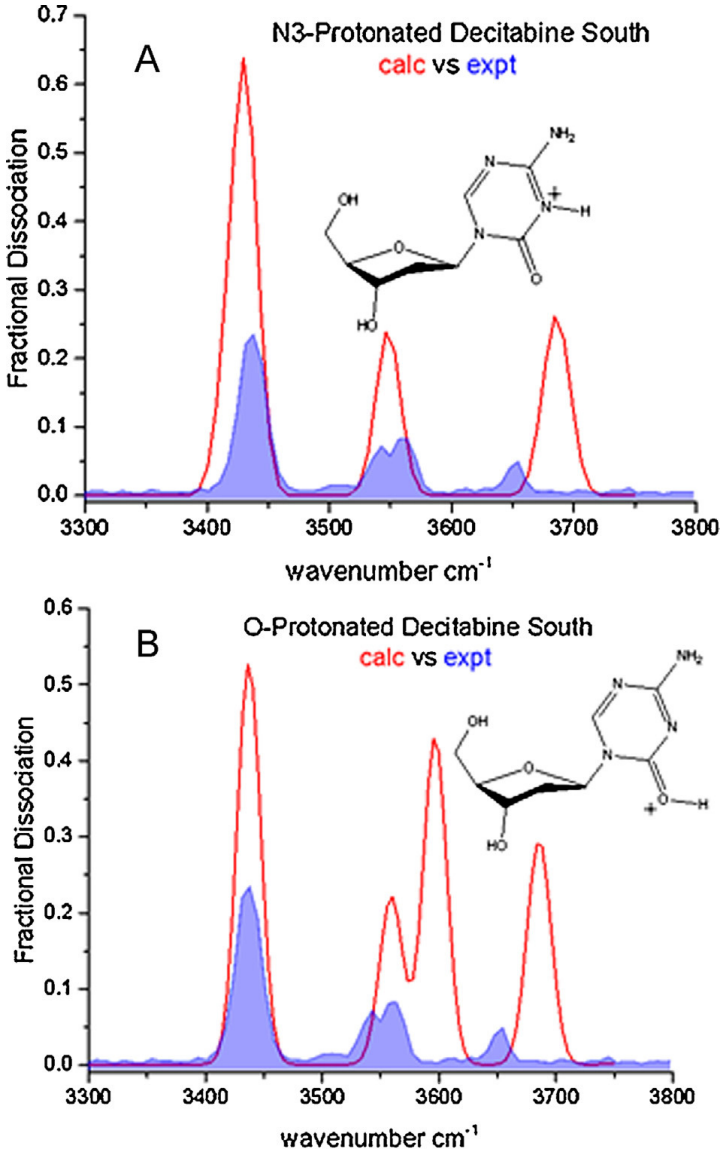


Figure 3.11: Scaled normal modes of **A**) *N3*-protonated decitabine and **B**) *O*-protonated decitabine calculated at the B3LYP/6-311++G** level (red) plotted against experimental IRMPD spectrum (blue). The sugar conformation was calculated in the South orientation because it produced a better fit of peaks against the experimental spectrum when compared to the North orientation.

Due to a lack of data in the fingerprint domain of decitabine, the fingerprint domain of 2'-deoxycytidine and gemcitabine have been investigated to determine the reliability of theoretical predictions. The predicted IR in the fingerprint domain

of 2'-deoxycytidine and gemcitabine were compared with the experimental IRMPD obtained by Speranza *et al.*, and our predicted scaled harmonic IR spectra match experimental fairly well. The calculated IR spectra of *N3*-Protonated 2'-deoxycytidine and *N3*-Protonated gemcitabine both display two strong absorptions corresponding to the amide C=O stretch at 1758 and 1765 cm^{-1}, respectively. These bands are observed experimentally by Speranza *et al.* at 1767 and 1765 cm^{-1}, respectively.

In the case of decitabine, a strong band is observed in the scaled harmonic IR spectrum, corresponding to the amide C=O stretch, at 1779 cm^{-1}. If the actual amide C=O stretch does fall within 10 cm^{-1}, as it did with 2'-deoxycytidine and gemcitabine, then an overtone of this fundamental band is expected at 3558 cm^{-1}. There are two bands that are present in experimental spectrum of decitabine at 3540 and 3560 cm^{-1}. If the band at 3560 cm^{-1} may be an overtone of the fundamental band predicted at 1779 cm^{-1}, then the experimental IRMPD spectrum would best represent the *N3*-Protonated spectrum. The N-H symmetric and asymmetric stretches are predicted at 3420 and 3550 cm^{-1}, respectively. The only other band that is not accounted for in that domain is at 3540 cm-1. This band may be the result of overlapping symmetric and asymmetric N-H stretches. The two O-H stretches from the furanose sugar ring is predicted to be very close at 3680 cm^{-1}, and is observed experimentally at 3650 cm^{-1}.

Partial deuteration experiments help suppress overtones and combination bands of fundamental bands. This method was employed on 2'-deoxycytidine and when compared to the undeuterated spectrum, two bands disappeared. Those bands were determined to be an overtone and a combination band from IR data obtained from the experimental IR previously obtained by Speranza *et al.* in the fingerprint domain. Since partial deuteration was not performed on gemcitabine, the overtone and combination bands can be inferred to exist by looking at the fingerprint IR domain previously obtained by Speranza et al.

Not only can the unnaturally occurring nucleosides form protonated monomers in the gas phase, they can form PBDs. These PBDs do not form naturally in biological systems, but they do provide an alternative view of the *i*-motif, which consists of the same hydrogen bonding motif. For every PBD, there are four different possibilities of ring orientations. It is possible for both sugar rings to be South-$H^{+\cdots}$South, South-$H^{+\cdots}$North, South$^{\cdots}H^{+}$-North, or North-$H^{+\cdots}$North. All four scenarios were accounted for and the predicted IR spectra calculated at the B3LYP/6-311++G** level are shown in **Figure 3.12**. Having both rings in the South-$H^{+\cdots}$South orientation best fits the experimental spectra since it best splits the O-H stretches.

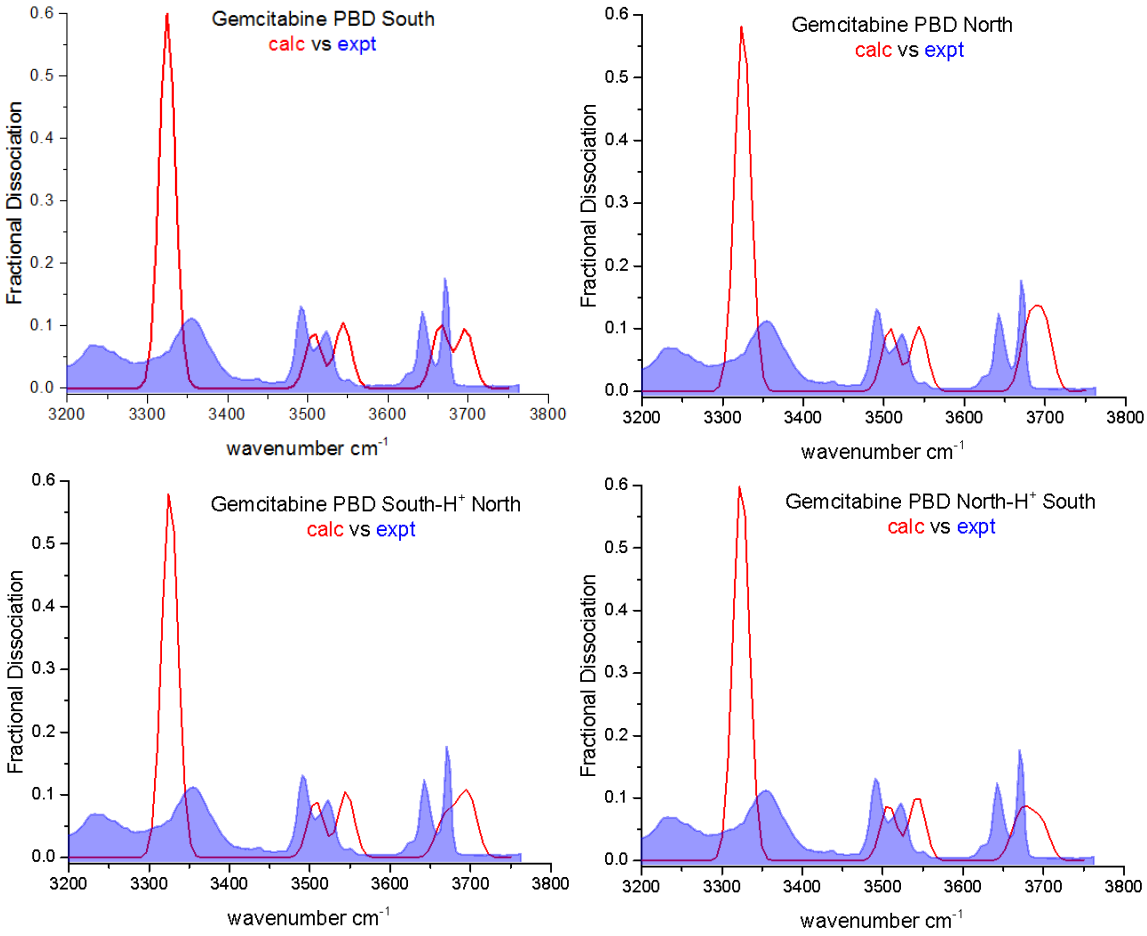


Figure 3.12: Comparsion of DFT predicted IRs with experiment of gemcitabine PBDs with varying sugar ring conformations. The red trace depicts the DFT predicted scaled normal modes while the blue silhouette depicts the experimental IRMPD spectrum. The top left panel has both furanose rings in the South orientation. The top right panel depicts the predicted IR of the PBD with both rings in the North orientation. The bottom left panel has the proton residing on the base with the sugar ring in the South orientation, and the bottom right panel depicts the bridging proton residing on the base with the sugar ring in the North orientation.

Figure 3.13 depicts the protonated PBD of gemcitabine in the CH/NH/OH domain.

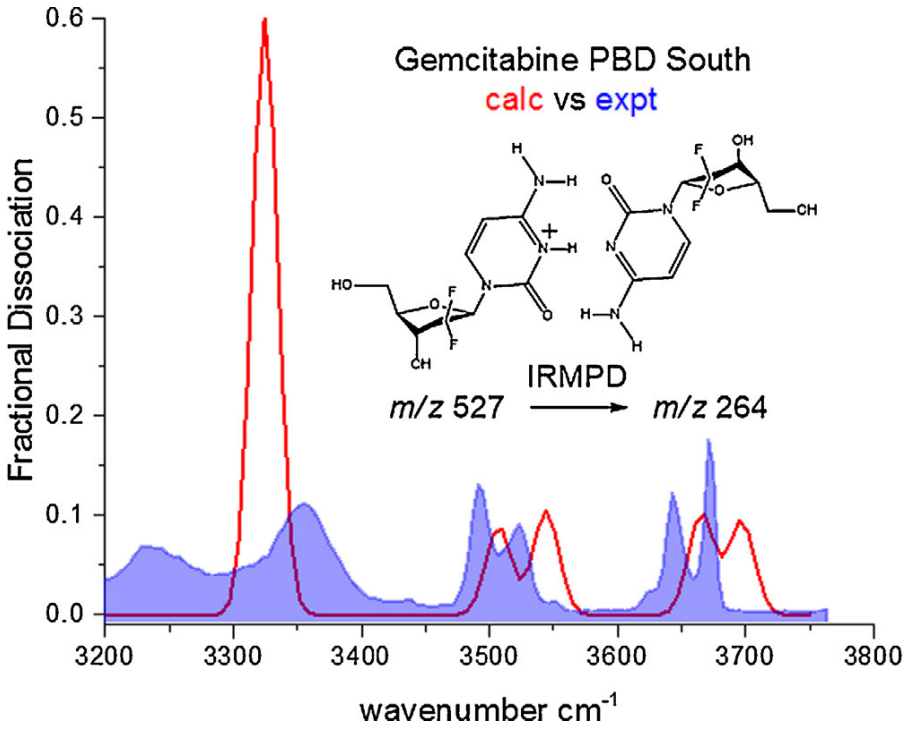


Figure 3.13: Experimental IRMPD spectrum of the PBD of gemcitabine in the CH/NH/OH stretching domain from 3200 – 3800 cm^{-1}. The red trace outlines the calculated scaled normal modes while the blue silhouette depicts the experimental IRMPD spectrum. The parent ion has *m/z* 527 and upon dissociation of the PBD, the protonated daughter ion has *m/z* 264.

The two NH symmetric stretches are predicted to be almost overlapping around 3320 cm^{-1}, and are observed experimentally to overlap at 3350 cm^{-1}. The N-H asymmetric stretches are predicted to not be clustered, but split into two separate resonances at 3500 and 3540 cm^{-1}. Those bands are observed experimentally at 3490 and 3525 cm^{-1}. The OH stretches vary depending if the furanose sugar ring was in the South or North orientation. According to DFT predictions, having the furanose sugar ring in the North orientation does not split the OH resonances, but having the sugar ring in the South position does, as displayed in **Figure 3.12**.

There is no fundamental band predicated near 3240 cm^{-1} where an experimental band is observed. Taking from what was observed with protonated monomers previously investigated, this band may be an overtone or a combination band. From prediction, there is a band at 1626 cm^{-1} observed for the gemcitabine PBD South. It is unclear what this band is until partial deuteration experiments can be conducted.

The IRMPD spectrum of the PBD of decitabine was also recorded and compared to theory. **Figure 3.14** displays the overlapped spectra between theory and experimental. Like that of the PBD of gemcitabine, the ring puckering orientation of the PBD was calculated to be in the South orientation. DFT predicted N-H symmetric stretches at 3360 cm^{-1} and is observed experimentally at 3370 cm^{-1}. The two asymmetric N-H stretches are predicted to be around 3490 and 3540 cm^{-1}. Experimentally, two resonances are observed at slightly lower frequencies at 3480 and 3520 cm^{-1}. The O-H resonances in this case are not split, as seen experimentally. The predicted O-H resonances are broadened out but are still close enough together to be a broad peak. Having the sugar ring in the North orientation brings the two O-H resonances closer together and sharpens the peak.

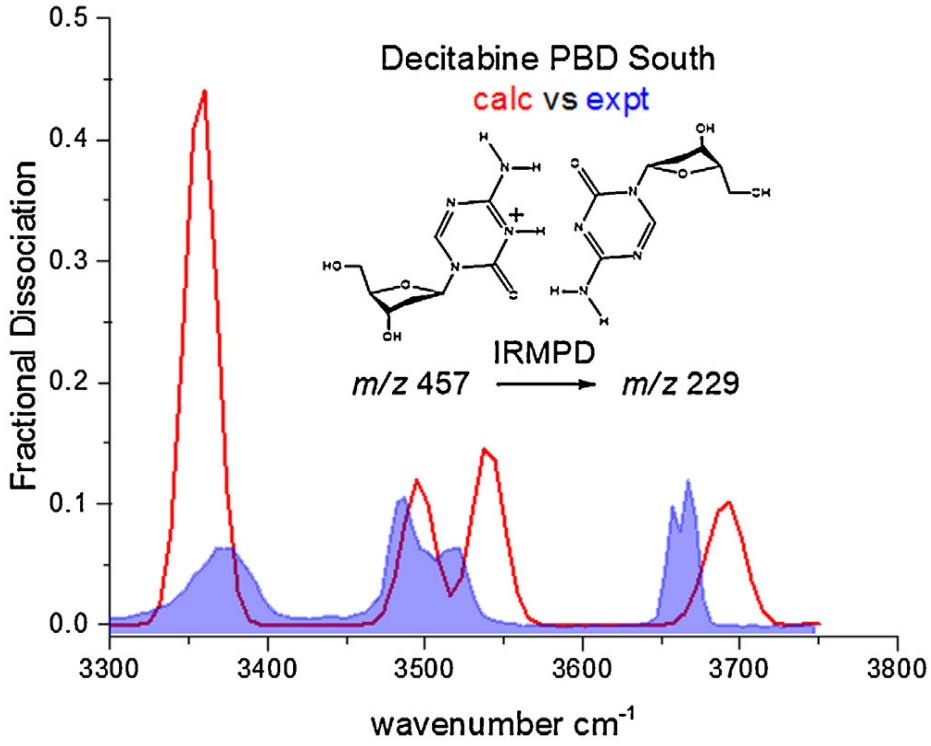


Figure 3.14: Experimental IRMPD spectrum (blue silhouette) plotted against DFT predicted scaled normal modes, calculated at the B3LYP/6-311++G** level (red trace), of the PBD of decitabine in the CH/NH/OH stretching domain from 3300 – 3750 cm^{-1}. The dissociation of the PBD parent ion (*m/z* 457) produces a neutral and a protonated daughter ion (*m/z* 229).

Conclusion

Comparison of the PBD equilibrium constants in **Table 3.1** with the gas phase proton affinities summarized in **Table 3.2**, homodimer/heterodimer equilibria show a markedly diminished sensitivity to basicity. For example, the nucleobase 1-methylcytosine has a proton affinity 40 kJ/mol less than that of 2'-deoxycytidine, yet the equilibrium constant for the 1-methylcytosine proton-bound homodimer going to the heterodimer with 2'-deoxycytidine in **Table 3.1** has a value of K_{eq} = 3.2, only a factor of 1.6 greater than expected on the basis of naïve

statistics. Similarly, conversion of that heterodimer to the proton-bound 2'-deoxycytidine homodimer gives K_{eq} = 1.14, only a factor of 2.3 greater than naïve statistics predict. The proton affinity difference between gemcitabine and decitabine is 20 kJ/mol, but equilibration of the proton-bound decitabine homodimer with the gemcitabine/decitabine heterodimer gives K_{eq} = 11.1, a factor of 5.5 greater than statistical, despite their proton affinity difference having half the value of that between 1-methycytosine and 2'-deoxycytidine. Hence, it does not appear that PBD equilibrium constants can be scaled to obtain basicities of the neutral monomers.

Due to ring puckering of the furanose sugar ring, a protonated monomer of a nucleoside may adopt several different conformations. DFT calculations were performed to compare the relative stabilities of the *N3*-protonated and *O*-protonated tautomer of all three nucleosides under investigation. Not only were the protonation sites examined, the North and South furanose ring puckering was also studied. Comparison of their relative stabilities showed that some nucleosides preferred to be *N3*-protonated, and some *O*-protonated. Having the furanose ring puckering in the South orientation provided the best fit of peaks when theory is compared with experimental spectra.

The IRMPD spectra of 2'-deoxycytidine and gemcitabine were previously published by Speranza *et al.* in the fingerprint and the CH/NH/OH stretching domain. We reproduced the IRMPD spectrum of 2'-deoxycytidine in the CH/NH/OH stretching domain and found that it is almost identical to that

previously published, except for the intensity of some of the stretches in both the spectra of 2'-deoxycytidine and gemcitabine.

Having determined the same bands exist in our spectra as in that previously reported by Speranza *et al.*, the next step is to conduct partial deuteration experiments were four of the five exchangeable protons are exchanged for deuteria. This method is employed to help suppress any overtones or combination bands that may be present in the CH/NH/OH domain.

Overtones and combination bands in the CH/NH/OH domain can be a problem when investigating the CH/NH/OH domain. Our results indicate that overtones and combination bands can successfully be suppressed when partial deuteration experiments are coupled with traditional gas phase IR experiments. In the case of protonated 2'-deoxycytidine, five bands are observed in the completely protonated IRMPD spectrum. When coupled with partial deuteration experiments, two bands disappear and were determined to be an overtone and a combination band. Partial deuteration also causes a shift of the N-H symmetric stretch to a lower frequency in the case of 2'-deoxycytidine.

Although partial deuteration experiments were not performed on gemcitabine, we can infer the same results since they share the same base. The protonated IRMPD spectrum of gemcitabine displayed at least seven bands in the CH/NH/OH domain. DFT predicted the IR spectrum of the *N3*-protonated monomer to have four bands, and the *O*-protonated tautomer to display five bands. Even accounting for an overtone and a combination band, there is still an

additional band unaccounted for when compared to the *N3*-Protonated tautomer. This leads us to believe the IRMPD spectrum of protonated gemcitabine best fit the spectrum of the *O*-protonated tautomer, reaffirming what was stated by Speranza *et al*.

Like that of gemcitabine, partial deuteration experiments were not carried out with decitabine. Experimental IRMPD spectrum in the fingerprint domain of decitbaine cannot be found in literature, so any assignments of overtones and combination bands are based on speculation. DFT predicted IR frequencies provide clues as to which bands may be present in the domain of interest, but cannot be said for certain if those bands actually exist. From the calculated IR frequencies, we can infer the band at 3540 cm^{-1} in the experimental could correspond to an overtone of the predicted fundamental band at 1770 cm^{-1}. If this prediction holds true, then the protonated IRMPD spectrum in the CH/NH/OH stretching domain best fits that of *N3*-protonated decitabine.

The comparison between the calculated IR spectrum and experimental spectrum on the PBD of gemcitabine shows a band that is not predicted around 3240 cm^{-1}. DFT predicted a fundamental band in the fingerprint domain near 1626 cm-1. An overtone of this fundamental band should have a band around 3252 cm^{-1}. The N-H symmetric and asymmetric stretches, and O-H stretches are predicted to be doubled, and is observed so experimentally.

The protonated PBD of decitabine matches that of calculations fairly well, except for the broad peak around 3690 cm^{-1}. That broad peak is calculated to be

the O-H stretches from the sugar rings of the PBD of decitabine. DFT calculations predict the O-H stretches to be close enough to each other that they overlap. Comparison of DFT calculations comparing the orientation of the furanose sugar ring affected the broadness of this peak. Having both furanose sugar rings in the South orientation provided the best fit of peaks, and separated the predicted O-H stretches the farthest from each other. Experimentally, the bands associated to the O-H stretches are doubled.

References

1. B. Yang, R.R.Wu, G. Berden, J. Oomens, M. T. Rodgers “Infrared Multiple Photon Dissociation Action Spectroscopy of Proton-Bound Dimers of Cytosine and Modified Cytosines: Effects of Modification on Gas-Phase Conformations” *J. Phys. Chem. B.* **2013**, 117, 14191 – 14201.

2. A. R. Moehlig, K. E. Djernes, V. M. Krishnan, R. J. Hooley “Cytosine Derivatives Form Hemiprotonated Dimers in Solution and the Gas Phase” *Org. Lett.* **2012**, 14, 2560 – 2563.

3. A. Filippi, C. Fraschetti, F. Rondino, S. Piccirillo, V. Steinmetz, L. Guidoni, M Speranza “Protonated Pyrimidine Nucleosides Probed by IRMPD Spectroscopy” *International Journal of Mass Spectrometry* **2013**, 354, 54 – 61.

4. H. U. Ung, A. R. Moehlig, R. A. Kudla, L. J. Mueller, J. Oomens, G. Berden, T. H. Morton “Proton-Bound Dimers of 1-Methylcytosine and its Derivatives: Vibrational and NMR Spectroscopy” *Phys. Chem. Chem. Phys.* **2013**, 15, 19001 – 19012.

5. G. Dellepiane, J. Overend “Vibrational Spectra and Assignment of Acetone, α,α,α Acetone-d_3 and Acetone-d_6” *Spectrochimica Acta* **1966**, 22, 593 – 614.

6. J-L. Mergny, L. Lavroix, X. Han, J-L Leroy, C. Helene “Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an i-DNA Motif” *J. Am. Chem. Soc.* **1995**, 117, 8887 – 8898.

7. E. Arias “Investigating Proton-Bound Nucleobase Dimers in the Gas Phase” Undergraduate Honors Thesis. University of California, Riverside, **2012**.

8. I. K. Webb, C. E. Muetterties, C. B. Platner, J. C. Poutsma “Gas-Phase Acidities of Lysine Homologues and Proline Analogs From the Extended Kinetic Method” *International Journal of Mass Spectrometry* **2012**, 316, 126 – 132.

9. Z. Wu, C. Fenselau, R. G. Cooks “Gas-Phase Basicities and Proton Affinities of Lysine and Histidine Measured from the Dissociation of Proton-Bound Dimers” *Rapid Communications in Mass Spectrometry* **1994**, 8, 777 – 780.

10. B.A. Cerda, C. Wesdemiotis "Li^+, Na^+, and K^+ Binding to the DNA and RNA Nucleobases. Bond Energies and Attachment Sites from the Dissociation of Metal Ions-Bound Heterodimers" *J. Am. Chem. Soc.* **1996**, 118, 11184 – 11892.

11. P. B. Armentrout "Entropy Measurements and the Kinetic Method: A Statistically Meaningful Approach" *Journal of the American Society for Mass Spectrometry* **2000**, 11, 371 – 379.

12. K. E. Ervin, P. B. Armentrout "Systematic and Random Errors in Ion Affinities and Activation Entropies From the Extended Kinetic Method" *Journal of Mass Spectrometry* **2004**, 39, 1004 – 1015.

Chapter IV: Cytosine and Isocytosine Derivatives

Introduction

According to the Center for Disease Control (CDC), cancer is the second leading cause of deaths in the United States in 2014 (with heart disease being the first)[1]. In 2014, 1,665,540 new cancer cases were diagnosed and 585,720 deaths reported[2].

Cancers of the stomach and the respiratory system are of interest to us due to the acidic environment of the two organs. Gastric acid in the stomach produces an acidic environment. High levels of carbon dioxide in the lungs are in equilibrium with water to produce carbonic acid and decrease the pH. Of the reported cancer cases in 2014, 264,770 of those cases are attributed to stomach cancer (22,220 cases) and respiratory cancer (242,550 cases)[2].

Some of the risk factors that can increase a person's chance of acquiring stomach cancer include: a diet high in smoked foods, eating aflatoxin fungus contaminated foods, a family history of stomach cancer, and smoking[3]. Some risk factors that may increase a person's chance of getting lung cancer include exposure to: smoking, secondhand smoke, asbestos, and a family history of lung cancer[4].

Currently, there is no cure for cancer, though a variety of palliative treatments are available. As previously discussed, the *i*-motif is believed to form under at or below pH 6 *in vitro*. If external conditions affect the cell nucleus, the acidic environments provided by the stomach and the respiratory system should

facilitate *i*-motif formation. Stabilization of the *i*-motif could halt the spread of cancerous cells.

Background

Due to the formation of the G-quadruplexes and *i*-motifs in promoter regions of DNA, the stabilization of these secondary structures could provide another avenue in combating cancer. Balasubramanian *et al.* showed G-quadruplex formation occurs in the human genome. Greater than 40% of the genes possessing one or more promoter regions have sequences compatible with G-quadruplex formation in one of the DNA strands[5]. For every G-quadruplex formed on one strand, there should be a sequence that can form an *i*-motif on the opposite strand.

Several examples of potential Class I *i*-motif-containing promoter regions are to be found in the VEGF, c-Kit, and Kras. Class I *i*-motifs a have relatively small number of bases in the loop regions. VEGF is an important growth factor involved in the production of new blood vessels[6]. VEGF is overexpressed in many tumors such as ovarian, bladder and pancreatic cancer and is vital to tumor growth[6,7]. Overexpression of the VEGF oncogene leads to self-induced angiogenesis[6]. Angiogenesis is the process in which new blood vessels can form from pre-existing blood vessels. For a tumor to have the capability to grow new blood vessels allows it to spread and grow uncontrollably throughout the host.

Two potential Class II *i*-motif-containing promoters (c-Myc and bcl-2) affect the growth of cancer cells in different ways. The c-Myc oncogene has been widely studied over the years. C-Myc falls into the Class II classification of *i*-motifs as it has large number of bases in the loop region. The protein associated with c-Myc promotes the activation of telomerase and forces uncontrollable cell growth[8]. This gene is normally regulated, but overexpression of this gene has been linked to tumor growth[9]. Several cancers can be linked to the over expression of c-Myc such as lymphomas, leukaemia, and lung, cervical, ovarian, breast, pancreatic, and gastric cancers[10]. Another oncogene that has been studied is bcl-2. Bcl-2 falls under Class II as it also has a large number of bases in the loop region. Bcl-2 affects the growth of cancer cells by disrupting the apoptotic function of the cell[11].

The use of organic ligands to bind to and stabilize the *i*-motif has been investigated[12,13]. One of the first examples of binding to the *i*-motif was reported by Hurley *et al.* in 2000, who reported using a tetraammonium porphyrin to promote *i*-motif formation[12]. The same tetraammonium porphyrin (**Figure 4.1**) was also found to down regulate the expression of c-Myc by interacting with the G-quadruplex [14]. Since the tetraammonium porphyrin can bind to different secondary structures of DNA, it is possible that the observed effects in c-Myc expression could be due to the relationship between G-quadruplex and *i*-motif. That is to say, a stabilization of one could lead to the stabilization of the other,

without specifying which one is stabilized through direct interaction with the ligand.

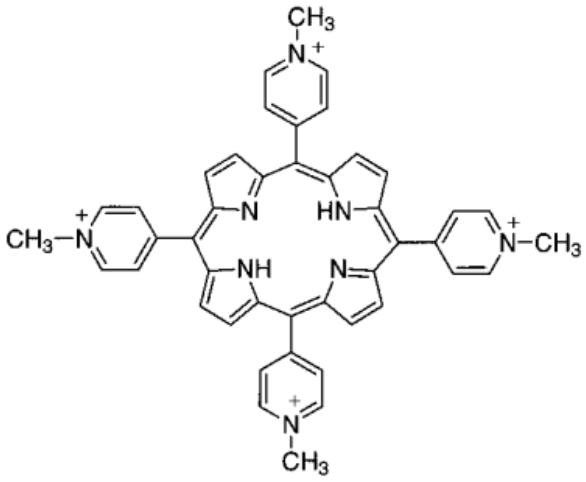


Figure 4.1: Tetraammonium porphyrin used to promote *i*-motif formation.

Another method of inducing i-motif formation is to use single-walled carbon nanotubes (SWNTs). In 2006, Qu *et al.* found SWNTs can inhibit duplex association of DNA and promote *i*-motif formation not only in physiological conditions, but also in slightly basic conditions at pH as high as 8[15].

Proton resonances between 15 - 16 ppm have been shown to correspond to the bridging proton [15,16]. At pH 5, several peaks are observed between 15 – 16 ppm (**Figure 4.2**). As the pH is increased to 8, i-motif formation is not observed. The introduction of SWNTs in the same pH 8 solution again shows the appearance of bands between 15 – 16 ppm. Since the SWNTs are relatively large in size compared to DNA (1.1 nm), binding is proposed to take place between SWNTs and the 5'-end of the *i*-motif.

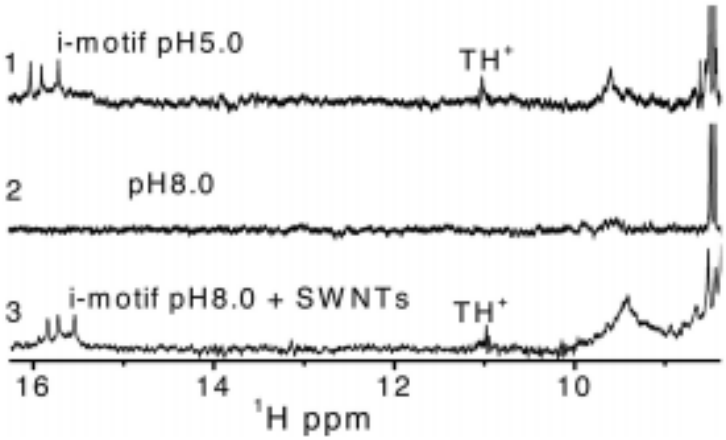
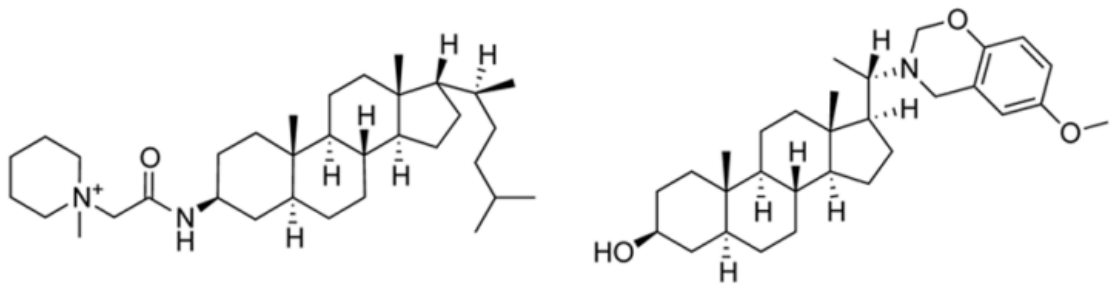


Figure 4.2: **1**) *i*-motif formation shown by three peaks between 15 – 16 ppm. **2**) i-motif formation is absent in basic conditions. **3**) the addition of SWNTs promote i-motif formation even in basic conditions. TH* denotes the imino group of thymine in the oligimer.

Hurley *et al.* recently examined the stability of the bcl-2 and showed bcl-2 can exist in equilibrium between hairpin structure and *i*-motif[17]. Two ligands (**Figure 4.3**) were identified using a FRET-based high throughput assay which screened 1990 compounds. These two ligands are selective toward bcl-2. Experiments with other i-motif forming DNA sequences, such as c-Myc and VEGF, did not show any significant change.



IMC-48 **IMC-76**

Figure 4.3: Two ligands investigated by Hurley *et al.* that was found to affect *i*-motif stability.

Using ^{1}H NMR analysis, IMC-48 was found to stabilize and favor *i*-motif formation, while IMC-76 destabilizes the *i*-motif and favors a hairpin structure (**Figure 4.4**). Hurley et al. suggest IMC-48 binds within the loop region of the i-motif due to stacking interactions with thymines within the loop region. This hypothesis is tested by using the c-Myc sequence to test IMC-48 binding. IMC-48 had no effect on c-Myc, but when a hybrid combination of c-Myc and bcl-2 was treated with IMC-48, the same effect was observed to that of pure bcl-2.

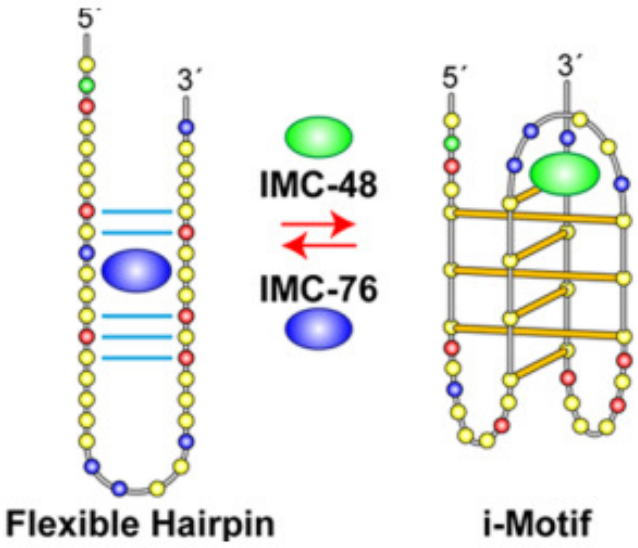
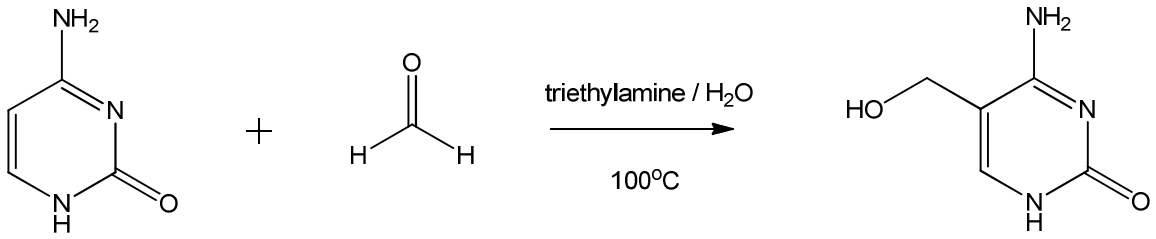


Figure 4.4: Artist rendition of IMC-48 stabiliing the *i*-motif and IMC-76 stabilizing the hairpin structure.

Guanidinylated nucleosides have been synthesized over a decade ago[18,19]. This chapter presents data on two target molecules. Our proposed target molecules differ from the previously investigated guanidinylated nucleosides in the sense that the guanidine is incorporated into a ring as to avoid unwanted hydrogen bonding interactions. Our target molecules also consist of three rings. The middle rings in our target molecules are envisioned to restrict the molecule into a conformation favorable for hydrogen bonding (**Figure 4.1**). These two molecules were designed to either bind to and stabilize the loop region of the *i*-motif, like IMC-48, or to the *i*-motif residues that flip away from the *i*-motif core due to charge repulsion. Since our target molecules are designed to bind to either protonated or neutral cytosine, they are candidates to bind to the residues that flip out of the *i*-motif core.

Synthesis and Characterization / Methods

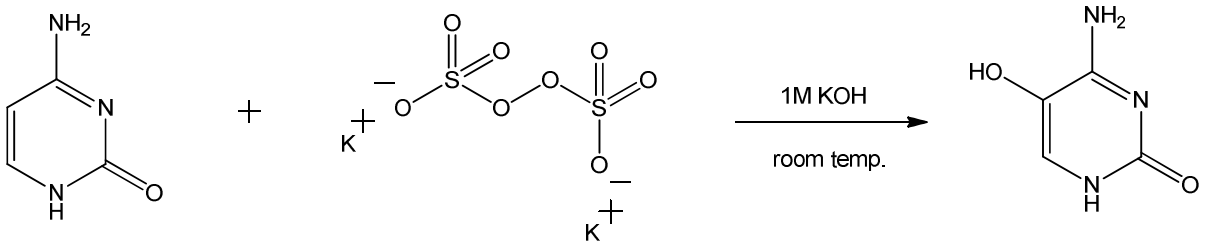
Many of the procedures described below do not give an isolable yield of the desired product. They are included so as to direct future investigators as to what has already been attempted.



Scheme 4.1[20]: Synthesis of 5-hydroxymethylcytosine starting from cytosine and paraformaldehyde.

The synthesis of 5-hydroxymethylcytosine followed a previously published procedure[20]. To 100 mg (0.00091 mol) of cytosine in 10 ml H_2O and 2 mL triethylamine in a 25 mL round bottom flask was added 94 mg paraformaldehyde (0.0031 mol). Paraformaldehyde is added every 12 hours in 15 mg increments. The resulting mixture was heated and stirred at 100°C for 72 hours. When the reaction has reached completion, the solution was concentrated under reduced pressure, leaving behind an oily solid in the flask. The mass of the oily solid exceeded that of 100% yield. It is possible for the solid within the oil to be repolymerized paraformaldehyde.

APCI-ESI MS: MH^+: 142.0607 (calc. 142.0610)

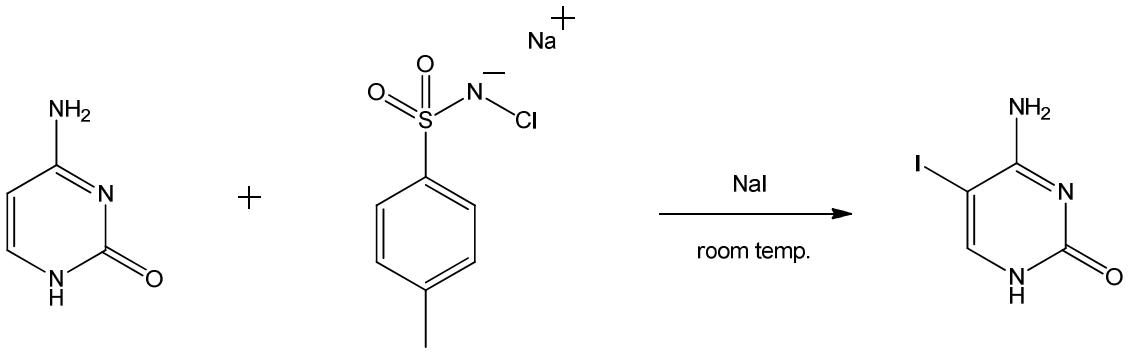


Scheme 4.2[21]: Synthesis of 5-hydroxycytosine.

The synthesis of 5-hydroxycytosine followed a previously published procedure[21].To 100 mg (0.0009 mol) of cytosine in 10 mL 1M aqueous KOH in a 25 mL round bottom flask, was added 365 mg (0.00135 mol) potassium persulfate. The reaction mixture was stirred at room temperature for 18 hours. After 18 hours of stirring, the reaction was quenched with 2 mL conc. HCl and the solid filtered. Mass spectrum of the white solid showed the desired product. A crude yield of 68 mg (60%) was obtained.

APCI-ESI MS: MH^+: 112.0510 (calc. 112.0505)

^{1}H NMR (300 MHz, D_2O): δ 7.83 (1H s)

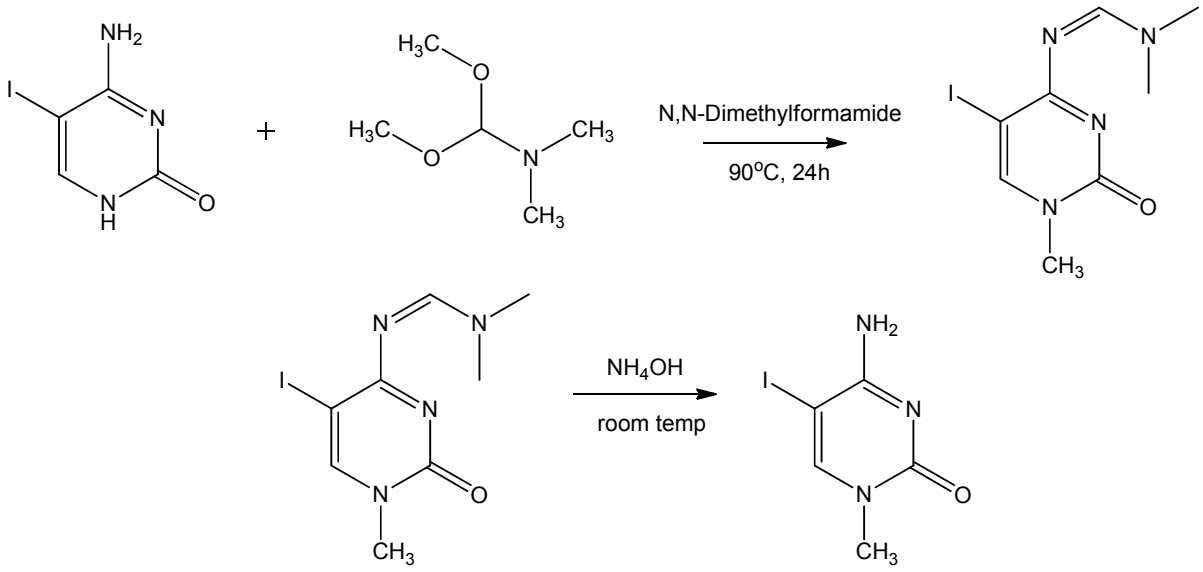


Scheme 4.3[22]: Synthesis of 5-iodo-cytosine.

The iodination of cytosine followed a previously published procedure which utilized sodium iodide and chloramine-T to iodinate phenol[22].To 1 g (0.009 mol) of cytosine in 10 ml DI H_2O in a 25 ml round bottom flask was added 2.7 g (0.018 mol) sodium iodide and 2.04 g (0.009 mol) chloramine-T. The reaction mixture was stirred at room temperature for 18 hours. After 18 hours of stirring, the solid was in the round bottom flask is vacuum filtered and recrystallized using hot DI H_2O. A yield of 910 mg (43%) was obtained.

APCI-ESI MS: MH^+: 237.9467 (calc. 237.9472)

^{1}H NMR (300 MHz, D_2O): δ 7.94 (1H s)

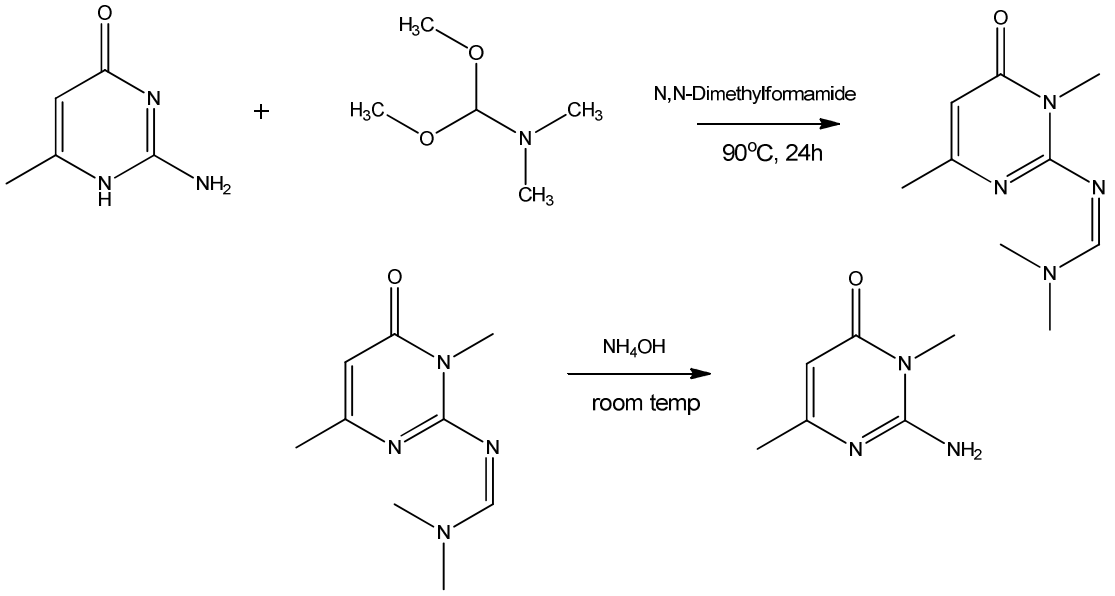


Scheme 4.4[23]: Synthesis of 5-iodo-1-methylcytosine by treating 5-iodocytosine with N,N-dimethylformamide dimethyl acetal.

To 910 mg (0.0038 mol) of 5-iodocytosine in 10 ml N,N-dimethylformamide (DMF) in a 100 ml round bottom flask equipped with a stir bar was added 2.55 ml (0.019 mol) of N,N-dimethylforamide dimethyl acetal. The resulting solution was heated for 12 hours at 90^{o}C with a reflux condenser attached. After cooling to room temperature, the solution was evaporated to dryness and 20 mL of conc. ammonium hydroxide is added. The solution was stirred at room temperature overnight. A yield of 790 mg (82%) was obtained after recrystallization from absolute ethanol.

APCI-ESI MS: MH^+: 251.9649 (calc. 251.9628)

^{1}H NMR (300 MHz, D_2O): δ 3.41 (3H s) δ 8.06 (1H s)



Scheme 4.5: Synthesis of 3,6-dimethylisocytosine by treating 6-methylisocytosine with N,N-dimethylformamide dimethyl acetal. Deprotection using conc. Ammonium hydroxide at room temperature yielded 3,6-dimethylisocytosine.

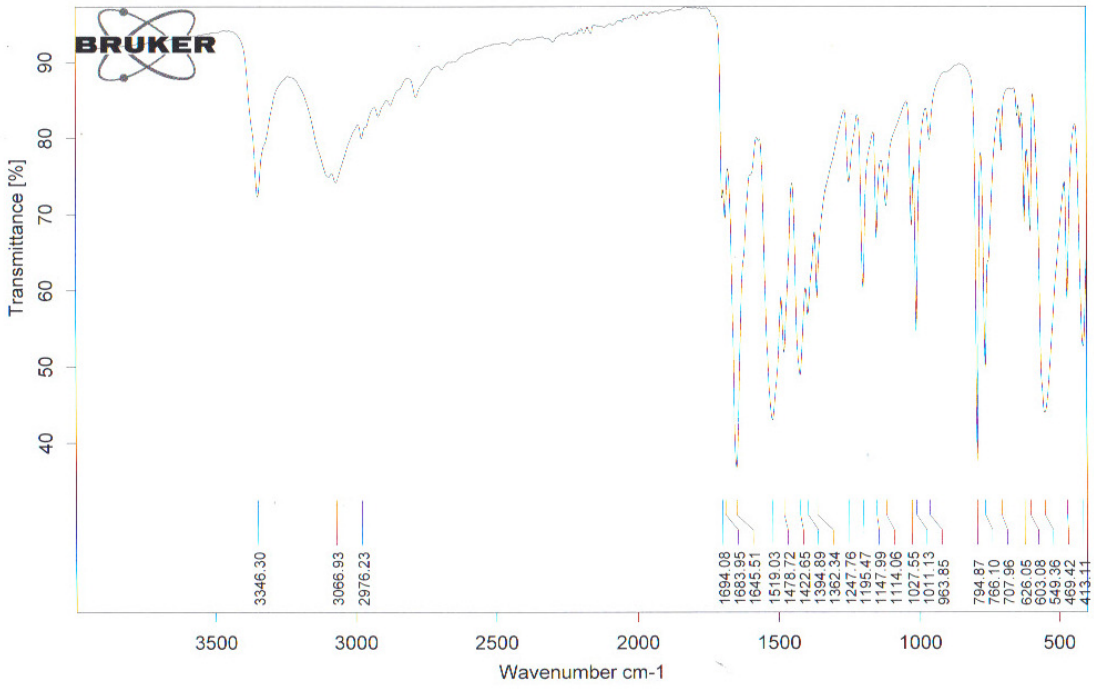
The original target was 1,6-dimethylisocytosine. Using the same methylation procedure as before to methylate cytosine, the end product turns out to be 3,6-dimethylcytosine. 3,6-Dimethylisocytosine was synthesized as depicted in **Scheme 4.5**. To 1.0g (0.008 mol) of 6-methylisocytosine in 10 mL N,N-dimethylformamide (DMF) in a 100 mL round bottom flask equipped with a stir bar was added 5.3 ml (0.04 mol) of N,N- dimethylformamide dimethyl acetal. The resulting solution was heated for 12 hours at 90^{o}C with a reflux condenser attached. After cooling to room temperature, the solution was evaporated to dryness and 20 mL of conc. ammonium hydroxide is added. The solution was stirred at room temperature overnight. A yield of 0.745g (67%) was obtained after recrystallization from absolute ethanol.

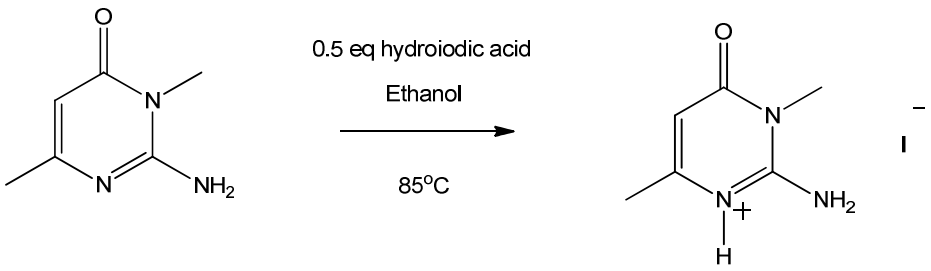
APCI-ESI MS: MH^+: 140.0824 (calc. 140.0818)

MP: 327 – 330°C

^{1}H NMR (400 MHz, DMSO-d_6): δ 1.98 (3H s) δ 3.20 (3H s) δ 5.48 (1H s) δ 7.03 (2H b)

^{13}C NMR (400 MHz, DMSO-d_6): δ 24.0 (1C) δ 28.0 (1C) δ 99.5 (1C) δ 156.4 (1C) δ162.4 (1C) δ 164.4 (1C)

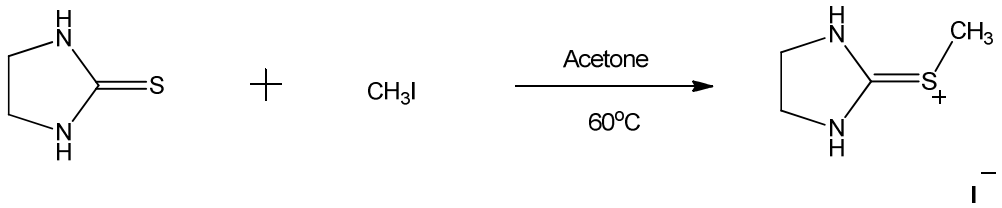




Scheme 4.6: Preparation of protonated 3,6-dimethylisocytosine iodide salt using 3,6-dimethylisocytosine and hydroiodic acid.

APCI-ESI MS: MH^+: 279.1562 (calc. 279.1563)

MP: 178-185°C

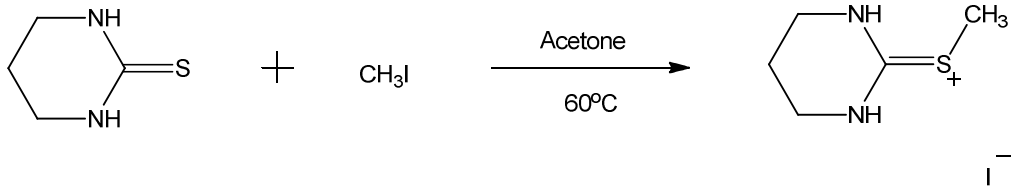


Scheme 4.7[24]: Synthesis of methylated ethylenethiourea iodide salt.

The synthesis of methylated ethylenethiourea iodide salt followed a previously published procedure[24]. To 2.5 g (0.024 mol) of ethylenethiourea in 25 mL acetone in a 50 ml round bottom flask was added 1.5 ml (0.024 mol) methyl iodide, dropwise. The reaction mixture was heated and stirred at 65°C for 30 minutes. After cooling to room temperature, the white solid is filtered. A yield of 4.52 g (77%) was obtained.

APCI-ESI MS: MH^+: 117.0499 (calc. 117.0481)

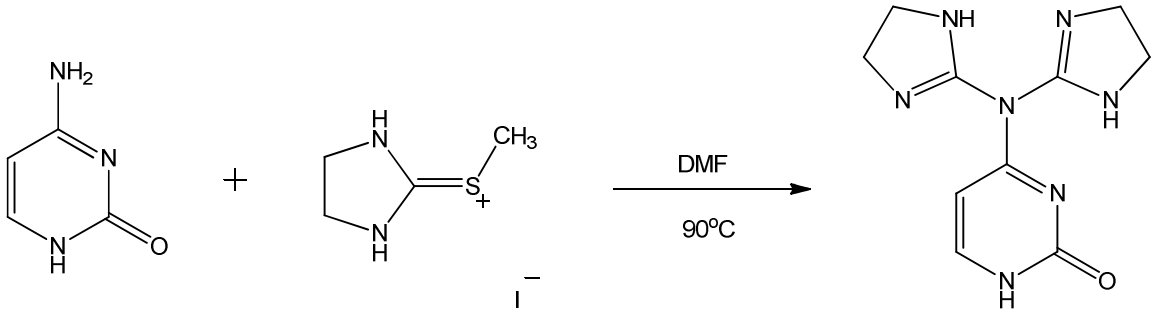
^{1}H NMR (300 MHz, D_2O): δ 2.65 (3H s) δ 3.97 (4H s)



Scheme 4.8: Synthesis of methylated propylenethiourea iodide salt.

The synthesis of methylated propylenethiourea iodide salt followed a previously published procedure[24], but propylenethiourea was used in place of ethylenethiourea. To 1 g (0.0086 mol) of propylenethiourea in 25 mL acetone, in a 50 mL round bottom flask was added dropwise 0.53 ml (0.0086 mol) methyl iodide. The reaction mixture was heated and stirred at 65°C for 30 minutes. After cooling to room temperature, the white solid is filtered. A yield of 2.01 g (92%) was obtained.

^{1}H NMR (300 MHz, D_2O): δ 2.04 (2H p) δ 2.58 (3H s) δ 3.49 (4H t)

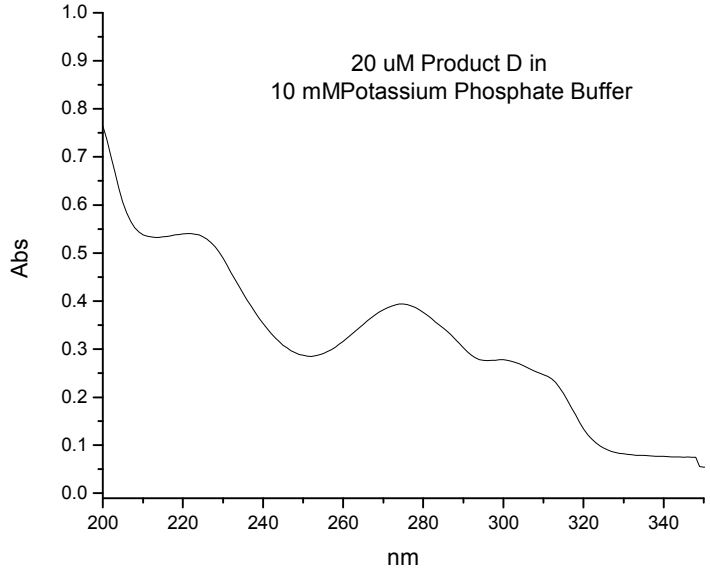


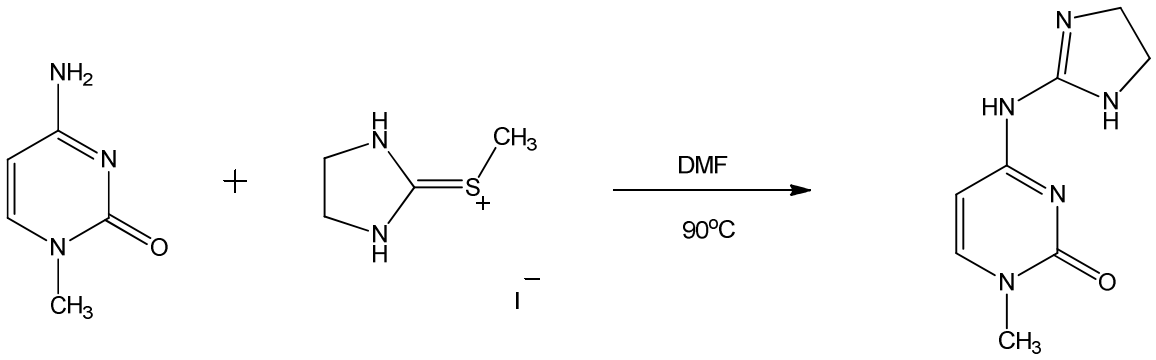
Scheme 4.9: Synthesis of N,N-bis-ethyleneguanidinecytosine.

To 500 mg (0.0045 mol) of cytosine in 5 mL DMF in a 25 mL round bottom flask was added 1.2 g (0.005 mol) methylated ethylenethiourea iodide salt. The reaction mixture was heated and stirred at 90°C for 48 hours. After 48 h, the reaction was cooled to room temperature and the solid filtered. Recrystallization of the filtered solid using absolute ethanol yields a white solid. A yield of 72 mg (6.5%) was obtained.

APCI-ESI MS: MH^+: 248.1288 (calc. 248.1254)

^{1}H NMR (400 MHz, D_2O): δ 3.90 (4H s) δ 3.95 (2H t) δ 4.06 (2H t) δ 6.34 (1H d) δ 7.74 (1H d)



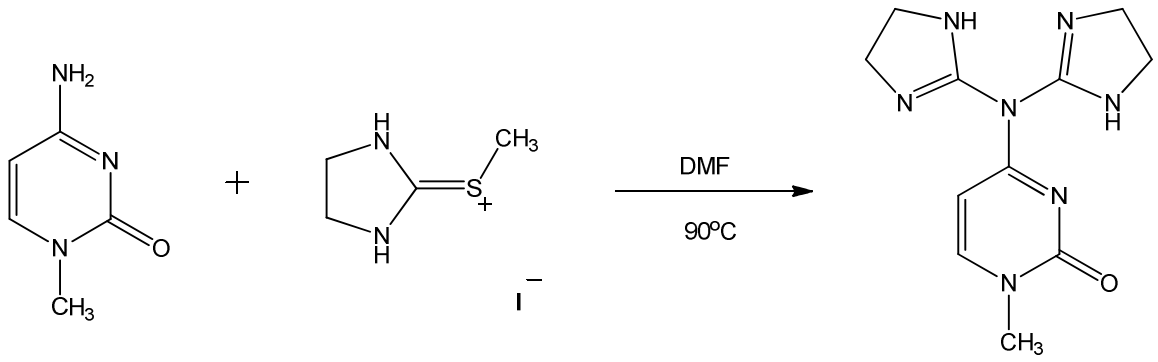


Scheme 4.10: Synthesis of N-ethyleneguanidine-1-methylcytosine.

To 500 mg (0.004 mol) of 1-methylcytosine in 5 mL DMF in a 25 mL round bottom flask was added 1.07 g (0.0044 mol) methylated ethylenethiourea iodide salt. The reaction mixture was heated and stirred at 90^{o}C for 48 hours. After 48 h the reaction was cooled to room temperature and the precipitated solid filtered. Mass spectrometric analysis of the crude product shows the desired product,

along with a side product (see **Scheme 4.11**). Product was not isolated because majority of crude is the side product as described below.

APCI-ESI MS: MH^+: 194.1045 (calc. 194.1036)



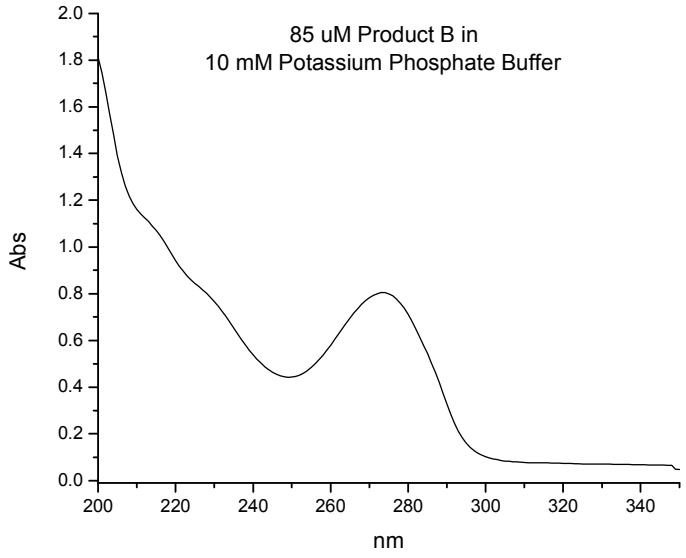
Scheme 4.11: Synthesis of N,N-bis-ethyleneguanidine-1-methylcytosine.

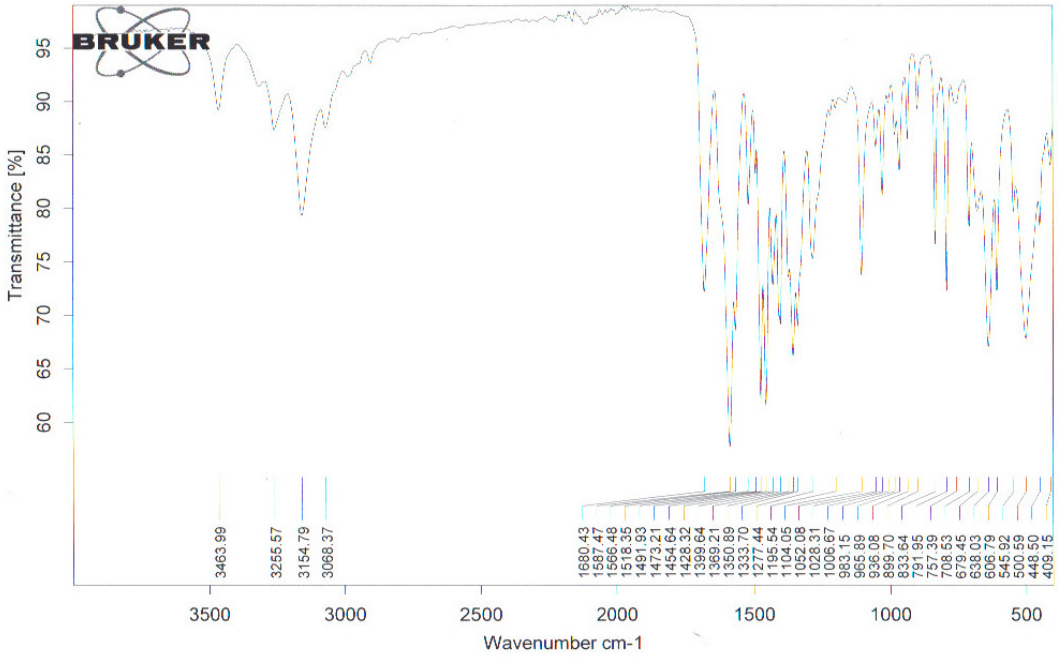
To 500 mg (0.004 mol) of 1-methylcytosine in 5 mL DMF in a 25 mL round bottom flask was added 1.07 g (0.0044 mol) methylated ethylenethiourea iodide salt. The reaction mixture was heated and stirred at 90°C for 48 hours. After 48 hours the reaction was cooled to room temperature and the solid filtered. Recrystallization of the filtered solid using absolute ethanol yields a white solid. A yield of 182 mg (17.5%) was obtained.

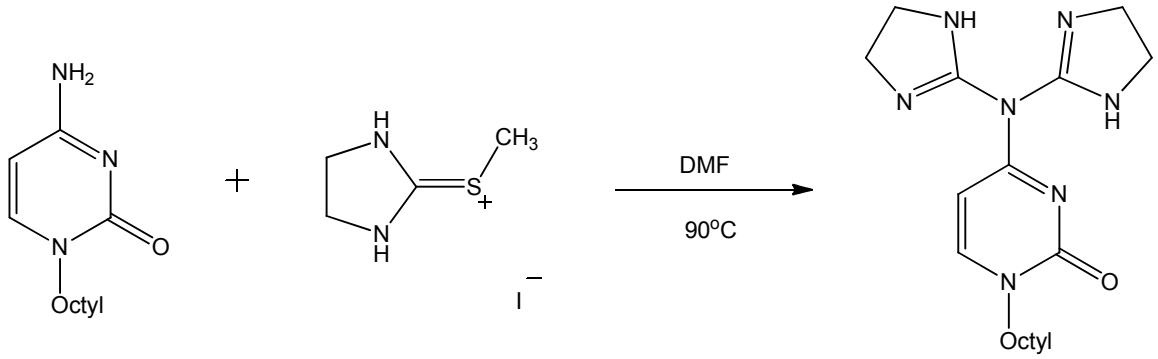
APCI-ESI MS: MH^+: 262.1437 (calc. 262.1411)

^{1}H NMR (400 MHz, D_2O): δ 3.5 (3H s) δ 3.95 (4H s) δ 4.07 (2H t) δ 4.06 (2H t) δ 6.35 (1H d) δ 7.84 (1H d)

^{13}C NMR (400 MHz, D_2O): δ 38.24 (1C) δ 40.99 (1C) δ 43.00 (2C) δ 43.36 (1C) δ 104.65 (1C) δ 149.00 (1C) δ 155.85 (1C) δ 159.13 (1C) δ 170.77 (1C)


85 uM Product B in
10 mM Potassium Phosphate Buffer
Abs
nm
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
200
220
240
260
280
300
320
340

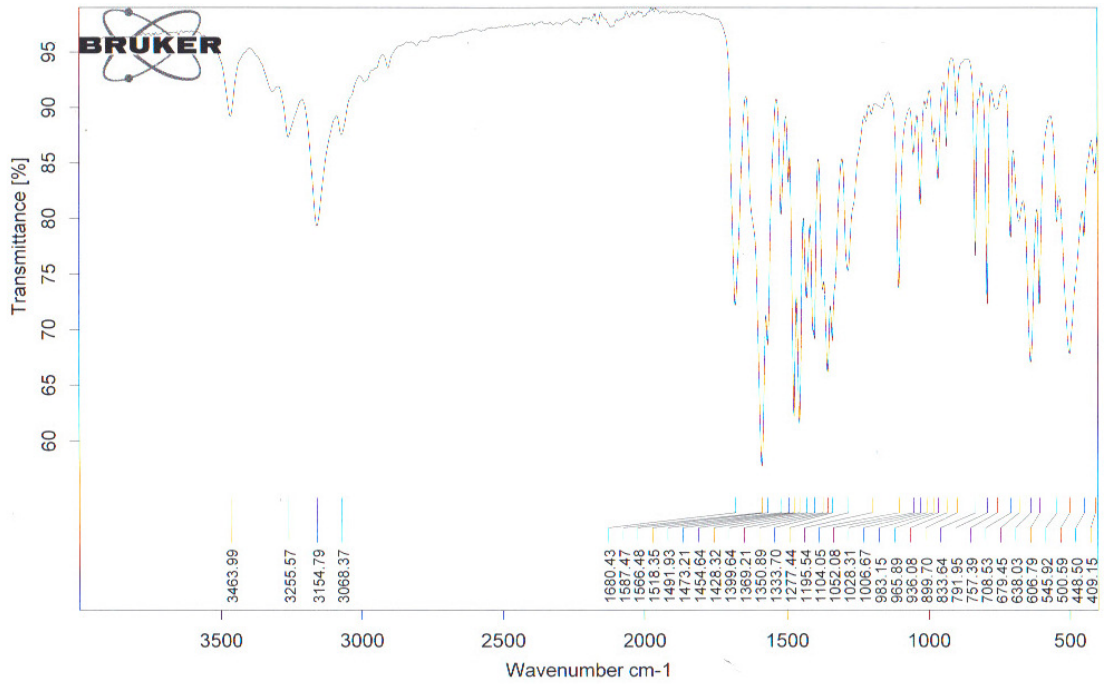

BRUKER
Transmittance [%]
95
90
85
80
75
70
65
60
3463.99
3255.57
3154.79
3068.37
1680.43
1587.47
1566.48
1518.35
1491.93
1473.21
1454.64
1428.32
1399.64
1389.21
1350.89
1333.70
1277.44
1195.54
1104.05
1052.08
1028.31
1006.67
983.15
965.89
936.08
899.70
833.64
791.95
757.39
708.53
679.45
638.03
606.79
545.92
500.59
448.50
409.15
3500
3000
2500
2000
1500
1000
500
Wavenumber cm-1

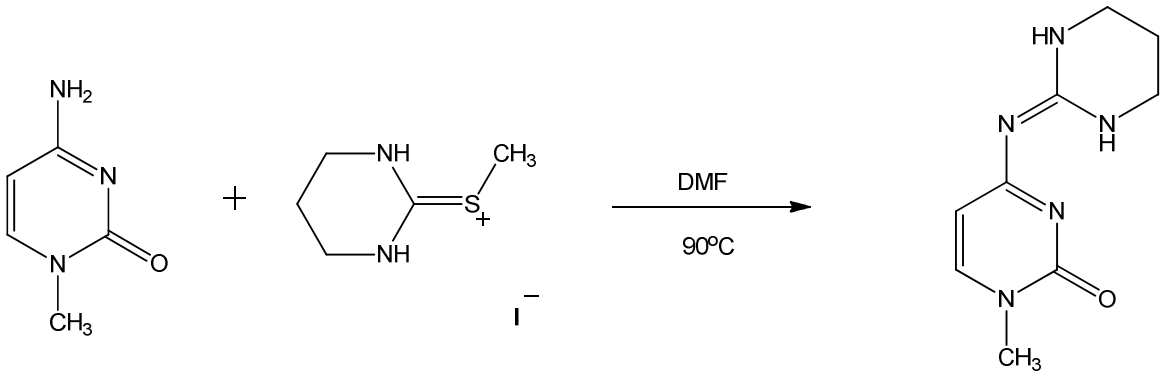


Scheme 4.12: Synthesis of N,N-bis-ethyleneguanidine-1-octylcytosine.

To 100 mg (0.0044 mol) of 1-octylcytosine in 5 mL DMF in a 25 mL round bottom flask was added 1.63 g (0.0066 mol) methylated ethylenethiourea iodide salt. The reaction mixture was heated and stirred at 90^{o}C for 48 hours. After 48 hours, the reaction was cooled to room temperature and the volume reduced under vacuum. ^{1}H NMR and mass spectrometry analysis of the crude product shows the presence of starting materials along with the desired product.

APCI-ESI MS: MH^{+}: 360.2496 (calc. 360.2506)



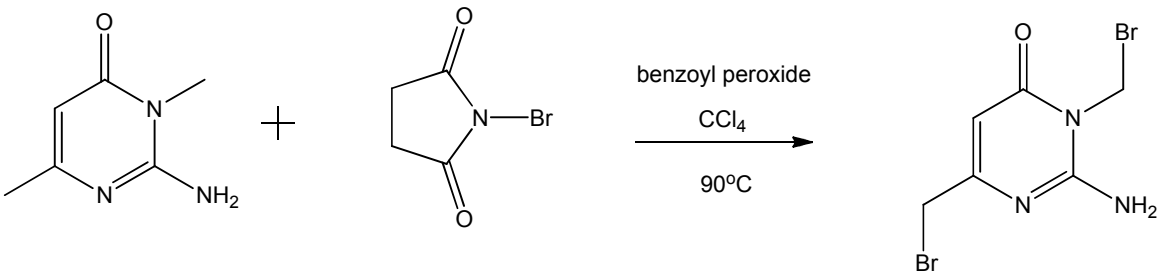


Scheme 4.13: Synthesis of N-propyleneguanidine-1-methylcytosine.

To 100 mg (0.0008 mol) of 1-methylcytosine in 4 mL DMF in a 25 mL round bottom flask was added 206 mg (0.0008 mol) methylated propylenethiourea iodide salt. The reaction mixture was heated and stirred at 90^{o}C for 48 hours. After 48 hours, the reaction was cooled to room temperature

and the solid filtered. Mass spectrum of the crude product showed the desired product, along with other side products. The intensities of the side products were much higher than the desired product.

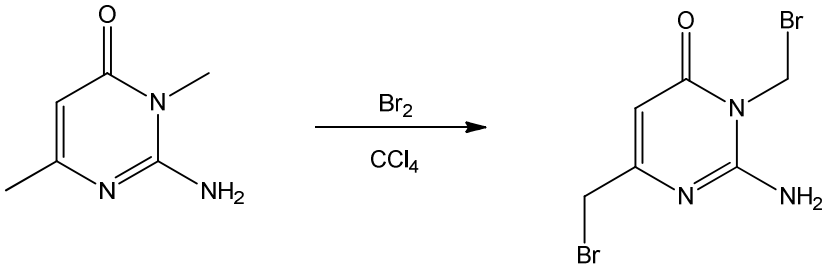
APCI-ESI MS: MH^+: 208.1203 (calc. 208.1193)



Scheme 4.14: Bromination of 3,6-dimethylisocytosine using N-bromosuccinimide and benzoyl peroxide.

To 100 mg (0.00072 mol) of 3,6-dimethylisocytosine in 10 mL carbon tetrachloride in a 25 mL round bottom flask is added 256 mg (0.00144 mol) N-bromosuccinimide. 70 mg Benzoyl peroxide (20 mol%) was added to the reaction mixture and the reaction mixture heated and stirred at 80°C for 16 hours. After stirring, the reaction was then quenched with 20 mL of DI H_2O. The organic layer was separated from the aqueous layer using a separatory funnel and dried with anhydrous sodium sulfate. The solution was decanted into a 50 ml round bottom flask and concentrated under reduced pressure, leaving behind an oily residue on the flask. An exact mass was not obtainable but the mass spectrum of the residue showed the desired product.

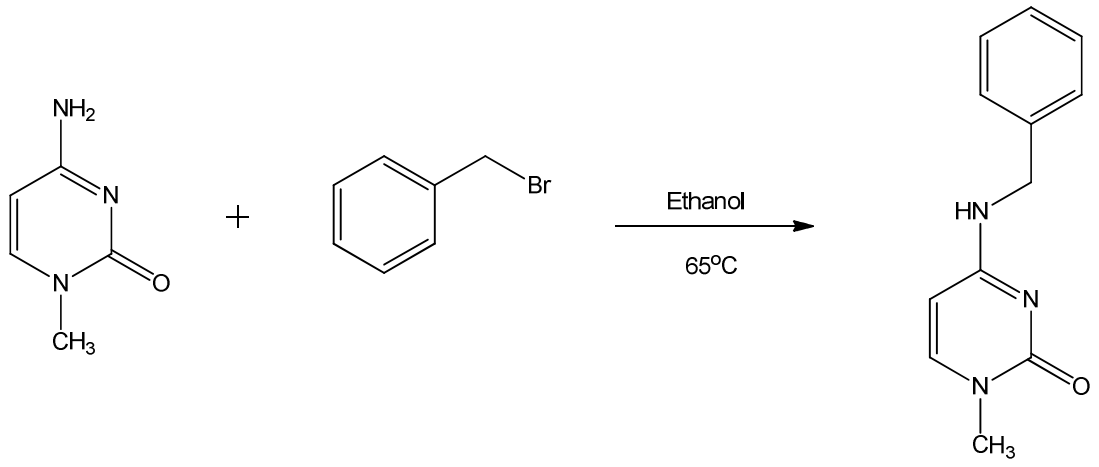
APCI-ESI MS: MH^+: 297.9007 (calc. 297.9008)



Scheme 4.15: Bromination of 3,6-dimethylisocytosine using bromine.

To 100 mg (0.00072 mol) of 3,6-dimethylisocytosine in 10 mL carbon tetrachloride in a 25 mL round bottom flask was added 0.25 mL (0.0049 mol) bromine. The resulting solution was stirred at room temperature for 30 minutes. After stirring the reaction was then quenched with 20 mL of DI H_2O. The organic layer was separated from the aqueous layer using a separatory funnel and dried with anhydrous sodium sulfate. The solution was decanted into a 50 mL round bottom flask and concentrated under reduced pressure, leaving behind an oily residue on the flask. An exact mass was not obtainable, but the mass spectrum of the residue showed the desired product.

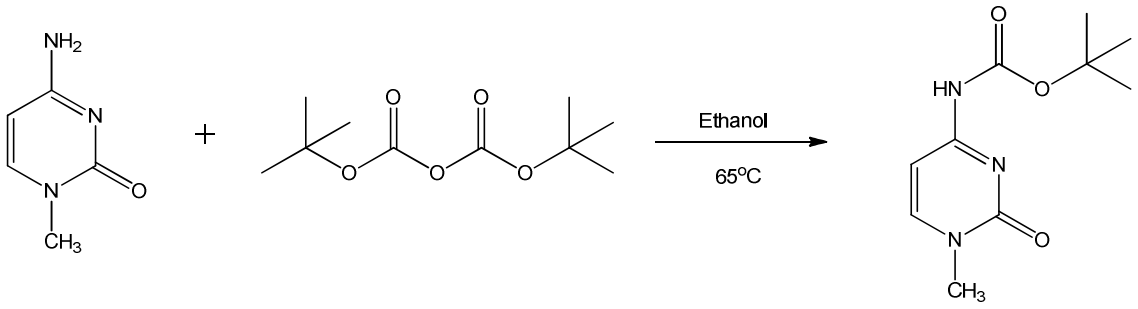
APCI-ESI MS: MH^+: 297.9015 (calc. 297.9008)



Scheme 4.16: Synthesis of N-benzyl-1-methylcytosine.

To 100 mg (0.0008 mol) of 1-methylcytosine in 20 mL absolute ethanol in a 50 mL round bottom flask was added 100 μL (0.0008 mol) benzyl bromide. The resulting solution was heated and stirred 65°C for 1.5 hours with an air reflux condenser attached. After stirring the reaction was cooled to room temperature and placed in an ice bath. White precipitate formed and was vacuum filtered using a Hirsch funnel. A crude yield of 60 mg (34.8%) was obtained. APCI-ESI MS showed the desired product, along with 1-methylcytosine. ^{1}H NMR showed the crude product consists of primarily 1-methycytosine starting material.

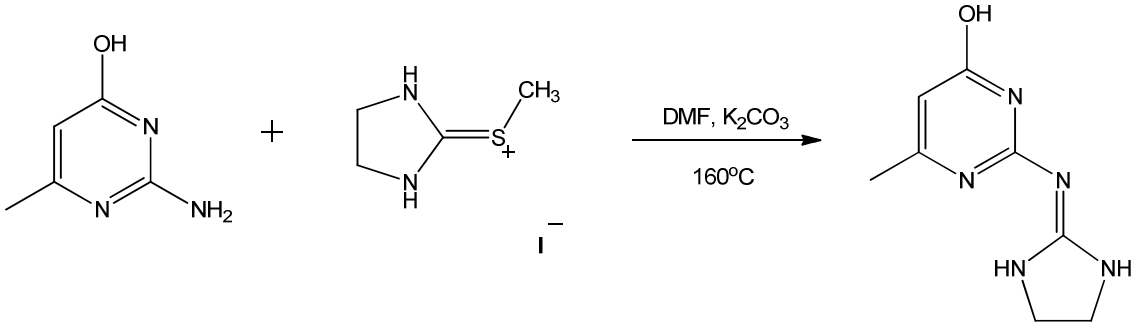
APCI-ESI MS: MH^+: 216.1123 (calc. 216.1131)



Scheme 4.17: Synthesis of N-BOC-1-methylcytosine.

To 100 mg (0.0008 mol) of 1-methylcytosine in 10 mL absolute ethanol in a 25 mL round bottom flask was added 202 μL (0.00088 mol) BOC anhydride. The resulting solution was heated and stirred 65°C for 12 hours with air reflux condenser attached. After stirring, the reaction is cooled to room temperature. The solution was evaporated to dryness under reduced pressure. ^{1}H NMR of the crude sample showed the composition of 1-methycytosine. A MS of the crude revealed only a small peak corresponding to the target product.

APCI-ESI MS: MH^+: 226.1186 (calc. 226.1186)



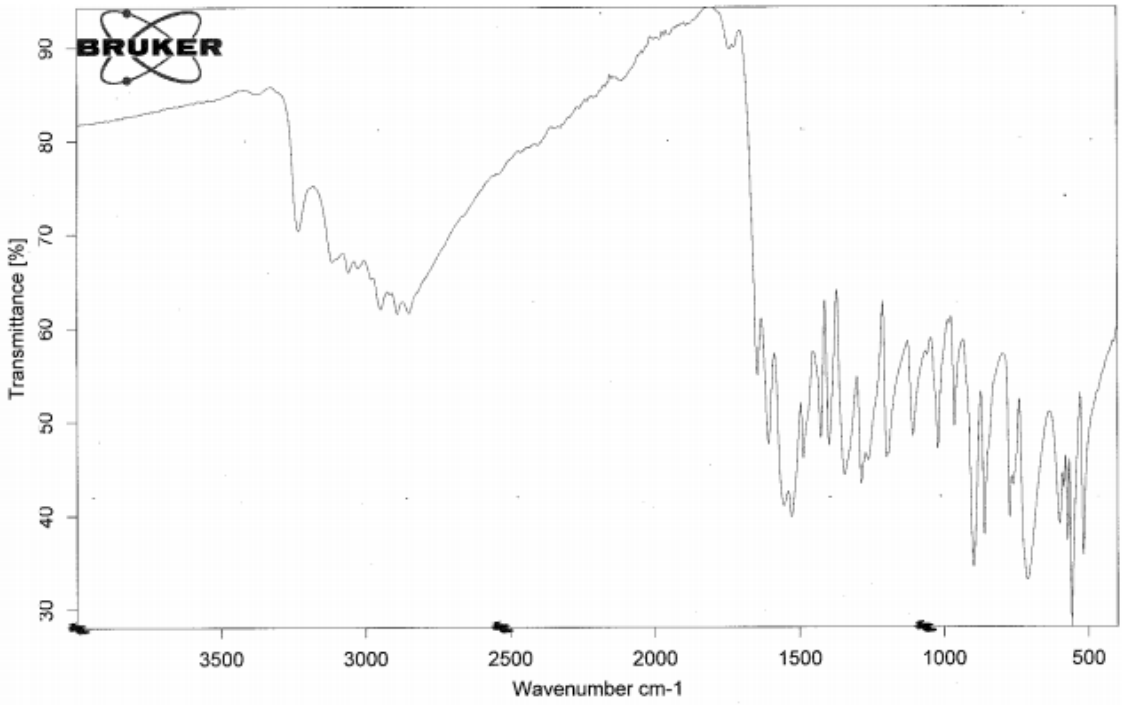
Scheme 4.18: Synthesis of N-ethyleneguanidine-6-methylisocytosine.

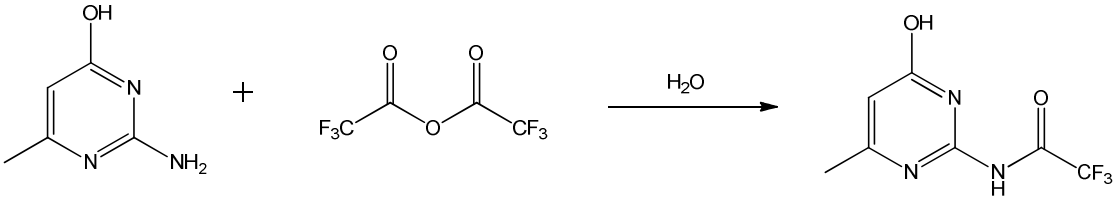
To 500 mg (0.004 mol) of 1-methylcytosine in 5 mL DMF in a 25 mL round bottom flask was added 950 mg (0.004 mol) methylated ethylene thiourea

iodide salt, and 550 (0.004) potassium carbonate. The resulting solution was heated and stirred 160°C for 12 hours with a water reflux condenser attached. After heating and stirring the reaction was cooled to room temperature. The white precipitate is filtered under vacuum and recrystallized using hot absolute ethanol. A yield of 421 mg (54%) was obtained.

APCI-ESI MS: MH^+: 194.1044 (calc. 194.1036)

^{1}H NMR (300 MHz, DMSO-*d*$_6$): δ 2.07 (3H s) δ 3.31 (2H s) δ 3.52 (2H s) δ 5.50 (1H s) δ 8.23 (2H s) δ 11.00 (1H s)



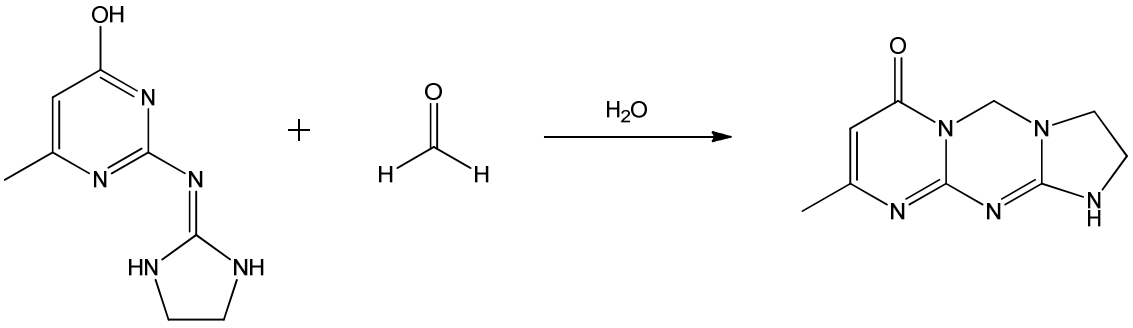


Scheme 4.19: Synthesis of N-trifluoroacetyl-6-methylisocytosine.

To 50 mg (0.0004 mol) of 1-methylcytosine in 3 mL DI H_2O in a 10 mL round bottom flask was added 113 μL (0.008 mol) trifluoroacetic anhydride, dropwise. The resulting solution was stirred at room temperature 12 hours. The solution is concentrated under reduced pressure. A yield of 60 mg (68%) was obtained.

^{1}H NMR (300 MHz, DMSO-*d*$_6$): δ 2.15 (3H s) δ 5.75 (1H s) δ 8.55 (1H b) δ 12.95 (1H b)

^{19}F NMR (300 MHz, DMSO-*d*$_6$): δ -74.71

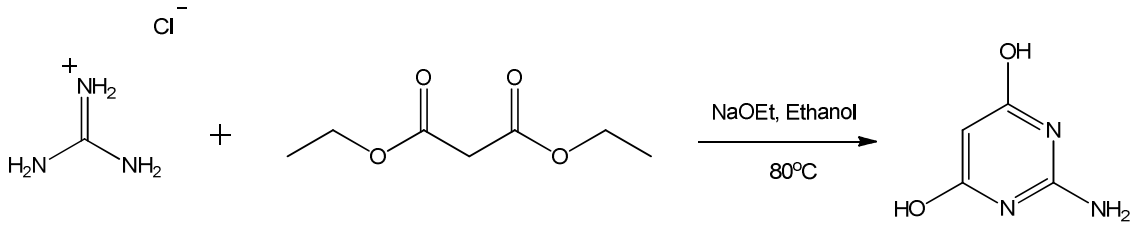


Scheme 4.20: Synthesis of 9-methyl-2,3-dihydro-1H-imidazo[1,2]pyrimido[2,1][1,3,5]triazin-7(5H)-one.

To 50 mg (0.00026 mol) of N-guanylated ethylene-6-methylisocytosine in 10 mL methanol in a 25 mL round bottom flask was added 2 ml formaldehyde (37%). The resulting solution was stirred at room temperature 24 hours. The solution was concentrated under reduced pressure. A MS of the crude product showed N-guanylated ethylene-6-methylisocytosine along with the desired

product. The intensity of N-guanylated ethylene-6-methylisocytosine was greater than the product. A yield was not obtained.

APCI-ESI MS: MH^+: 206.1041 (calc. 206.1036)



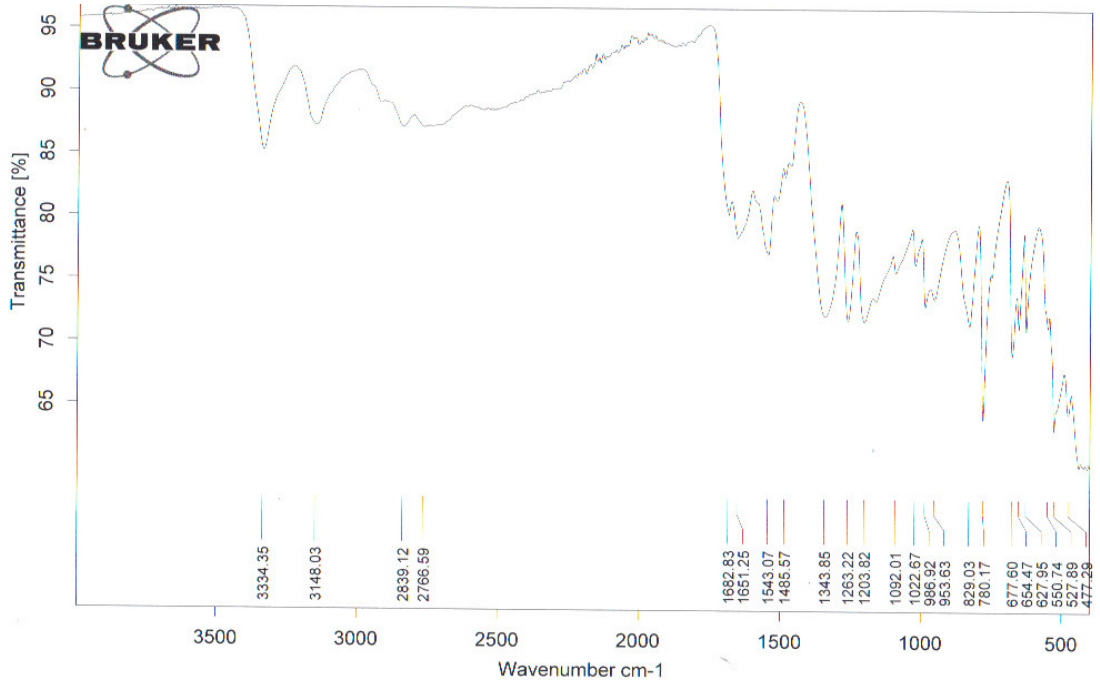
Scheme 4.21[25]: Synthesis of 6-hydroxyisocytosine.

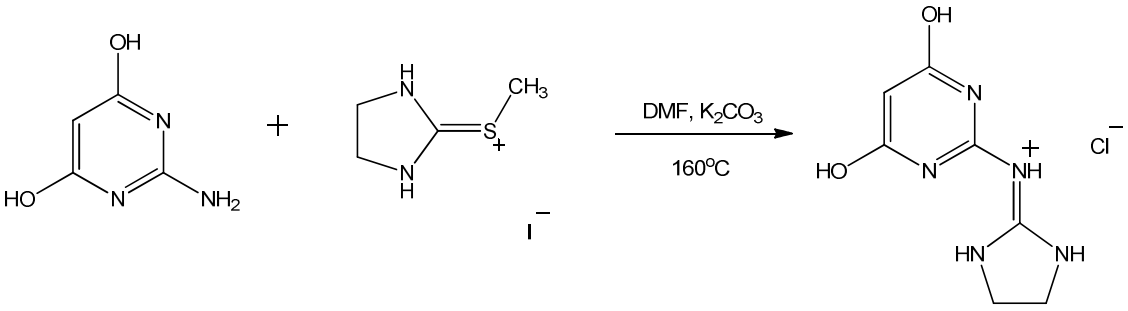
The synthesis of 6-hydroxyisocytosine followed a previously published procedure[25]. To 577 mg (0.025 mol) sodium in 30 mL ethanol, in a 50 mL round bottom flask with a stir bar was added 2.4 grams (0.025) guanidinium HCl and 3.8 mL diethyl malonate (0.025 mol). The resulting solution was heated and stirred at 80°C for 3 hours. The solution was cooled to room temperature after heating and hot DI H_2O is added until solids dissolve. Upon cooling to room temperature, glacial acetic acid was added until solution is cloudy. The round bottom flask was placed in an ice bath for 10 minutes and the solids filtered under vacuum. A yield of 2.21 grams (69%) was obtained.

APCI-ESI MS: MH^-: 126.0312 (calc. 126.0309)

^{1}H NMR (400 MHz, D_2O): δ 4.14 (1H s) δ 7.58 (2H b) δ 11.12 (2H b)

^{13}C NMR (400 MHz, D_2O): δ 78.63 (1C) δ 155.03 (1C) δ 167.33 (1C)





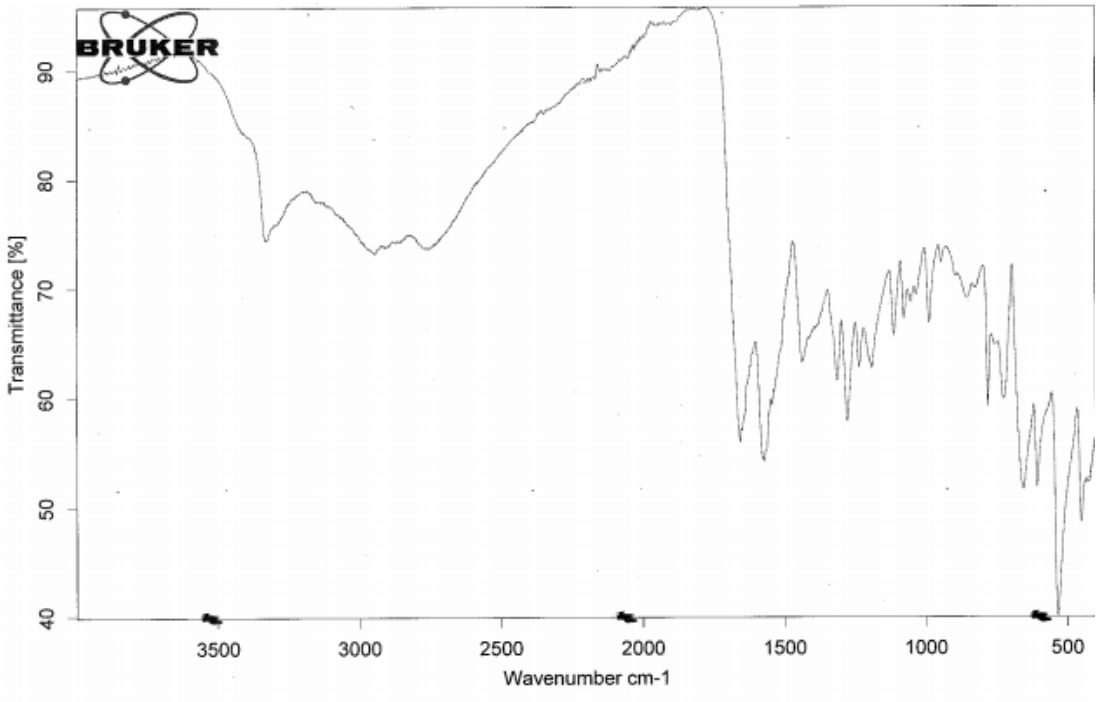
Scheme 4.22: Synthesis of protonated N-ethyleneguanidine-6-hydroxyisocytosine chloride salt.

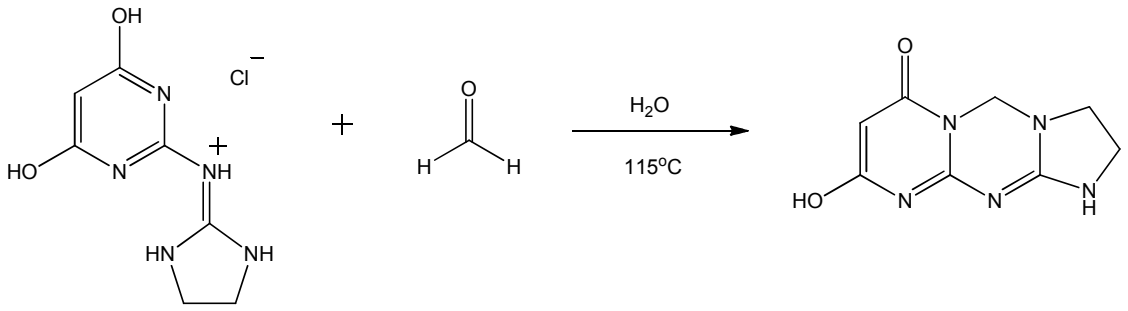
To 866 mg (0.0068 mol) of 6-hydroxyisocytosine in 7 mL DMF in a 25 mL round bottom flask was added 1.66 gram (0.0068 mol) methylated ethylene thiourea iodide salt and 900 (0.0068) potassium carbonate. The resulting solution

was heated and stirred 160^{o}C for 12 hours**,** with a water reflux condenser attached. After heating and stirring, the reaction was cooled to room temperature. Hot DI H_2O was added until solids dissolve. Concentrated HCl is added until solution is acidic and the solid filtered. A yield of 411 mg (26%) was obtained.

APCI-ESI MS: MH^+: 196.0833 (calc. 196.0829)

^{1}H NMR (300 MHz, DMSO-*d_6*): δ 3.65 (4H s) δ 8.04 (2H b) δ 8.81 (1H s) δ 11.67 (2H b)

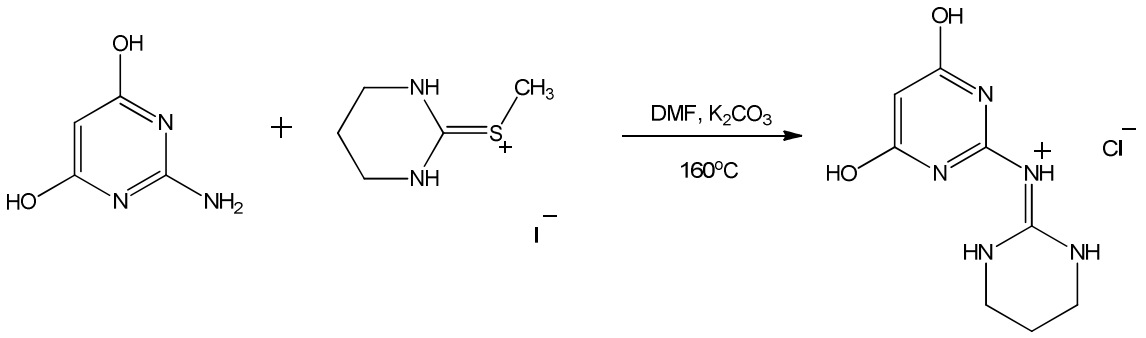




Scheme 4.23: Synthesis of 9-hydroxy-2,3-dihydro-1H-imidazo[1,2]pyrimido[2,1][1,3,5]triazin-7(5H)-one

To 378 mg (0.00016 mol) of protonated guanylated ethylene-6-hydroxyisocytosine in 10 mL DI H_2O in a 25 mL round bottom flask was added 2 mL formaldehyde (37%). The resulting solution was heated and stirred at 115°C 24 hours. The solution was cooled to room temperature and the solid filtered under vacuum. A MS of the crude product showed guanylated ethylene-6-hydroxyisocytosine along with the desired product. The intensity of guanylated ethylene-6-hydroxyisocytosine was more intense compared to the product. A yield was not obtained.

APCI-ESI MS: MH^+: 208.0834 (calc. 208.0829)



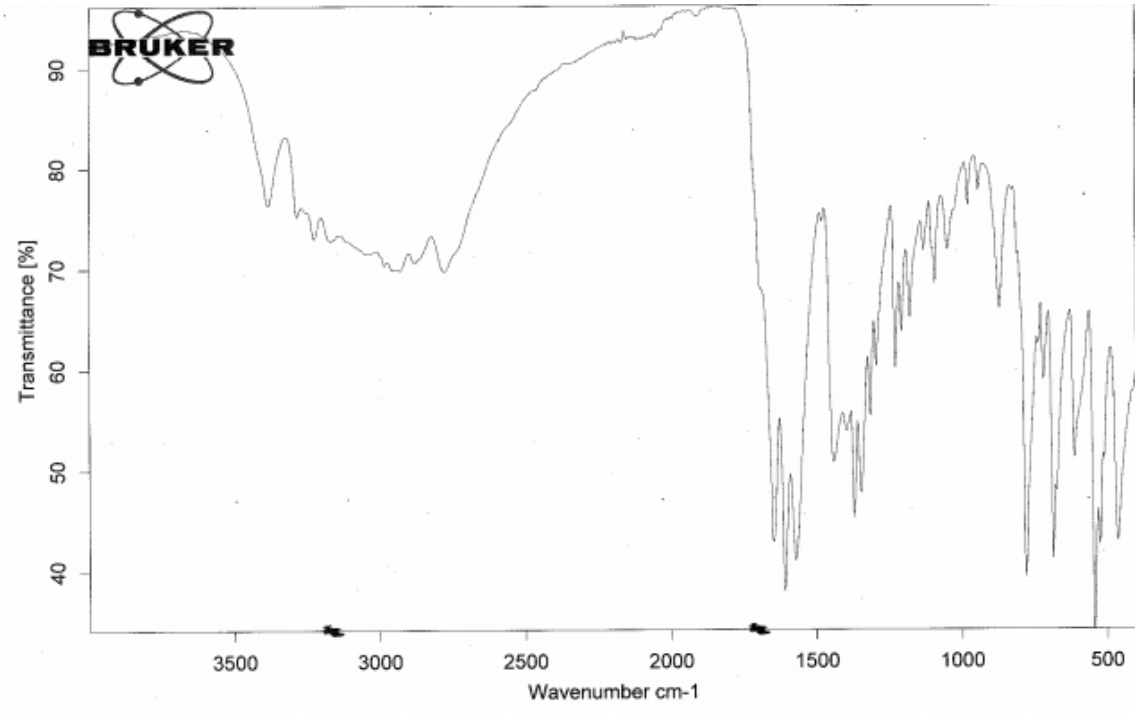
Scheme 4.24: Synthesis of protonated N-propyleneneguanidine-6-hydroxyisocytosine chloride salt.

To 850 mg (0.0066 mol) of 6-hydroxycytosine in 7 mL DMF in a 25 mL round bottom flask was added 1.72 grams (0.0066 mol) methylated propylene thiourea iodide salt, and 900 mg (0.0066) potassium carbonate. The resulting solution was heated and stirred 160^{o}C for 12 hours, with a water reflux condenser attached. After heating and stirring, the reaction was cooled to room temperature. Hot DI H_2O is added until solids dissolve. Concentrated hydrochloric acid was added until solution is acidic and the solid filtered. A yield of 550 mg (34%) was obtained.

APCI-ESI MS: MH^+: 210.0987 (calc. 210.0986)

^{1}H NMR (400 MHz, DMSO-*d*$_6$): δ 1.81 – 1.86 (2H p) δ 3.37 – 3.40 (4H p) δ 8.23 (2H b) δ 9.82 (1H s) δ 11.83 (2H b)

^{13}C NMR (400 MHz, DMSO-*d*$_6$): δ 18.49 (1C) δ 37.91 (2C) δ 79.32 (1C) δ 150.84 (1C) δ 157.48 (1C) δ 162.50 (2C)



Results

5-Hydroxy-1-methylcytosine:

5-Hydroxy and 5-hydroxymethyl cytosine were synthesized to determine if their proton-bound heterodimer would be more or less tightly bound compared to 1-methylcytosine proton-bound homodimer. 5-hydroxycytosine was synthesized, but methylation at the 1-position proved to be difficult. Methylation using the same method as before for cytosine did not yield the desired product. Refluxing 5-hydroxycytosine in methyl iodide also proved unsuccessful. Surprisingly, neither the hydroxyl group at the 5-position nor amino group underwent nucleophilic attack with methyl iodide. 5-Hydroxy-1-methylcytosine was never successfully isolate, nor observed by mass spectrometry.

Tricyclic Cytosine and Isocytosine

The target molecules are tricyclic compound consisting of either cytosine or isocytosine base core. This tricyclic cytosine heterocycle is designed in a manner that upon the acquisition of a proton, it forms a D-D-A-A hydrogen bonding motif. This bonding motif is unique in the sense that it is capable of forming a three-point hydrogen bond with either cytosine or protonated cytosine (**Figure 4.5**). Protonated targets 1, 2, and 3 not only have the capability to form a proton-bound dimer with neutral or protonated cytosine, they can also form a proton-bound homodimer with their corresponding neutral counterpart.

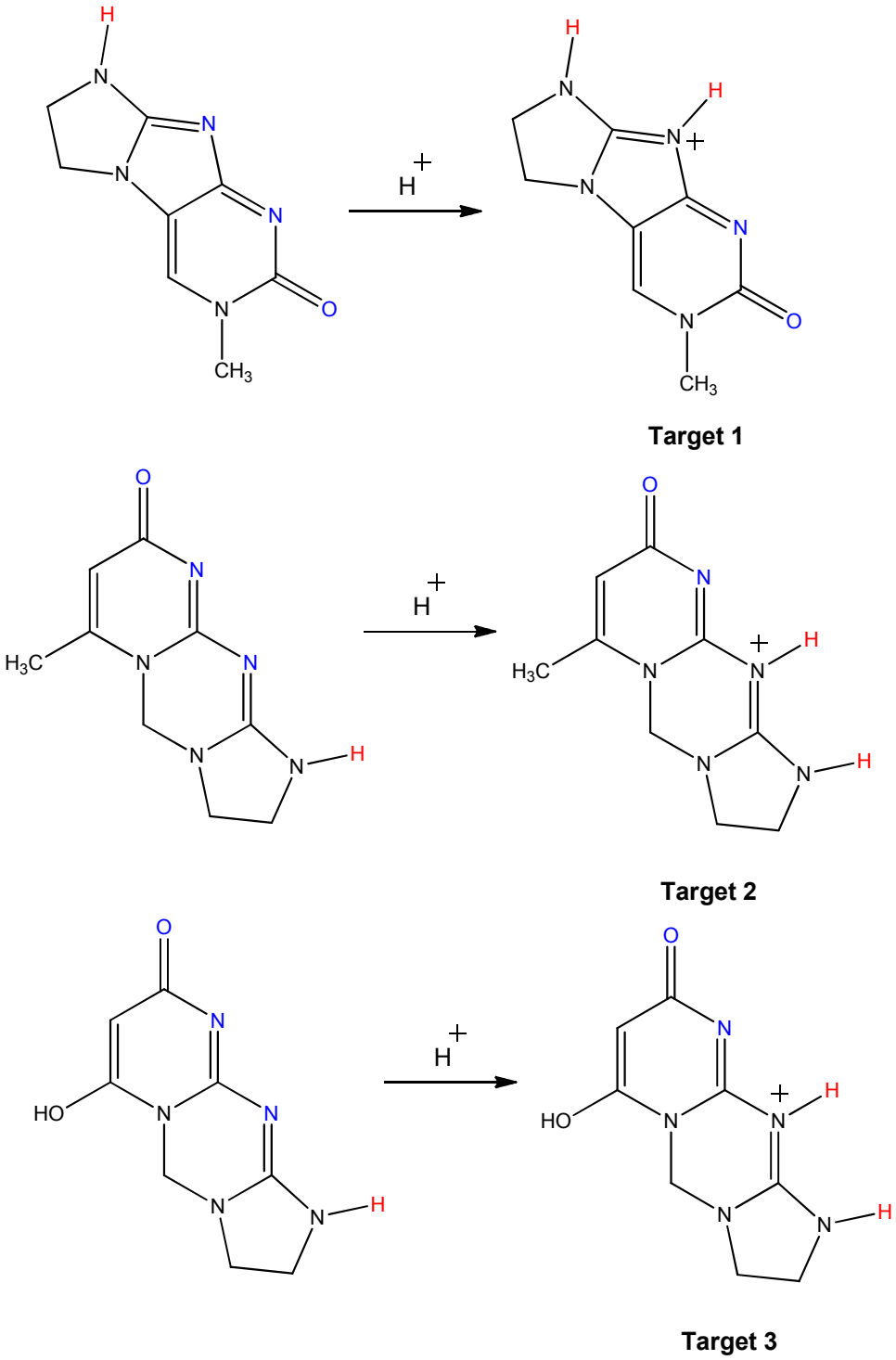
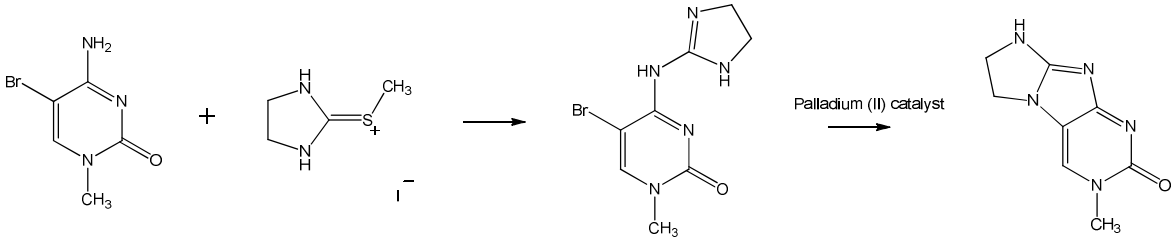


Figure 4.5: Targets **1**,**2**, and **3** are all capable of forming hydrogen bonding D-D-A-A hydrogen bonding motif upon protonation.

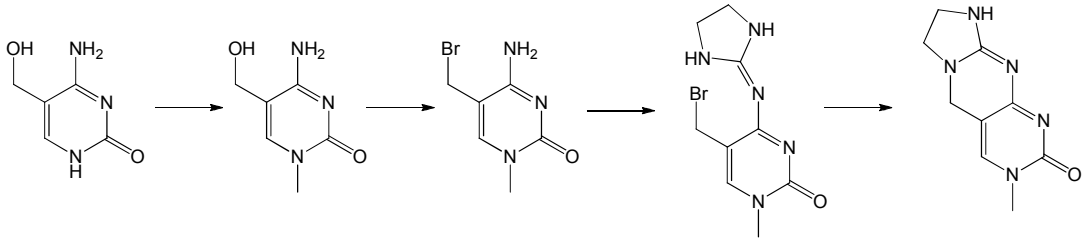
The envisioned synthesis of **Target 1** began with methylation of 5-bromocytosine using the previously described method[23]. After obtaining 1-methyl-5-bromocytosine, the next step was to react it with methylated ethylene thiourea, resulting in a guanylated 5-bromo-1-methylcytosine. A palladium catalyzed Buchwald-Hartwig cross coupling reaction would be utilized afterward to produce **Target 1** and close the middle ring **(Scheme 4.24)**.



Scheme 4.25: Proposed synthesis of **Target 1**.

1-Methyl-5-bromocytosine was synthesized following the published procedure[20], but the addition of the cyclic guanidine group turned out to be more difficult than initially anticipated. 1-Methyl-5-bromo-guanylated cytosine was detected using ESI-MS, but was never isolated. The reaction also produced other unidentified side products. The reaction did not proceed to completion as both starting materials were present in the ESI-MS. In attempt to increase the yield and efficiency of the reaction, different solvents and reaction condition were performed using 1-methylcytosine in place of 1-methyl-5-bromo-cytosine due to the cost of 5-bromocytosine.

Although 5-bromocytosine can be purchased, the cost of it was mounting, especially when it is the main component. To save on costs, different functional groups were envisioned to replace bromine. The purpose of positioning a halogen at the 5-postion is to perform a Buchwald-Hartwig palladium catalyze coupling, but the same result can be achieved if an iodide was at the 5-position. The treatment of 1 equivalent of cytosine with 2 equivalent of chloramine-T and 1 equivalent of sodium iodide produced 5-iodocytosine as described in **Scheme 4.3**. Methylation at the 1-position of 5-iodocytosine proceeded without complication. Addition of the cyclic guanidine group did not work as anticipated. It is possible the electron withdrawing nature of the halogen at the 5-position renders the amino group non-nucleophilic.

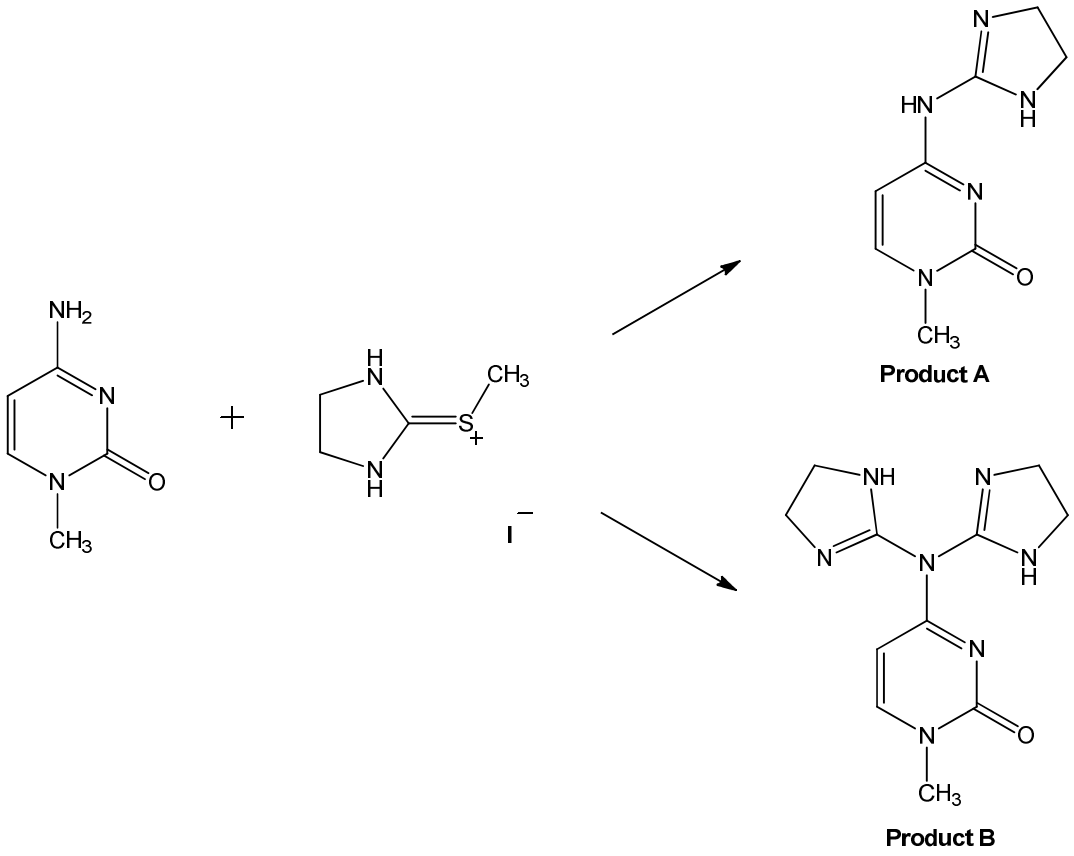


(11)

Another functional group that was thought to be a suitable replacement for bromine at the 5-position was a hydroxyl methyl group. The treatment of 5-hydroxymethylcytosine with phosphorus tribromide would result in 5-bromomethylcytosine. It was though that the position of the bromine on a sp^3 hybridized carbon would allow for spontaneous cyclization after the addition of the cyclic guanidine group (**Eq 11**). The synthesis of 5-hydroxymethylcytosine is outlined in **Scheme 4.1**. 5-Hydroxymethylcytosine was observed by MS, but not

isolable. Methylation at the 1-position of 5-hydroxymethylcytosine did not proceed, like that of 5-hydroxycytosine.

Scheme 4.26 shows the reaction of 1-methylcytosine with methylated ethylene thiourea iodide salt. The reaction attaches both one and two guanidine groups to 1-methylcytosine, with the addition of two guanidine groups being the major product. ESI-MS shows the presence of both **Product A** and **Product B**. **Product A** were not isolable, whereas **Product B** was.



Scheme 4.26: Reaction of 1-methylcytosine with methylated ethylene thiourea iodide salt results in two products (A and B).

After isolation of **Product B**, a 225 μM solution was made and added to a solution containing DNA oligomers to see any adduct formation. **Table 4.1** outlines the DNA sequences chosen for investigation and **Figure 4.6** reproduces the ESI-MS of a solution containing the symmetrical loop region DNA sequence without (top) and with (bottom) the addition of **Product B**.

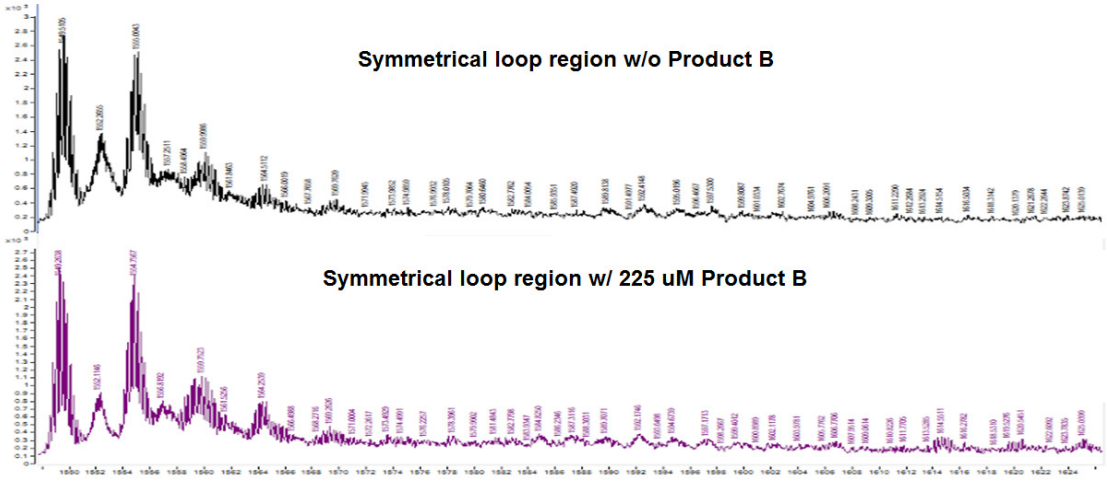


Figure 4.6: ESI-MS of symmetrical loop region DNA sequence without **Product B** (top) and with 225 μM solution of **Product B**.

The addition of **Product B** didn't show any immediate adduct with the symmetrical loop region DNA sequence by ESI-MS. When the solution was left idle for a period of several months, DNA-**Product B** adduct was observed.

Table 4.1: DNA sequences of interest along with their base codes.

DNA Sequences	Base Code
Symmetrical loop region	5'-(CCCTAACCCTAACCCTAACCC)-3'
Sequence from HIF-1α gene	5'-(TCCCGCCCCCTCTCCCCTCCCC)-3'
VEGF promoter region	5'-(ACCCCGCCCCCGGCCCGCCCCG-)3'
VEGF cis-regulatory region	5'-(CCCCGCCCCTGACCCGGGTACCC)-3'
c-KIT promoter region	5'-(CCCTCCTCCCAGCGCCCTCCCT)-3'
K-Ras promoter region	5'-(CCTCCCCCTCTTCCCTCTTCCCACACCGCCCT)-3'

The next experiment probed whether **Product B** had any effect on the melting point of DNA oligomers. To determine if the compounds have an effect on the DNA melting points, the absorbance of a wavelength corresponding to the DNA sequence or compound is tracked as a function of temperature. The term DNA melting does not correspond to a phase change, but rather a conformational change from its random coil into a stable secondary structure. The melting point is at the equivalence point.

The addition of **Product B** with the symmetrical loop region sequence and HIF-1α, from **Table 4.1**, showed an increase in the DNA's melting point. **Figure 4.7** reproduces the melting point plots of the symmetrical loop region and HIF-1α oligomer sequence with and without the presence of **Product B**. The red and blue curves depict the melting point of the symmetrical loop region and HIF-1α at 32 and 54^{o}C, respectively. When **Product B** is introduced, the melting point curves of both sequences increased by 10 and 4^{o}C, respectively.

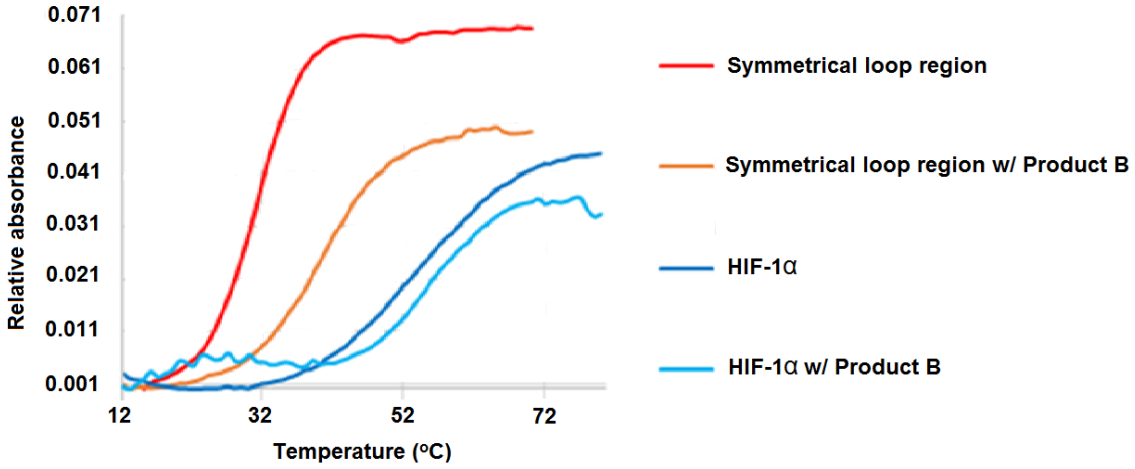
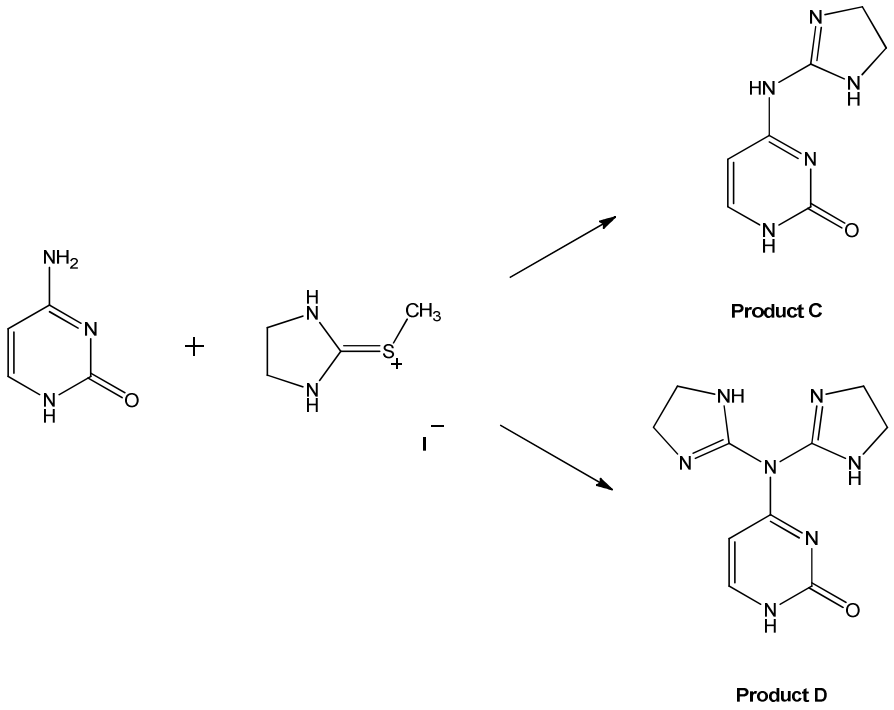


Figure 4.7: DNA melting points of the symmetrical loop region and HIF-1α with and without **Product B**.

Since **Product B** increased the melting points of several DNA oligomers, another analog of **Product B** was synthesized and investigated. This analog is not methylated at the 1-position of cytosine, and its reaction scheme is outlined in **Scheme 4.27**.



Scheme 4.27: Reaction of cytosine with methylated ethylene thiourea iodide salt.

Similar to that of **Scheme 4.25**, two products are possible but only **Product D** was observed by ESI-MS and isolated. **Product D** was also found to increase the melting point of the symmetrical loop region. **Figure 4.8** reproduces the melting point curves of the symmetrical loop region with the addition of **Product D**. The original melting point diagram of the symmetrical loop region by itself, from **Figure 4.7,** shows a conformation change of the DNA sequence around 32^{o}C. The addition of **Product D** with the symmetrical loop region sequence increased the melting point to 43^{o}C, an increase of 11^{o}C. Another analog of bis-guanylated cytosine is being developed has an octyl group at the 1-position. The synthesis of this compound proved to be more challenging than **Product B** or **D**, and efforts are still being made to isolate this compound.

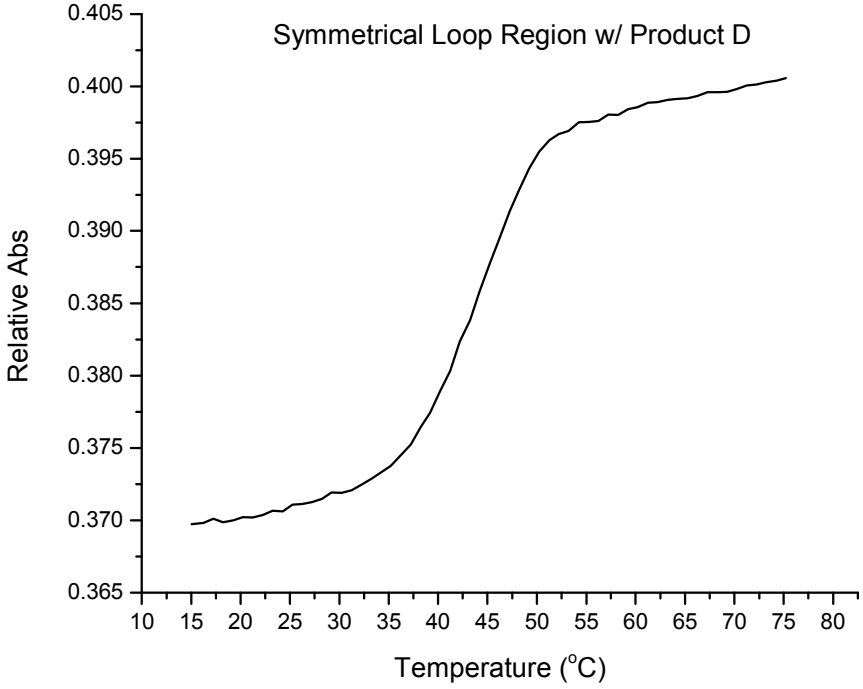
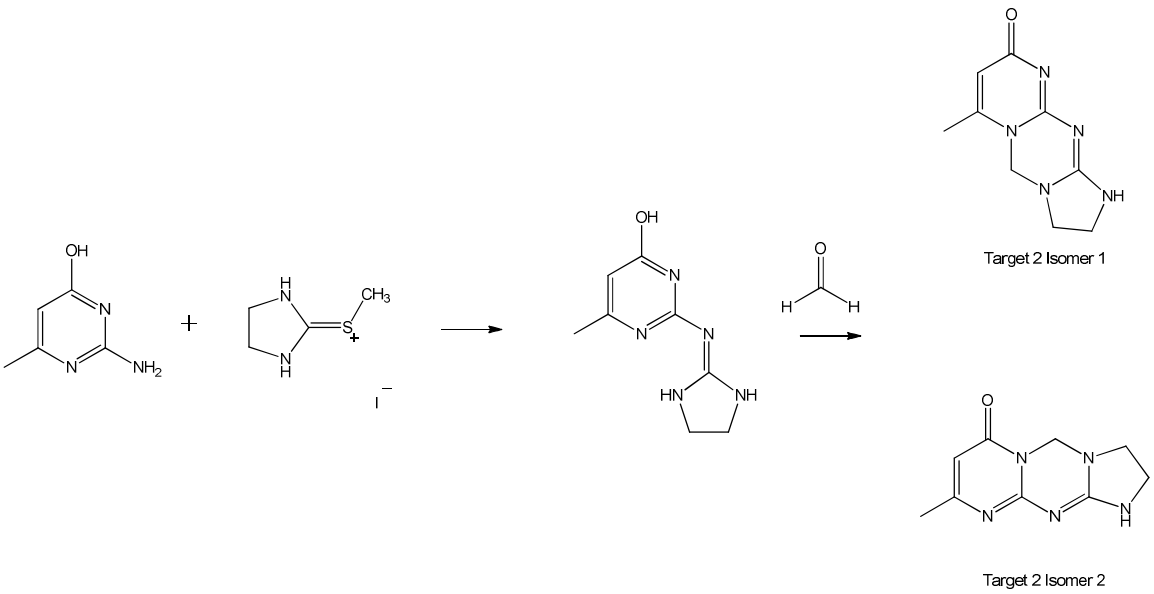


Figure 4.8: DNA melting point of the symmetrical loop region with **Product D**.

Even though **Product B** and **Product D** cannot form a three-point hydrogen bond with cytosine (due to the placement of the two guanidine groups), it does somehow affect the stability of the *i*-motif. One hypothesis of its interaction with the *i*-motif is a base stacking interaction. It is possible for the cytosine base core of **Product B** and **D** to base stack on the bases of the loop region.

Target 2:

The synthesis of **Target 2** is outlined in **Scheme 4.28**. The reaction between commercially available 6-methylisocytosine with methylated ethylene thiourea iodide salt to produce N-ethyleneguanidine-6-methylisocytosine is described in the experimental section. Pumping gaseous formaldehyde (by heating paraformaldehyde) using nitrogen as a carrier gas into a solution of N-ethyleneguanidine-6-methylisocytosine should close the middle ring.



Scheme 4.28: Proposed synthesis of **Target 2**.

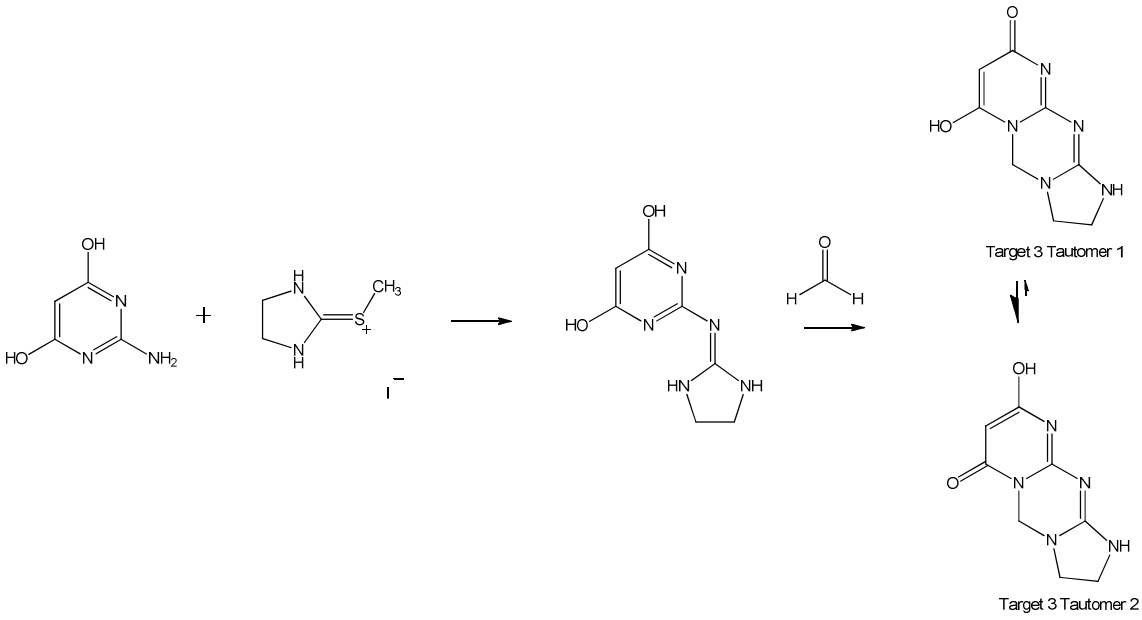
The reaction of N-ethyleneguanidine-6-methylisocytosine with gaseous formaldehyde produces isomers of **Target 2**. **Target 2** Isomer 1 is capable of forming a A-A-D-D hydrogen bonding motif upon protonation while **Target 2** Isomer 2 can form a A-D-D motif. Since a complementary protonated Target 2 Isomer 1 has the same hydrogen bonding motif, it can form a diprotonated PBD

with itself. This event can be observed using mass spectrometry. A diprotonated dication would have its ^{13}C natural abundance peak separated by ½ amu. A crude ESI-MS of the reaction between N-ethyleneguanidine-6-methylisocytosine and gaseous formaldehyde was obtained and a peak corresponding to protonated **Target 2** was observed. Its ^{13}C natural abundance peak was also observed but separated by 1 amu. This led us to conclude that the major product of the reaction is **Target 2** Isomer 2. To overcome this obstacle, we utilized a molecule with a pane of symmetry so that cyclization does not produce two isomers.

Target 3:

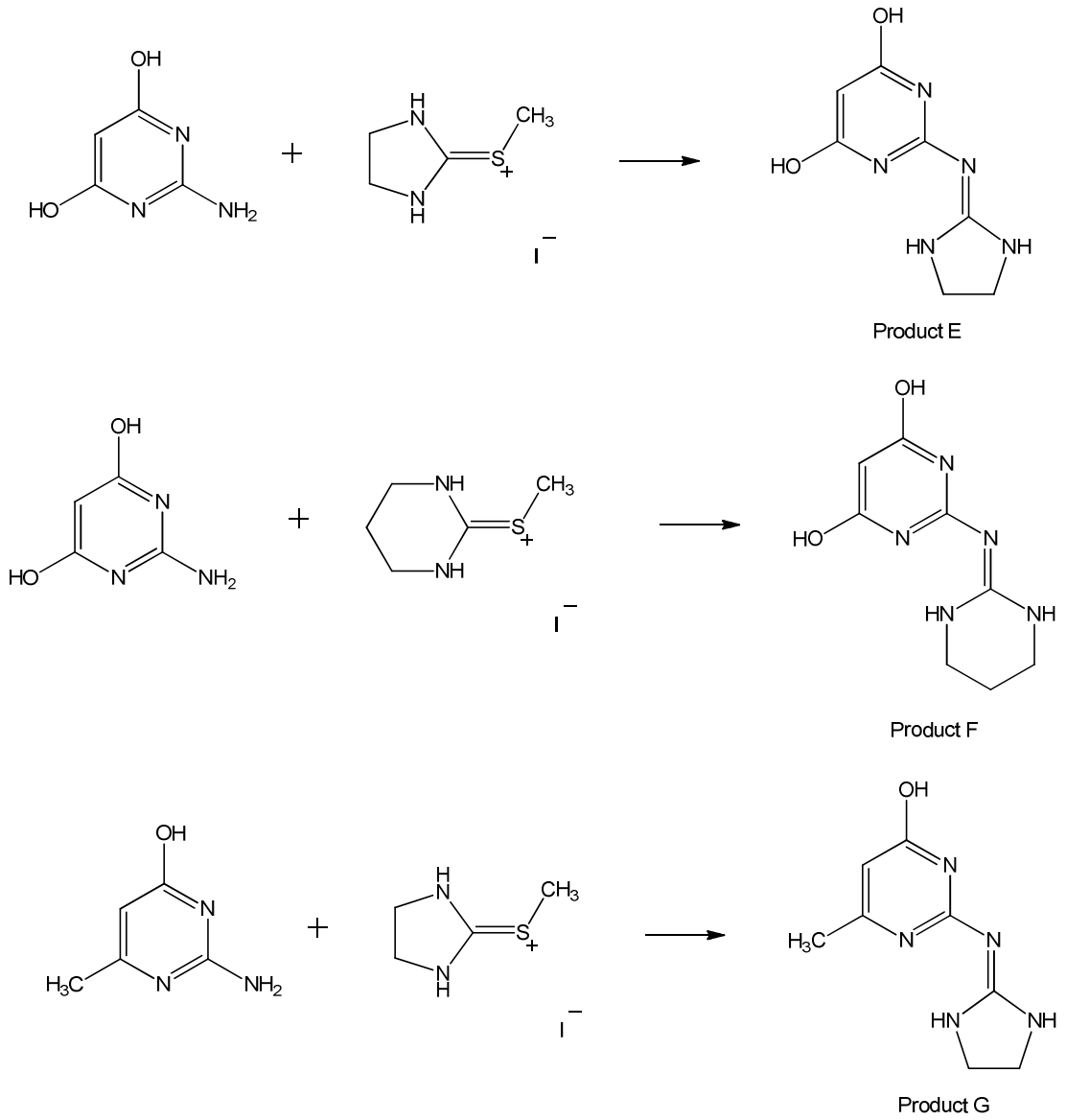
The synthesis of **Target 3** was envisioned to take the place of **Target 2** since the synthesis of **Target 2** proved to be more difficult than initially thought. Like **Target 1**, protonated **Target 3** is still able to form a proton-bound dimer with both neutral and protonated cytosine. **Scheme 4.29** outlines the synthetic pathway to produce **Target 3**. **Target 3** exists as a single isomer, but it can tautomerize between two stable tautomers. DFT calculations at B3LYP/6-31G** level predicts tautomer 2 to be more stable than tautomer 1 with a ΔH = 5.3 kcal/mol. One explanation for tautomer 2 possessing a higher stability compared to tautomer 1 is its conjugation with the carbonyl group. Since tautomer 2 is the more stable of the two, a proton-bound homodimer might not be possible

depending on the orientation of the hydroxyl group. An ESI-MS spectrum of **Target 3** did not show proton-bond homodimer formation.



Scheme 4.29: Proposed synthesis of **Target 3**.

The road to **Target 2** and **Target 3** branched off into three other isolatable products. **Scheme 4.30** outlines the synthesis of **Product E, F,** and **G**. Reaction of 6-hydroxyisocytosine or 6-methylisocytosine with methylated ethylene thiourea iodide and methylated propylene thiourea iodide salts produce the intermediate of which cyclization of the middle rings produce **Target 2**.



Scheme 4.30: Synthesis of **Product E**, **F**, and **G**.

Unlike **Product B**, the ESI-MS of **Product G** shows binding to the symmetrical loop region DNA sequence (**Figure 4.9)**. The deprotonated ion of the symmetrical loop region DNA sequence is observed around *m/z* 1549 and its sodiated ion around *m/z* 1554. Binding of **Product G** to the DNA sequence is

observed near *m/z* 1597 and its sodiated ion around *m/z* 1603. Since there are more hydrogen bond donating and accepting sites available for **Product G**, it is hypothesized that the binding occurs with either the loop region of the DNA sequence. DNA melting point experiments with **Product G** does not indicate an increase of stability with any of the DNA sequences from **Table 4.1**. **Product E** and **F** have not been tested to determine if they for a complex with any DNA sequences.

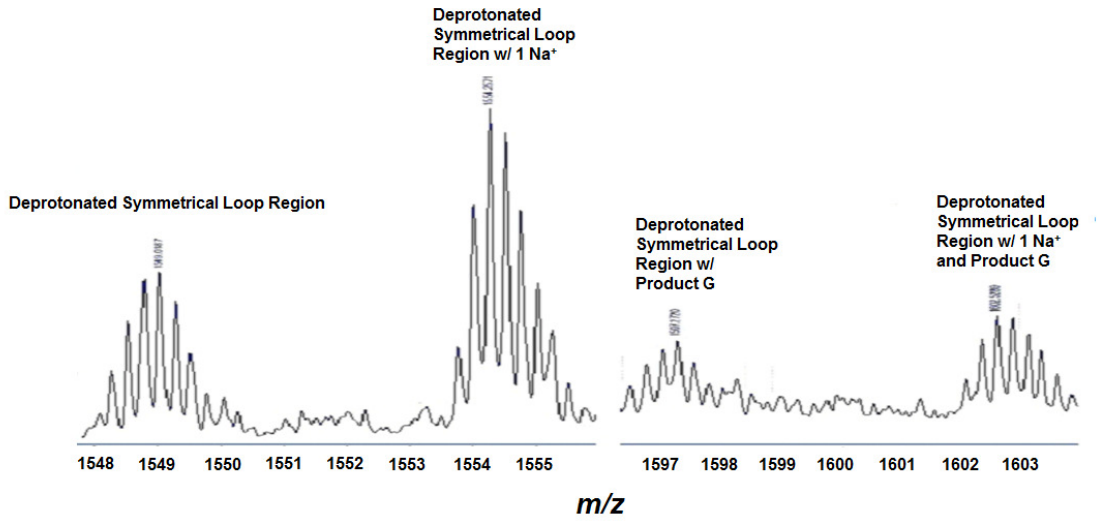


Figure 4.9: ESI-MS of the symmetrical loop region DNA sequence with **Product G**.

It is possible **Product G** is binding to a cytosine or protonated cytosine of the *i*-motif core. One way to investigate this possibility to conduct solution phase NMR experiments of the DNA sequence with **Product G**. The bridging proton of the *i*-motif has characteristic proton shifts between 15 – 16 ppm, and the number of proton resonance represents the number of bridging protons. A disappearance of

a band with the introduction of **Product G** would mean that proton-bound dimer is separated.

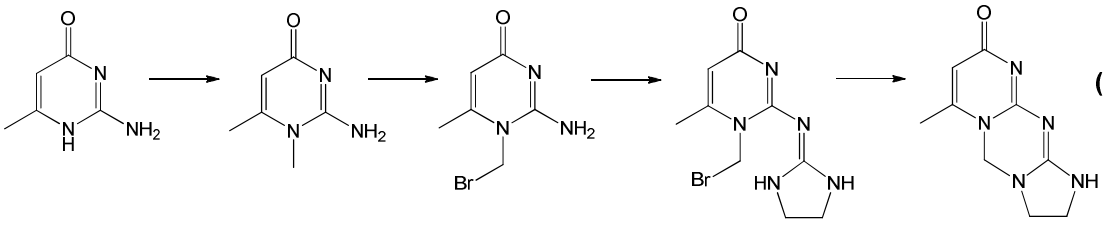
AVO

An undergraduate student from Prof. Michael Marsella's lab, Aiden Aceves, performed computational studies on **Products B, D** and **E** using a software their lab developed called AVO. AVO simulates a molecule's behavior in a biological environment and measures how the molecule responds to each situation. Using this program, **Product E** showed favorable binding to cell division protein kinase 4 (CDK4) and received a score of 2.112, comparable to that of Palbociclib which received a score of 2.173 and was approved by the FDA in 2015 to treat breast cancer[26].

Products B and **D** did not show any promising results from AVO even though they influence the melting points of several DNA strands by UV melting. We decided to introduce aqueous solutions of **Products B** and **D** (at various concentrations ranging from 1 nM – 100 μM) to cancerous breast cells. The cells are incubated with and without **Product B** and **D** for 4 days at physiological temperature and in 5% CO_2 atmosphere to maintain pH balance. A MTS assay is used to determine the number of live cells remaining in each culture. Eric Commendatore is performing these experiments in Prof. Ameae Walker's laboratory.

Ongoing Work

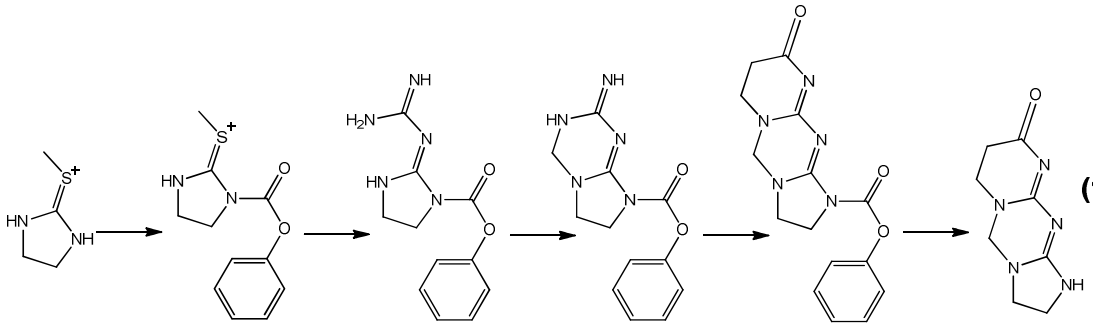
Other synthetic pathways were explored to achieve **Targets 1, 2, and 3** but proven unsuccessful. **Eq 12** depicts one of the envisioned pathways to **Target 2** was to brominate the methyl group at the 1-position of 1,6-dimethylisocytosine before the addition of the cyclic guanidine group. With the 1-position of 1,6-dimethylisocytosine brominated, one of the nitrogen on the cyclic guanidine might undergo spontaneous nucleophilic displacement of the bromine to form the middle ring. This pathway proved to be unsuccessful because methylation of 6-methylisocytosine using the same procedure as previously used for methylating cytosine at the 1-position[20] methylated the nitrogen at the 3-position rather than the 1-position. X-ray diffraction confirms the position of the methyl group at the 3-position. Bromination of the methyl group also proved to be challenging. It was difficult to control whether one of the methyl groups was going to be brominated or both. An ESI-MS of the crude reaction product showed mono and dibrominated product.



(12)

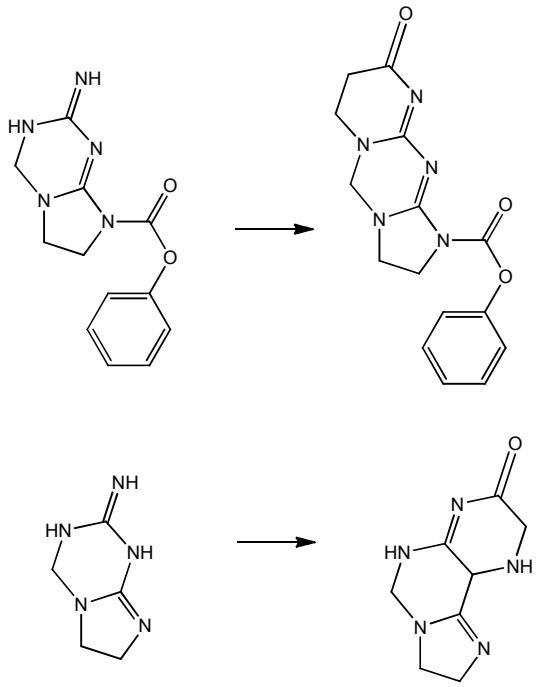
Other synthetic pathways involving the use of protecting groups were investigated but not yet completed. Since methylation prefers the nitrogen at the 3-position, positioning a protecting group at the 3-position should allow for modification at the 1-position with ease. Of all the protecting groups investigated,

none of which protected the nitrogen at the 3-position because doing so would break aromaticity. **Eq 13** outlines the synthetic pathway that makes use of benzoylchloroformate as the protecting group.



(13)

This synthetic pathway starts with the protection of methylated ethylene thiourea. The addition of guanidine free base is added next along with cyclization of the middle ring. A protecting group is necessary before the cyclization of the top ring using 3-bromopropionyl choride. Without the protecting group, the formation of two different products is possible (**Eq 14**). Having one side of the protonated ethylene thiourea protected allows for only one side of the guanidine group available for modification. Without the protecting group, 3-bromopropionyl chloride has the possibility of reacting with two different sites, resulting in a different tricyclic compound.



(14)

Other ongoing work with AVO and cell culture includes performing computation studies on **Targets 1**, **2**, or **3** once a pure product is obtained. Other additional tests include testing **Products B** and **D** on healthy breast cells to see if they cause healthy cells to undergo apoptosis.

Conclusions

In 2014, an imidazolinopyrimidinone was studied due to its ability to kill cancer cells while showing little toxicity against healthy cells[27]. This imidazolinopyrimidinone was assigned to be in a linear conformation. While investigating the linear conformation, it was discovered that the active structure was misassigned. Another constitutional isomer (angular) of this

imidazolinopyrimidinone was synthesized and it was discovered that the angular form exhibit anti-cancer activity, while the previously published form is inactive.

As shown in **Scheme 4.28**, the cyclization to achieve **Target 2** produces two isomers, similar to that of the previously investigated imidazolinopyrimidinones[27]. The use of 6-hydroxycytosine in place of 6-methylisocytosine was envisioned to overcome this obstacle. This project is still ongoing as other avenues are being investigated in attempt to successfully synthesize and characterize **Targets 1**, **2**, and **3**. **Product B** and **D** showed an increase of stability of several *i*-motif forming DNA sequences. These compounds were side products isolated from the synthesis of **Target 1**. **Product B** and **D** didn't show any binding to single stranded DNA by MS within a short time frame, but does show binding if the solution of DNA and **Product B** is left idle over a period of several months. The addition of **Product B** and **D** to several DNA sequences demonstrated additional stability of the DNA sequences by UV melting.

Contrary to **Product B** and **D**, **Product E** does binding to single stranded DNA by MS, but does not show an increase in stabilization of single-stranded DNA by UV melting. Using AVO, Product E showed promise in becoming an anti-breast cancer agent. Additional trial runs need to be performed on **Product B**, **D**, and **E** as to get an average number of cancerous cells these compounds kill. A variety of concentrations also need to be screen to determine the optimal concentration.

References

1. Centers for Disease Control and Prevention, accessed 7 September 2015, <http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>

2. American Cancer Society, accessed 7 September 2015, <http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/>

3. Mayo Clinic, accessed 7 September 2015, <http://www.mayoclinic.org/diseases-conditions/stomach-cancer/basics/risk-factors/con-20038197>

4. Mayo Clinic, accessed 7 September 2015, <http://www.mayoclinic.org/diseases-conditions/lung-cancer/basics/risk-factors/con-20025531>

5. J.L. Huppert, S. Balasubramanian "G-Quadruplexes in Promoters Throughout the Human Genome" *Nucleic Acids Res.*, **2007**, 35, 406 – 413.

6. W. Wu, X. Shu H. Hovsepyan, R. D. Mosteller, D. Broek "VEGF Receptor Expression and Signaling in Human Bladder Tumors" *Oncogene*, **2003**, 22, 3361 – 3370.

7. C. A. Boocock, D. S. Charnock-Jones, A. M. Sharkey, J. McLaren, P. J. Barker, K. A. Wright, P. R. Twentyman, S. K. Smith "Expression of Vascular Endothelial Growth Factor and Its Receptros flt and KDR in Ovarian Carcinoma" *Journal of the National Cancer Institute*, **1995**, 87, 506 – 516.

8. C. V. Dang "C-Myc Target Genes Involved in Cell Growth, Apoptosis, and Metabolism" *Mol Cell Biol.*, **1999**, 19, 1 – 11.

9. M. Henriksson, B. Luscher "Proteins of the Myc Network: Essential Regulators of Cell Growth and Differentiation" *Adv Cancer Res.*, **1996**, 68, 82 – 109.

10. D. J. Slamon, J. B. deKernion, I. M. Verma, M. J. Cline "Expression of Cellular Oncogenes in Human Malignancies" *Science*, **1984**, 224, 256 – 262.

11. D. L. Vaux, S. Cory, J. M. Adams “Bcl-2 Gene Promotes Haemopoietic Cell Survival and Cooperates with C-myc to Immortalize Pre-B Cells” *Nature*, **1988**, 335, 440 – 442.

12. O. Y. Fedoroff, A. Rangan, V. V. Chemeris, L. H. Hurley “Cationic Porphyrins Promote the Formation of *i*-Motif DNA and Bind Peripherally by a Nonintercalative Mechanism” *Biochemistry*, **2000**, 39, 15083 – 15090.

13. X. Li, Y. Peng, J. Ren, X. Qu “Carboxyl-Modified Single-Walled Carbon Nanotubes Selectively Induce Human Telomeric *i*-Motif Formation” *PNAS*, **2006**, 103, 19658 – 19663.

14. A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley “Direct Evidence for a G-Quadruplex in a Promoter Region and its Targeting with a Small Molecule to Repress c-MYC Transcription” *PNAS*, **2002**, 99, 11593 – 11598.

15. K. Gehring, J. Leroy, M. Gueron “A Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs” *Nature*, **1993**, 363, 561 – 565.

16. J. Leroy, M. Gueron “Solution Structures of the *i*-Motif Tetramers of d(TCC), d(5methylCCT) and d(T5methylCC): Novel NOE Connections Between Amino Protons and Sugar Protons” *Structure*, **1995**, 3, 101 – 120.

17. S. Kendrick, H. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht, L. H. Hurley “The Dynamic Character of the BCL2 Promoter *i*-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure” *J. Am. Chem. Soc.*, **2014**, 136, 4161 – 4171.

18. S. O. Doronina, J Behr “Synthesis of 4-Guanidinopyrimidine Nucleosides for Triple Helix-Mediated Guanine and Cytosine Recognition” *Tetrahedron Letters*, **1998**, 39, 547 – 550.

19. J. Robles, A. Grandas, E. Pedroso “Synthesis of Modified Oligonucleotides Containing 4-Guanidino-2-pyrimidione Nucleobases” *Tetrahedron Letters*, **2001**, 57, 179 – 194.

20. R. H. E. Hudson, Y. Liu, F. Wojciechowski “Hydrophilic Modifications in Peptide Nucleic Acid – Synthesis and Properties of PNA Possessing 5-Hydroxymethyluracil and 5-Hydroxymethylcytosine” *Can. J. Chem.*, **2007**, 85, 302 – 312.

21. R. C. Moschel, E. J. Behrman “Oxidation of Nucleic Acid Bases by Potassium Peroxodisulfate in Alkaline Aqueous Solution” *J. Org. Chem.*, **1974**, 14, 1983 – 1989.

22. T. Kometani, D. S. Watt, T. Ji “Iodination of Phenols Using Chloramine T and Sodium Iodide” *Tetrahedron Letters*, **1985**, 26, 2043 – 2046.

23. D. L. Helfer II, R. S. Hosmane, N. J. Leonard “Selective Alkylation and Aralkylation of Cytosine at the 1-Position” *J. Org. Chem.,* **1981**, 46, 4803 – 4804.

24. K. J. Kennedy, T. L. Simandan, T. A. Dix “A Facile Route to Cyclic and Acyclic Alkyl-Arginines” *Synthetic Communications*, **1998**, 28, 741 – 746.

25. P. G. Baraldi, A. Bovero, F. Fruttarolo, R. Romagnoli, M. A. Tabrizi, D. Preti, K. Varani, P. A. Borea, A. R. Moorman “New Strategies for the Synthesis of A_3 Adenosine Receptor Antagonists” *Bioorg. Med. Chem.*, **2003**, 11, 4161 – 4169.

26. Pfizer, accessed 7 September 2015, http://www.pfizer.com/news/press-release/press-release-detail/pfizer_receives_u_s_fda_accelerated_approval_of_ibrance_palbociclib

27. N. T. Jacob, J. W. Lockner, V. V. Kravchenko, K. D. Janda “Pharmacophore Reassignment for Induction of the Immunosurveillance Cytokine TRAIL” *Angew. Chem. Int. Ed.*, **2014**, 53, 6628 – 6631.