Base–Pairing Energies of Protonated Nucleobase Pairs and Proton Affinities of 1–Methylated Cytosines: Model Systems for the Effects of the Sugar Moiety on the Stability of DNA i –M…

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JANUARY 2015

Impact Factor: 3.3 · DOI: 10.1021/acs.jpcb.5b00035 · Source: PubMed

4 AUTHORS, INCLUDING:

Bo Yang
Vanderbilt University
16 PUBLICATIONS **103** CITATIONS
SEE PROFILE

Mary T Rodgers
Wayne State University
125 PUBLICATIONS **4,182** CITATIONS
SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Base-Pairing Energies of Protonated Nucleobase Pairs and Proton Affinities of 1-Methylated Cytosines: Model Systems for the Effects of the Sugar Moiety on the Stability of DNA *i*-Motif Conformations

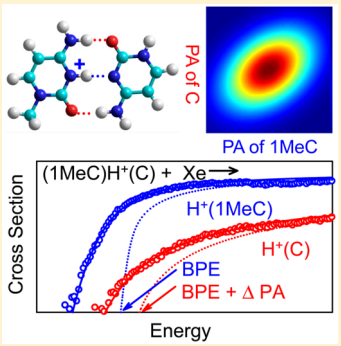
Bo Yang,[†] Aaron. R. Moehlig,[‡] C. E. Frieler,[†] and M. T. Rodgers*[,†]

[†]Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States

[‡]Department of Chemistry, Adams State University, Alamosa, Colorado 81101, United States

Supporting Information

ABSTRACT: Expansion of $(CCG)_n \cdot (CGG)_n$ trinucleotide repeats leads to hypermethylation of cytosine residues and results in Fragile X syndrome, the most common cause of inherited intellectual disability in humans. The $(CCG)_n \cdot (CGG)_n$ repeats adopt *i*-motif conformations that are preferentially stabilized by base-pairing interactions of noncanonical protonated nucleobase pairs of cytosine ($C^+ \cdot C$). Previously, we investigated the effects of 5-methylation of cytosine on the base-pairing energies (BPEs) using threshold collision-induced dissociation (TCID) techniques. In the present work, we extend our investigations to include protonated homo- and heteronucleobase pairs of cytosine, 1-methylcytosine, 5-methylcytosine, and 1,5-dimethylcytosine. The 1-methyl substituent prevents most tautomerization processes of cytosine and serves as a mimic for the sugar moiety of DNA nucleotides. In contrast to permethylation of cytosine at the 5-position, 1-methylation is found to exert very little influence on the BPE. All modifications to both nucleobases lead to a small increase in the BPEs, with 5-methylation producing a larger enhancement than either 1-methyl or 1,5-dimethylation. In contrast, modifications to a single nucleobase are found to produce a small decrease in the BPEs, again with 5-methylation producing a larger effect than 1-methylation. However, the BPEs of all of the protonated nucleobase pairs examined here significantly exceed those of canonical G·C and neutral C·C base pairs, and thus should still provide the driving force stabilizing DNA *i*-motif conformations even in the presence of such modifications. The proton affinities of the methylated cytosines are also obtained from the TCID experiments by competitive analyses of the primary dissociation pathways that occur in parallel for the protonated heteronucleobase pairs.



INTRODUCTION

Epigenetic traits, defined as stable heritable phenotypes caused by alterations in a chromosome without changes in the DNA sequence,[1] have drawn great attention among biologists in the past several decades. DNA methylation is one of the most widely investigated epigenetic modifications of the genome that can regulate chromatin status and directly influence the interactions between transcription factors and DNA. In the context of gene promoters, hypomethylation of cytosine residues is associated with active, constitutively expressed genes, whereas hypermethylation of cytosine residues is associated with silenced genes.[2] Indeed, cytosine methylation is a significant contributor to the generation of germline[3] and somatic mutations that cause diseases and cancer.[4] Previous investigations have found that 5-permethylation of cytosine causes Fragile X syndrome.

Fragile X syndrome is the most common inherited cause of intellectual disabilities in humans, and leads to developmental problems and physical handicaps.[5,6] Fragile X syndrome occurs in approximately 1 of every 4000 males and 1 of every 6000–8000 females, and is common to all races and ethnicities.[7] It has previously been shown that the expansion of $(CCG)_n \cdot (CGG)_n$ trinucleotide repeats to more than 230 trinucleotides leads to marked methylation of both the CCG and CGG trinucleotide repeats as well as the *Fragile X mental retardation 1* (*FMR1*) gene promoter on the X chromosome, resulting in an inability to express the protein that is necessary for normal neural development.[8−10] The single $(CCG)_n$ and $(CCG)_n \cdot (CGG)_n$ strands are believed to adopt noncanonical DNA *i*-motif conformations that may be the cause of the disease.[11,12] The DNA *i*-motif conformation was first observed in NMR studies of DNA oligomers[13] and later in $(CCG)_n \cdot (CGG)_n$ trinucleotide repeats.[14] The DNA *i*-motif is a tetrameric DNA secondary structure consisting of parallel-stranded DNA duplexes held together by intercalated protonated nucleobase pairs of cytosine ($C^+ \cdot C$) in an antiparallel orientation.[13] Ever since the first report of the DNA *i*-motif, the biomedical functions of DNA *i*-motif conformations and their fascinating material properties for nanotechnology applications have received great attention. The structure of the *i*-motif is also conserved in the gas phase when the complex of interest is generated by electrospray ionization (ESI),[15] suggesting that gas-phase

Received: October 9, 2014
Published: January 7, 2015

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

studies of ESI generated species may provide valuable information regarding solution-phase structure and function.

The structure of the protonated nucleobase pair of cytosine is well established. It involves N3-protonation of the canonical form of cytosine and is noncovalently bound via three hydrogen-bonding interactions.[13,16,17] Recent work has shown that methylation of a single cytosine residue of a single DNA strand allows formation of the DNA *i*-motif conformation under proper physiological conditions.[18] However, the diseased state of Fragile X syndrome is related to hypermethylation of cytosine residues in the DNA *i*-motif conformations.[8,9] Given the roles that DNA *i*-motif conformations may play in the development of cancer and several human diseases including lung carcinoma,[19] breast carcinoma,[20] and Burkitt's lymphomas,[21] a comprehensive and pedagogic investigation is necessary to determine the base-pairing energies (BPEs) of protonated nucleobase pairs of cytosine and methylated cytosines to characterize the influence of methylation on the BPEs. Methylation also affects the proton affinities (PAs) of the nucleobases. Accurate determination of the proton affinities (PAs) of methylated cytosines provides valuable information regarding the intrinsic nucleobase reactivity and the effects of solvent on that reactivity. However, thus far, PAs have only been reported for cytosine (C)[22,23] and 1-methylcytosine (1MeC).[24]

The B3LYP/def2-TZVPPD level of theory has been found to provide accurate energetic descriptions of similar hydrogen-bonded systems.[25,26] Therefore, in the present work, theoretical calculations are performed on the protonated nucleobase pair of cytosine (C^+·C) and the canonical G·C and neutral C·C nucleobase pairs using this level of theory. The BPE calculated for C^+·C is 170.1 kJ/mol, and is 79 and 163% larger than the BPEs of the canonical G·C and neutral C·C base pairs, respectively, consistent with previous results and indicating that the very strong base-pairing interactions in the protonated C^+·C nucleobase pair are likely the driving force that leads to double stranded DNA undergoing conversion to noncanonical DNA *i*-motif conformations.

Recently, the gas-phase conformations of the protonated homo- and heteronucleobase pairs of cytosine (C) and 5-methylcytosine (5MeC) were examined via complementary infrared multiple photon dissociation (IRMPD) action spectroscopy experiments and theoretical electronic structure calculations.[17] The structure of the protonated C^+·C nucleobase pair was found to be the same as that observed in condensed-phase NMR studies.[13] In addition, 5-methylation of cytosine residues does not alter the preferred hydrogen-bonding pattern. The BPEs of these protonated nucleobase pairs were quantitatively determined using threshold collision-induced dissociation (TCID) techniques.[25,26] It was determined that 5-permethylation of cytosine leads to an increase in the BPE and should therefore stabilize DNA *i*-motif conformations, whereas 5-methylation of only one of the cytosine residues decreases the BPE and would therefore slightly destabilize the *i*-motif. In related work, the gas-phase structure of the protonated homonucleobase pair of 1MeC was also studied using IRMPD action spectroscopy techniques.[16] As found for 5-methylation, 1-methylation does not change the preferred base-pairing interactions. However, to date, the effects of 1-methylation on the BPEs have not been examined. In order to achieve a comprehensive understanding of the effects of methylation on the stabilities of protonated nucleobase pairs of cytosine and determine PAs of the methylated cytosines, we expand the systems examined to include 1MeC and 1,5-dimethylcytosine (15dMeC) in the present work. The bulky methyl group at the N1 position serves as a mimic for the sugar moiety such that implications for the effects of the 2′-deoxyribose moiety on the BPE can be elucidated as well. The BPEs of the protonated homo- and heteronucleobase pairs of cytosine and methylated cytosines generated by ESI are determined using a guided ion beam tandem mass spectrometer and TCID techniques. Relative N3 proton affinities (PAs) of the methylated cytosines are also obtained from the TCID experiments from competitive analyses of the two primary dissociation pathways that occur in parallel for the protonated heteronucleobase pairs of cytosine and methylated cytosines. Absolute N3 PAs of the methylated cytosines are then obtained via a maximum likelihood statistical analysis using the relative PAs determined here and the PAs of C[22,23] and 1MeC[24] reported in the literature. The measured values are compared with calculated values to evaluate the accuracy of the energetics predicted by each level of theory.

EXPERIMENTAL AND COMPUTATIONAL APPROACHES

Experimental Procedures. Seven protonated (xC)H^+(yC) nucleobase pairs, including (1MeC)H^+(1MeC), (15dMeC)H^+(15dMeC), (1MeC)H^+(C), (15dMeC)H^+(C), (1MeC)H^+(5MeC), (15dMeC)H^+(5MeC), and (15dMeC)H^+(1MeC), were studied using TCID techniques. Measurements were carried out using a guided ion beam tandem mass spectrometer that has been described in detail previously,[27] and that has since been coupled to an electrospray ionization (ESI) source. The (xC)H^+(yC) protonated nucleobase pairs were generated by ESI using solution conditions similar to those described for related protonated nucleobase pairs.[25,26] The ions are desolvated, focused, and thermalized in an rf ion funnel/hexapole ion guide collision cell interface, and extracted, accelerated, and focused into a magnetic sector momentum analyzer for mass analysis. Ions of interest are mass-selected, slowed down to a well-defined kinetic energy, and focused into a radio frequency (rf) octopole ion beam guide that provides a radial trapping field for ions[28−30] such that scattered reactant (xC)H^+(yC) and CID product ions, H^+(xC) and H^+(yC), are not lost as they drift through the octopole. The octopole passes through a static gas cell containing Xe at low pressure (~0.05−0.10 mTorr) such that the (xC)H^+(yC) protonated nucleobase pairs undergo collision-induced dissociation (CID) with Xe[31−33] under conditions that generally result in at most a single collision between the protonated nucleobase pair and Xe. After the collision, product H^+(xC) and H^+(yC) ions and undissociated (xC)H^+(yC) precursor ions are focused into a quadrupole mass filter for mass analysis, and detected using a detector of the Daly type and standard pulse counting techniques. C and 5MeC were acquired from Alfa Aesar (Massachusetts, USA), whereas 1MeC, 1Me-d_3-C, and 15dMeC were provided by Professor T. H. Morton of the University of California, Riverside.

Theoretical Calculations. The stable low-energy tautomeric conformations of xC, H^+(xC), and (xC)H^+(yC) protonated nucleobase pairs, where (xC)H^+(yC) = (C)H^+(C), (5MeC)H^+(C), and (5MeC)H^+(5MeC), have previously been examined and described in detail.[17,25,26] In the present study, the stable low-energy tautomeric conformations of xC and H^+(xC), where xC = C, 5MeC, 1MeC, 1Me-d_3-C, and

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

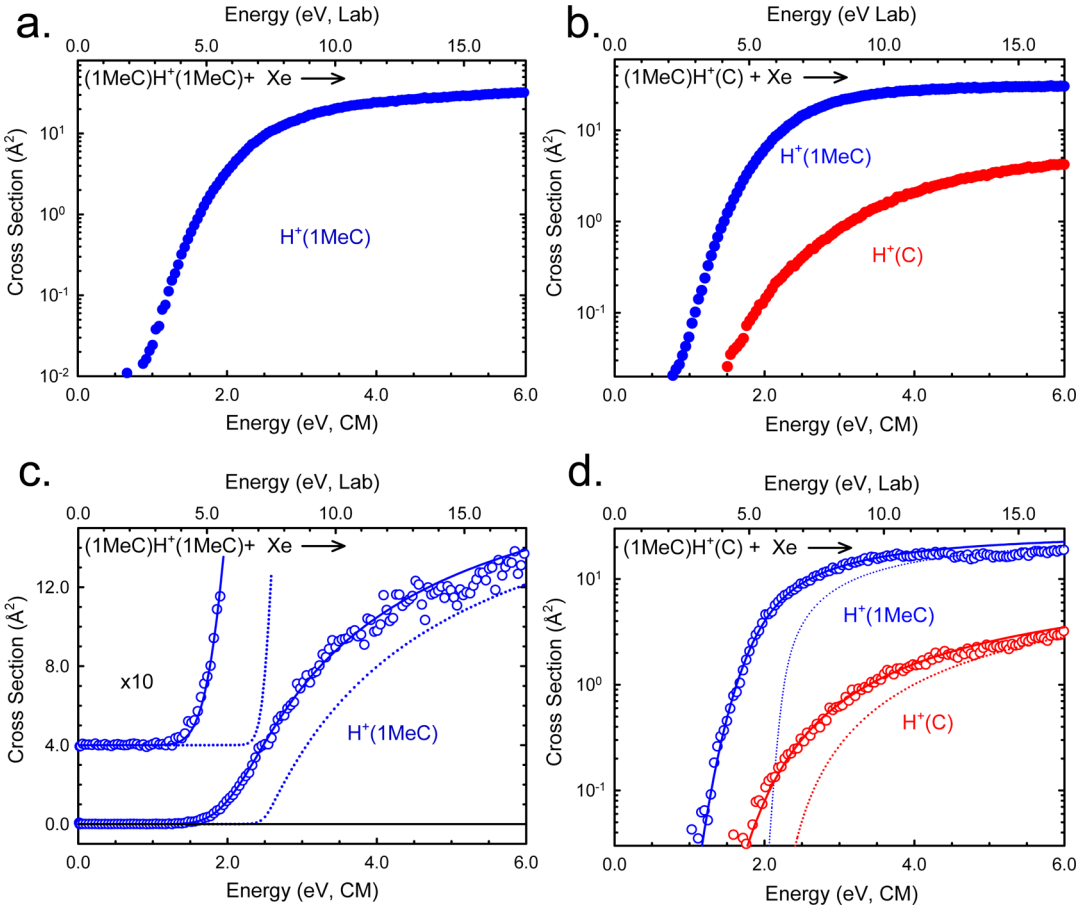


Figure 1. Cross sections for collision-induced dissociation of the $(1MeC)H^+(1MeC)$ and $(1MeC)H^+(C)$ complexes with Xe as a function of collision energy in the center-of-mass frame (lower *x*-axis) and laboratory frame (upper *x*-axis), parts a and b, respectively. Data are shown for a Xe pressure of ~0.1 mTorr. Zero-pressure-extrapolated cross sections for collision-induced dissociation of the $(1MeC)H^+(1MeC)$ and $(1MeC)H^+(C)$ complexes with Xe in the threshold region as a function of kinetic energy in the center-of-mass frame (lower *x*-axis) and the laboratory frame (upper *x*-axis), parts c and d, respectively. The solid lines show the best fit to the data using the models of eqs S1 and S2 (Supporting Information) convoluted over the neutral and ion kinetic and internal energy distributions. The dotted lines show the model cross sections in the absence of experimental kinetic energy broadening for the $(1MeC)H^+(1MeC)$ and $(1MeC)H^+(C)$ complexes with an internal temperature of 0 K. The data and models are shown expanded by a factor of 10 and offset from zero in the inset of part c.

15dMeC, and seven protonated nucleobase pairs, including $(1MeC)H^+(1MeC)$, $(15dMeC)H^+(15dMeC)$, $(1MeC)H^+(C)$, $(15dMeC)H^+(C)$, $(1MeC)H^+(5MeC)$, $(15dMeC)H^+(5MeC)$, and $(15dMeC)H^+(1MeC)$, were geometry optimized and vibrational frequencies were calculated at three different levels of theory, B3LYP/6-31G*, B3LYP/def2-TZVPPD, and MP2-(full)/6-31G*, using the Gaussian 09[34] suite of programs. The def2-TZVPPD basis set[35] is a balanced triple-ζ basis set that includes polarization and diffuse functions, which was obtained from the EMSL basis set exchange library.[36,37]

Polarizability is one of the intrinsic factors that impacts the strength of noncovalent interactions. The phase space limit transition state (PSL TS) model used for thermochemical analysis of the CID cross sections requires knowledge of the polarizabilities of the departing neutral nucleobases. Therefore, the polarizabilities of the neutral nucleobases are calculated at the PBE1PBE/6-311+G(2d,2p) level of theory, which was chosen as it has been shown to provide polarizabilities that exhibit better agreement with measured values than the four levels of theory used in this work for structures and energetics.[38] To obtain candidate structures for the transition states (TSs) for dissociation of the N3-protonated ground conformers of the protonated nucleobase pairs to produce O2-protonated ground conformers of the protonated nucleobase products, relaxed potential energy surface (PES) scans were calculated using the B3LYP/6-31G* level of theory. The actual TSs were obtained using the quasi synchronous transit method, QST3,[39] and the input from the relevant minima (reactants and products) and an estimate for the TS taken as the least stable structures along the relaxed PES scans at the B3LYP/6-31G*, B3LYP/def2-TZVPPD, and MP2(full)/6-31G* levels of theory. Single point energies of the *x*C, $H^+(x\mathrm{C})$, TSs, and seven $(x\mathrm{C})H^+(y\mathrm{C})$ protonated nucleobase pairs were calculated at the B3LYP/6-311+G(2d,2p), B3LYP/def2-TZVPPD, and MP2(full)/6-311+G(2d,2p) levels of theory using the B3LYP/6-31G*, B3LYP/def2-TZVPPD, and MP2(full)/6-31G* optimized structures, respectively. Single point energy calculations performed at the MP2(full)/def2-TZVPPD level make use of geometries determined at the B3LYP/def2-TZVPPD level of theory because frequency analyses at this level require computational resources beyond those available to us. Zero-point energy (ZPE) corrections were determined using B3LYP and MP2(full) calculated vibrational frequencies scaled by factors of 0.9804 and 0.9646, respectively.[40] The calculated BPEs also include basis set superposition error (BSSE)

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

corrections using the full counterpoise approach to obtain accurate energetics.[41,42]

Thermochemical Analysis of CID Cross Sections. The threshold regions of the measured CID cross sections of seven protonated nucleobase pairs are modeled using approaches developed and described in detail elsewhere.[43−49] The data handling and analysis procedures explicitly include the internal energy distributions of the protonated nucleobase pairs, the translation energy distributions of both reactants, the effects of multiple ion-neutral collisions, and the lifetime of the dissociating protonated nucleobase pairs. These approaches have been shown to reproduce CID cross sections well and enable extraction of accurate thermochemical data.[50−54] Data handling and analysis procedures are summarized in detail in the Supporting Information associated with this article.

Maximum-Likelihood Estimate of the N3 Proton Affinities. The absolute N3 PAs of C, 5MeC, 1MeC, and 15dMeC are determined by performing a statistical analysis that determines the most probable values, i.e., the maximum likelihood estimate (MLE), and the uncertainties in these values. This method assumes that the data is reliable and follows Gaussian statistics, that the measured values represent good estimates of the population means, and that the reported uncertainties represent good estimates of the standard deviations in these values. The MLE PA values are extracted on the basis of the literature values reported for the absolute PAs of C[22,23] and 1MeC,[24] and the relative PAs of these four nucleobases derived from our TCID studies of the protonated heteronucleobase pairs measured here as well as the value for the C, 5MeC pair reported in our earlier work.[25] In the analysis, the reported uncertainties are assumed to describe the width of the normal probability distributions. Because each of the measurements is independent, the combined MLE for all four nucleobases is simply the product of the likelihood of each of the measurements. Normalization of the resulting likelihood distribution transforms it into a formal probability density function. The peak of this function is the MLE for the PAs, whereas the 2-D cross sections of the probability density function provide the residual uncertainties in these estimates.

RESULTS

Cross Sections for Collision-Induced Dissociation. TCID cross sections were measured for the interaction of Xe with seven $(x\mathrm{C})\mathrm{H}^+(y\mathrm{C})$ protonated nucleobase pairs, $(1\mathrm{MeC})\mathrm{H}^+(1\mathrm{MeC})$, $(15\mathrm{dMeC})\mathrm{H}^+(15\mathrm{dMeC})$, $(1\mathrm{MeC})\mathrm{H}^+(\mathrm{C})$, $(15\mathrm{dMeC})\mathrm{H}^+(\mathrm{C})$, $(1\mathrm{MeC})\mathrm{H}^+(5\mathrm{MeC})$, $(15\mathrm{dMeC})\mathrm{H}^+(5\mathrm{MeC})$, and $(15\mathrm{dMeC})\mathrm{H}^+(1\mathrm{MeC})$. The energy dependent CID cross sections of the $(1\mathrm{MeC})\mathrm{H}^+(1\mathrm{MeC})$ and $(1\mathrm{MeC})\mathrm{H}^+(\mathrm{C})$ protonated nucleobase pairs are shown in Figure 1. The TCID behavior of the other five $(x\mathrm{C})\mathrm{H}^+(y\mathrm{C})$ protonated nucleobase pairs is highly similar, and figures showing data for these systems are included in the Supporting Information as Figure S1. Over the collision energy range examined, typically ~0−6 eV, the only CID pathway observed for the protonated homonucleobase pairs corresponds to cleavage of the three hydrogen-bonding interactions that stabilize these species, and resulting in loss of the neutral nucleobase in CID reaction 1.

$$(x\mathrm{C})\mathrm{H}^+(x\mathrm{C}) + \mathrm{Xe} \rightarrow \mathrm{H}^+(x\mathrm{C}) + x\mathrm{C} + \mathrm{Xe} \quad (1)$$

CID of the $(x\mathrm{C})\mathrm{H}^+(y\mathrm{C})$ protonated heteronucleobase pairs leads to two dissociation pathways that occur in parallel and compete with one another, CID reactions 2 and 3.

$$(x\mathrm{C})\mathrm{H}^+(y\mathrm{C}) + \mathrm{Xe} \rightarrow \mathrm{H}^+(x\mathrm{C}) + y\mathrm{C} + \mathrm{Xe} \quad (2)$$

$$(x\mathrm{C})\mathrm{H}^+(y\mathrm{C}) + \mathrm{Xe} \rightarrow \mathrm{H}^+(y\mathrm{C}) + x\mathrm{C} + \mathrm{Xe} \quad (3)$$

This behavior is consistent with fragmentation via IRMPD[16,17] and CID[25,26] of similar protonated nucleobase pairs. Production of the protonated nucleobase having the higher PA is energetically favored over production of the protonated nucleobase with the lower PA such that the CID results indicate that the relative N3 PAs follow the order 15dMeC > 1Me-d_3-C, 1MeC > 5MeC > C.

Theoretical Results. As mentioned in the Theoretical Calculations section, the stable tautomeric conformations of the neutral xC and protonated $\mathrm{H}^+(x\mathrm{C})$ nucleobases and the $(x\mathrm{C})\mathrm{H}^+(y\mathrm{C})$ protonated nucleobase pairs, where $(x\mathrm{C})\mathrm{H}^+(y\mathrm{C})$ = $(\mathrm{C})\mathrm{H}^+(\mathrm{C})$, $(5\mathrm{MeC})\mathrm{H}^+(\mathrm{C})$, and $(5\mathrm{MeC})\mathrm{H}^+(5\mathrm{MeC})$, have previously been examined using the B3LYP and MP2 levels of theory with 6-31G* and def2-TZVPPD basis sets in the IRMPD and TCID studies we previously reported.[17,25,26] These calculations are extended in this work to include structures for neutral and protonated 1MeC, 1Me-d_3-C, and 15dMeC as well as the protonated homo- and heteronucleobase pairs of these species with C and 5MeC optimized at the same levels of theory as described in the Theoretical Calculations section. The geometry-optimized structures of the three most stable tautomeric conformations computed and their 0 K enthalpies and 298 K relative Gibbs free energies of neutral xC and protonated $\mathrm{H}^+(x\mathrm{C})$ are included in Figure S2 of the Supporting Information. The relative stabilities of the two most stable tautomeric conformations of these species are influenced by electronic and steric effects arising from the 1- and 5-methyl substituents, as summarized in Table S3. The various stable low-energy tautomeric conformations of these species are described using lowercase Roman numerals for the neutral nucleobase and uppercase Roman numerals with a "+" sign for the protonated nucleobases. All species are ordered on the basis of the relative 298 K Gibbs free energies of the analogous low-energy tautomeric conformations of C and $\mathrm{H}^+(\mathrm{C})$. The B3LYP/def2-TZVPPD optimized structures of the ground-state conformations of the seven protonated nucleobase pairs examined here are compared in Figure 2. The ground structures of all seven protonated nucleobase pairs are bound by three hydrogen-bonding interactions. The protonated and neutral nucleobases adopt an antiparallel configuration, which is the most commonly observed conformation in multistranded DNAs. In the ground-state tautomeric conformers of the heterodimers, the excess proton is bound to the nucleobase having the higher PA. This ground-state conformer is designated as **II**$^+$···**i_3a** to indicate that the excited **II**$^+$ tautomeric conformation of the protonated nucleobase, $\mathrm{H}^+(x\mathrm{C})$, is bound to the ground-state **i** tautomeric conformation of the neutral nucleobase, xC or yC via three hydrogen-bonding interactions, and the protonated and neutral nucleobases are bound in an antiparallel configuration. It is unknown whether these protonated nucleobase pairs will undergo tautomerization to the O2-protonated nucleobases, **I**$^+$, upon CID. Therefore, PES scans and explicit TS calculations were performed to estimate the height of the tautomerization barriers. The reaction coordinate diagrams for fragmentation of the $(1\mathrm{MeC})\mathrm{H}^+(1\mathrm{MeC})$ and $(1\mathrm{MeC})\mathrm{H}^+(\mathrm{C})$ protonated nucleobase pairs to produce N3-protonated **II**$^+$ or O2-protonated **I**$^+$ products are shown in Figure 3. Highly parallel results are found for the other five protonated nucleobase pairs

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

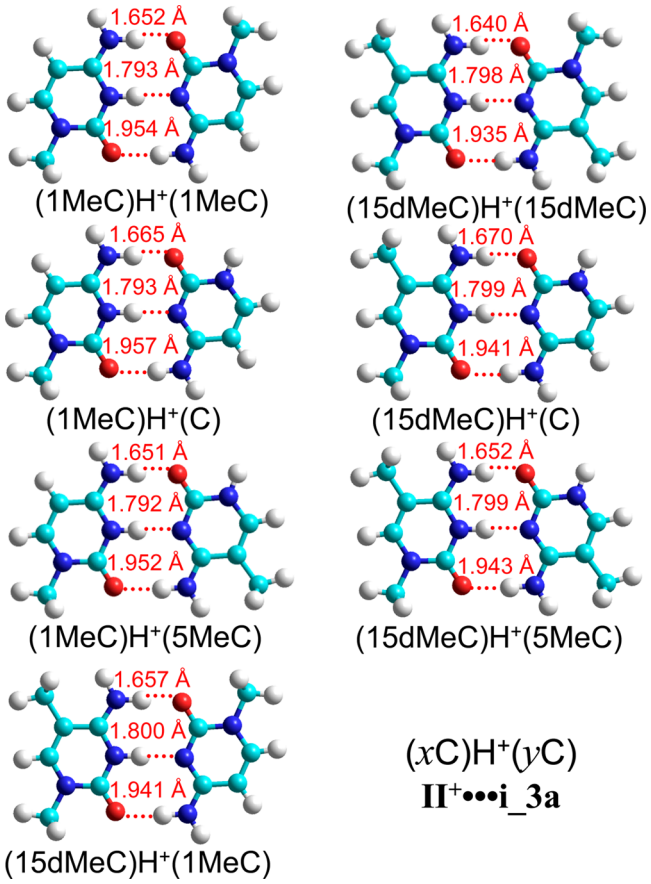


Figure 2. B3LYP/def2-TZVPPD optimized geometries of the ground-state **II**$^+$**•••I_3a** conformations of seven $(x\text{C})\text{H}^+(y\text{C})$ protonated nucleobase pairs.

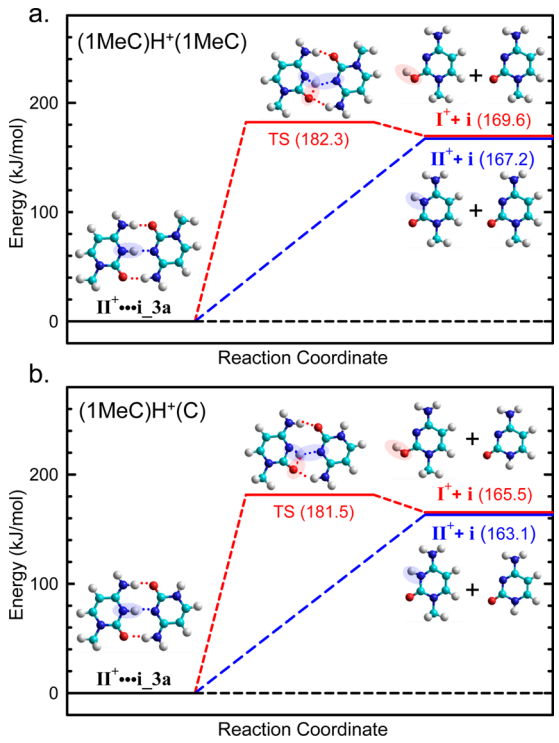


Figure 3. B3LYP/def2-TZVPPD potential energy surfaces for dissociation of the ground-state **II**$^+$**•••I_3a** conformations of the $(1\text{MeC})\text{H}^+(1\text{MeC})$ and $(1\text{MeC})\text{H}^+(\text{C})$ complexes to produce 1MeC_**i** or C_**i** and protonated $\text{H}^+(1\text{MeC})$_**I**$^+$ products and neutral 1MeC_**i** or C_**i** and protonated $\text{H}^+(1\text{MeC})$_**II**$^+$ products, parts a and b, respectively.

and are included in the Supporting Information as Figure S3. The relative energies along the PESs for the dissociation of all seven protonated nucleobase pairs to produce N3-protonated **II**$^+$ or O2-protonated **I**$^+$ products calculated here are summarized in Table S4 of the Supporting Information. The TSs of all seven protonated nucleobase pairs involve bidentate interaction of the excess proton with the O2 and N3 atoms. The barriers to tautomerization (182.3 and 181.5 kJ/mol) in the PESs of Figure 3 exceed the dissociation energies for simple cleavage of the three hydrogen bonds (167.2 and 163.0 kJ/mol) of the $(1\text{MeC})\text{H}^+(1\text{MeC})$ and $(1\text{MeC})\text{H}^+(\text{C})$ complexes by 15.1 and 18.5 kJ/mol, respectively, indicating that at threshold energies tautomerization cannot occur. The B3LYP/def2-TZVPPD calculated tautomerization barriers for the other $(x\text{C})\text{H}^+(y\text{C})$ dimers lie in the range 179.9−183.5 kJ/mol, and also exceed the dissociation energies for simple noncovalent bond cleavage (161.5−171.1 kJ/mol) by 11.1−19.2 kJ/mol. To examine whether or not the TS barriers are significantly impacted by the basis sets and level of theory used, the tautomerization barriers are also predicted using the B3LYP/6-311+G(2d,2p), MP2(full)/6-311+G(2d,2p), and MP2(full)/def2-TZVPPD levels of theory. As can be seen in Table S4 (Supporting Information), regardless of the level of theory employed, the calculated tautomerization barriers exceed the dissociation energies for simple noncovalent bond cleavage for all seven protonated nucleobase pairs, confirming that tautomerization will not occur upon dissociation at threshold energies. Therefore, the BPEs for simple cleavage of the three hydrogen bonds of the $(x\text{C})\text{H}^+(y\text{C})$ protonated nucleobase pairs and site-specific N3 PAs of *x*Cs are determined from the experiments. BPEs including ZPE and BSSE corrections calculated for the dissociation pathway that produces the N3-protonated product (**II**$^+$) at the B3LYP/6-311+G(2d,2p), B3LYP/def2-TZVPPD, MP2(full)/6-311+G(2d,2p), and MP2(full)/def2-TZVPPD levels of theory are summarized in Table 1. The BPEs calculated for the $(1\text{MeC})\text{H}^+(5\text{MeC})$ and $(1\text{Me-}d_3\text{-C})\text{H}^+(5\text{MeC})$ protonated nucleobase pairs and the N3 PAs of 1MeC and 1Me-d_3-C are equal, indicating that TCID experiments for the $(1\text{Me-}d_3\text{-C})\text{H}^+(5\text{MeC})$ complexes can be used to determine the BPE for the $(1\text{MeC})\text{H}^+(5\text{MeC})$ complex and relative N3 PAs of 1MeC and 5MeC.

Threshold Analysis of CID Cross Sections. The thresholds for reaction 1 for two $(x\text{C})\text{H}^+(x\text{C})$ protonated homonucleobase pairs are modeled using eq S1, whereas the thresholds for reactions 2 and 3 for five $(x\text{C})\text{H}^+(y\text{C})$ protonated heteronucleobase pairs are modeled using eq S2, as described in detail in the Supporting Information. As concluded from the computed barriers to tautomerization, dissociation at threshold energies will produce the neutral and protonated nucleobase products in the same tautomeric forms as in the protonated nucleobase pairs, **II**$^+$ and **i**. Theoretical results also find that the hydrogen bond involving the excess proton provides the majority of the stabilization energy (~100 kJ/mol) for the protonated nucleobase pair, whereas the two additional neutral hydrogen bonds each contribute ~30 kJ/mol of additional stabilization. The dissociation coordinate involves lengthening and cleavage of the N3−H$^+$···N3 hydrogen bond, which results in simultaneous lengthening and cleavage of the two neutral hydrogen bonds. On the basis of the

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

Table 1. Base-Pairing Energies of the (*x*C)H$^+$(*y*C) Complexes at 0 K in kJ/mol[a]

		B3LYP[b]		MP2(full)[c]	
*x*C, *y*C	TCID	D_0	$D_{0,BSSE}$[d]	D_0	$D_{0,BSSE}$[d]
C, C	169.9 (4.6)[e]	171.7[e]	168.9[e]	155.2[e]	136.7[e]
		170.1[e]	***169.2***[e]	***149.3***[e]	***136.0***[e]
5MeC, 5MeC	177.4 (5.3)[e]	176.3[e]	173.4[e]	160.0[e]	141.0[e]
		174.2[e]	***173.3***[e]	***153.3***[e]	***140.4***[e]
1MeC, 1MeC	170.7 (5.3)	169.7	166.8	144.6	125.5
		167.2	***166.3***	***138.7***	***125.1***
1SdMeC, 1SdMeC	172.3 (5.8)	169.7	166.8	144.6	125.6
		171.1	***170.2***	***136.6***	***122.8***
5MeC, C	163.6 (5.1)[f]	169.7[f]	166.8[f]	153.9[f]	141.0[f]
		168.0[f]	***167.1***[f]	***145.5***[f]	***132.3***[f]
1MeC, C	166.6 (4.5)	164.1	161.8	148.5	129.8
		163.1	***162.2***	***140.2***	***126.7***
1SdMeC, C	163.4 (4.6)	163.1	160.3	147.6	128.7
		161.5	***160.6***	***138.7***	***125.6***
1MeC, 5MeC	170.1 (4.5)	170.8	167.9	154.2	135.5
		168.9	***167.9***	***149.3***	***134.6***
1SdMeC, 5MeC	163.6 (5.2)	169.0	166.2	153.3	134.7
		167.0	***166.1***	***147.1***	***133.9***
1SdMeC, 1MeC	160.9 (4.7)	166.9	164.0	151.1	132.3
		165.4	***164.5***	***146.1***	***132.7***
AEU/MAD[g]	4.9 (0.5)	2.6 (2.3)	3.6 (1.2)	17.6 (7.0)	36.5 (7.2)
		2.8 (1.4)	***3.0 (0.8)***	***23.0 (10.3)***	***38.0 (7.8)***

[a]Present results, uncertainties are listed in parentheses. [b]Calculated at the B3LYP level of theory including ZPE corrections. Values obtained using the 6-311+G(2d,2p) basis set are shown in standard font, whereas those computed using the ***def2-TZVPPD*** basis set are shown in bold-italic. [c]Calculated at the MP2(full) level of theory including ZPE corrections. Values obtained using the 6-311+G(2d,2p) basis set are shown in standard font, whereas those computed using the ***def2-TZVPPD*** basis set are shown in bold-italic and make use of the B3LYP/def2-TZVPPD optimized geometries. [d]Also includes BSSE corrections. [e]Values taken from ref 26. [f]Values taken from ref 25. [g]Average experimental uncertainty (AEU). Mean absolute deviation (MAD) between the measured and computed values.

computational results, a loose phase space limit (PSL) TS model[48] is employed. The results of threshold analysis are summarized in Table S5 of the Supporting Information. The threshold energies determined are also listed in Table 1, and representative fits to the CID cross sections of the (1MeC)H$^+$(1MeC) and (1MeC)H$^+$(C) protonated nucleobase pairs are displayed in Figure 1. Similar results are obtained for the other five (*x*C)H$^+$(*y*C) protonated nucleobase pairs and are included in Figure S4 of the Supporting Information. For the homodimers, the measured cross sections for reaction 1 are accurately reproduced using a loose PSL TS model[48] for the **II$^+$···i_3a** → **II$^+$** + **i** CID pathway. In the cases of the heterodimers, the measured cross sections for reactions 2 and 3 are accurately reproduced using the loose PSL TS model[48] for both **II$^+$···i_3a** → **II$^+$** + **i** CID pathways. The relative N3 PAs of cytosine and the methylated cytosines are also determined from competitive analyses of these dissociation pathways for the protonated heteronucleobase pairs. The TCID measured and B3LYP/def2-TZVPPD calculated relative N3 PAs are summarized in Table 2. The measured CID cross sections are also analyzed using the tight TS associated with tautomerization of the protonated nucleobase from the **II$^+$** to the **I$^+$** tautomeric conformation upon dissociation of the protonated nucleobase pair (see Figure 3 and Figure S3, Supporting Information) to ensure proper interpretation of the TCID data. However, the thresholds extracted are ∼50 kJ/mol lower, and the resultant entropies of activation are negative. The threshold values determined from these latter analyses exhibit extremely poor agreement with all of the calculated values, confirming that these protonated nucleobase pairs dissociate via a loose PSL TS

Table 2. Relative N3 PAs of Cytosine and Methylated Cytosines at 298 K in kJ/mol

base pair	ΔN3 PA[a]	ΔN3 PA_{calc}[b]
1MeC, C	14.6 (2.0)	14.6
5MeC, C	15.5 (2.1)[c]	10.5
1SdMeC, C	27.6 (3.4)	24.2
1MeC, 5MeC	1.8 (2.0)	4.1
1SdMeC, 1MeC	12.5 (2.0)	15.7
1SdMeC, 5MeC	18.5 (3.7)	11.6

[a]Adjusted to 298 K using thermal corrections based on the ΔH_{298} values listed in Table S6 of the Supporting Information. Uncertainties are listed in parentheses. [b]Values calculated at the B3LYP/def2-TZVPPD level of theory. [c]Value taken from ref 25.

to produce the N3-protonated nucleobase, **II$^+$**, at threshold energies. The threshold energies, E_0, obtained when RRKM lifetime effects are not incorporated in the analysis, are also included in Table S5 (Supporting Information). Comparison of the E_0 and E_0(PSL) values provides a measurement of the kinetic shift associated with the finite experimental time window.

The entropy of activation, $\Delta S^{\ddagger}$, is a reflection of the size and complexity of the system and also provides a measure of the looseness of the TS. The entropy of activation is largely determined from the molecular constants used to model the energized complex and the TS, which are listed in Tables S1 and S2 (Supporting Information), but also depends on the threshold energy, E_0(PSL). The $\Delta S^{\ddagger}$(PSL) values at 1000 K are included in Table S5 (Supporting Information), and lie between 91 and 104 J·K^{-1}·mol^{-1} for these systems. The large

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

positive entropies of activation determined result from the fact that three noncovalent interations that conformationally constrain the reactant protonated nucleobase pair are eliminated such that upon dissociation there is a large increase in entropy.

DISCUSSION

Comparison of Measured and Calculated BPEs. The TCID measured BPEs of the protonated nucleobase pairs at 0 K are listed in Table 1. Also summarized in Table 1 are the BPEs of the protonated nucleobase pairs calculated at the B3LYP/6-311+G(2d,2p), B3LYP/def2-TZVPPD, MP2(full)/6-311+G(2d,2p), and MP2(full)/def2-TZVPPD levels of theory, and including independent ZPE and BSSE corrections. The B3LYP/def2-TZVPPD calculated and TCID measured BPEs of the protonated homo- and heteronucleobase pairs are compared in Figure 4. Parallel comparisons for all four levels of theory examined here are shown in Figure S5 of the Supporting Information. The calculated and measured BPEs for the protonated nucleobase pairs of cytosine and 5MeC are taken from refs 25 and 26 and are included for comparison. Overall, very good agreement is achieved for the B3LYP results with the measured BPEs, whereas the MP2(full) theory systematically underestimates the BPEs. The mean absolute deviations (MADs) between the B3LYP/def2-TZVPPD and B3LYP/6-311+G(2d,2p) calculated results and the TCID measured BPEs are 3.0 ± 0.8 and 3.6 ± 1.2 kJ/mol, respectively, and are smaller than the average experimental uncertainty (AEU) in these values, 4.9 ± 0.5 kJ/mol. These comparisons suggest that the B3LYP theory provides accurate descriptions of the base-pairing interactions responsible for the binding in these protonated nucleobase pairs, with the def2-TZVPPD basis set providing slightly better results. The MP2(full) theory does not describe the energetics of binding nearly as well. The MADS between experiment and the MP2(full)/def2-TZVPPD and MP2(full)/6-311+G(2d,2p) levels of theory are 38.0 ± 7.8 and 36.5 ± 7.2 kJ/mol, respectively, and significantly exceed the MADs for the B3LYP values and the AEU. The MADs between the MP2(full) calculated and TCID measured BPEs decrease to 23.0 ± 10.3 and 17.7 ± 7.0 kJ/mol when BSSE corrections are not included, consistent with previous TCID studies on similar protonated nucleobase pairs.[25,26] Similar observations were made in other theoretical investigations of hydrogen-bonded complexes,[55−62] which have determined that basis sets of at least triple-ζ quality are necessary to provide accurate descriptions of complexes involving intramolecular hydrogen-bonding interactions, and the BSSE corrections can be quite significant for MP2(full) results when flexible yet unsaturated basis sets are employed. On the basis of present comparisons, it can be concluded that B3LYP provides accurate descriptions of the strength of the base-pairing interactions in these protonated nucleobase pairs, whereas MP2(full) systematically underpredicts the strength of the base-pairing interactions in all seven of the protonated nucleobase pairs examined here.

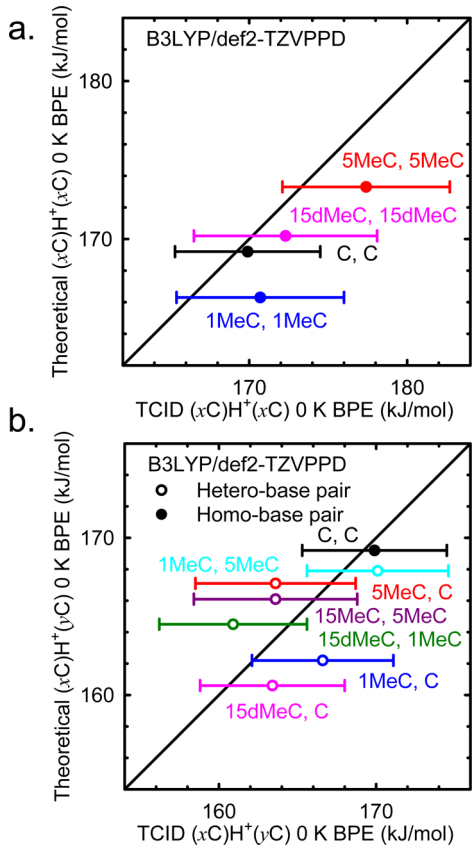


Figure 4. B3LYP/def2-TZVPPD calculated BPEs of $(x\text{C})\text{H}^+(y\text{C})$ at 0 K (in kJ/mol), where xC, yC = C, 5MeC, 1MeC, and 15dMeC, plotted versus TCID measured values including ZPE and BSSE corrections. The results for protonated homo- and heteronucleobase pairs are shown in parts a and b, respectively. BPEs of the $(5\text{MeC})\text{H}^+(5\text{MeC})$ and $(5\text{MeC})\text{H}^+(\text{C})$ protonated nucleobase pairs are taken from refs 26 and 25, respectively. The black solid diagonal line indicates the values for which the calculated and measured BPEs are equal.

Comparison of Measured and Calculated N3 PAs. The TCID measured and B3LYP/def2-TZVPPD calculated relative N3 PAs are listed in Table 2 and compared pictorially in Figure 5a. As can be seen in Figure 5a, the B3LYP/def2-TZVPPD level of theory accurately estimates the relative N3 PAs of C versus 1MeC and 15dMeC, whereas the theory underestimates the relative N3 PAs of C versus 5MeC. The MAD between theory and experiment for the relative N3 PAs is 2.8 ± 2.6 kJ/mol, which just slightly exceeds the AEU in these values, 2.5 ± 0.8 kJ/mol. The B3LYP/def2-TZVPPD predicted N3 PA of 1MeC is slighter greater than that of 5MeC, whereas TCID experiments using cytosine as the reference base find that the PAs of these two nucleobases differ by less than the experimental error in either measurement. The TCID experiment examining the $(1\text{Me-}d_3\text{-C})\text{H}^+(5\text{MeC})$ protonated nucleobase pair helps solve this problem and suggests that the N3 PA of 1MeC slightly exceeds that of 5MeC, consistent with the B3LYP/def2-TZVPPD results.

Absolute N3 PAs at 298 K of the three methylated cytosines are derived from a comprehensive maximum likelihood statistical analysis of the TCID results for all five protonated heteronucleobase pairs examined here as well as the $(5\text{MeC})\text{H}^+(\text{C})$ protonated heteronucleobase pair previously investigated[25] and the PAs of C and 1MeC obtained from the NIST Chemistry Webbook and ref 24. The results of the maximum likelihood statistical analyses are summarized in Table 3, whereas the absolute N3 PAs and the 2-D cross sections of the uncertainties in the PAs determined are shown in Figure 6. Each 2-D cross section plot in Figure 6 illustrates the correlation in the MLE of the N3 PAs and the corresponding uncertainties for each nucleobase pair. The

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

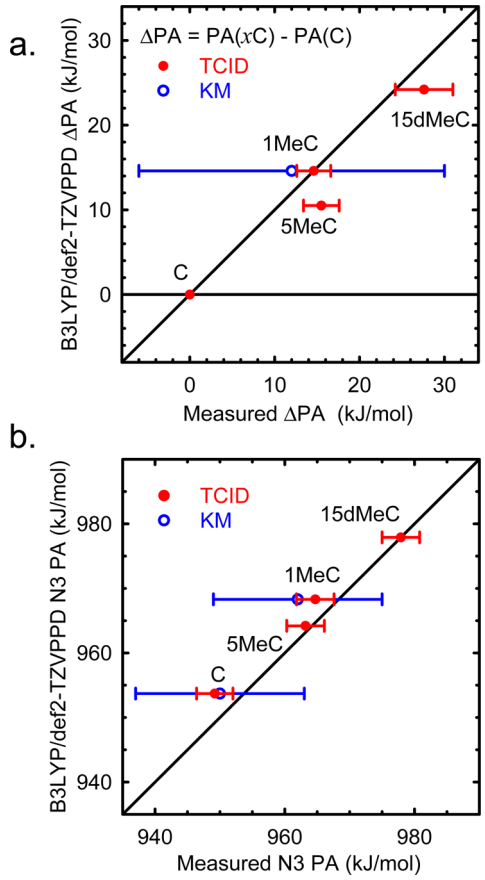


Figure 5. B3LYP/def2-TZVPPD calculated relative and absolute N3 PAs plotted versus TCID results at 298 K (in kJ/mol), parts a and b, respectively. Values determined here are indicated with closed symbols, whereas the kinetic method results of Liu et al.[24] are indicated with open symbols.

Table 3. Absolute N3 PAs of Cytosine and Methylated Cytosines at 298 K in kJ/mol

	N3 PA			
nucleobase	TCID[a]	KM	NIST	calc[b]
C	949.2 (2.8)	950 (13)[c]	949.9 (8.0)[d]	953.7
5MeC	963.2 (2.9)			964.2
1MeC	964.7 (2.9)	962 (13)[c]		968.3
15dMeC	977.9 (2.9)			977.9

[a]Present results, uncertainties are listed in parentheses. [b]Calculated at the B3LYP/def2-TZVPPD level of theory and including ZPE corrections. [c]Values taken from ref 24. [d]Value taken from refs 22 and 23.

absolute N3 PAs determined from the MLE are compared with B3LYP/def2-TZVPPD calculated values in Figure 5b. The MAD between theory and experiment for the absolute N3 PAs is 2.3 ± 2.1 kJ/mol, smaller than the AEU in these values, 2.9 ± 0.1 kJ/mol. Liu et al. previously reported PAs of C (950 ± 13 kJ/mol) and 1MeC (962 ± 13 kJ/mol) using the extended Cooks kinetic method.[24] These values are consistent with the PA listed in the NIST Chemistry Webbook for C, 949.9 ± 8 kJ/mol, and those for C and 1MeC measured here, 949.2 ± 2.8 and 963.2 ± 2.9 kJ/mol. However, the additional data provided by the present measurements combined with the MLE significantly reduce the uncertainty in the absolute N3 PAs from ~8 to ~3 kJ/mol.

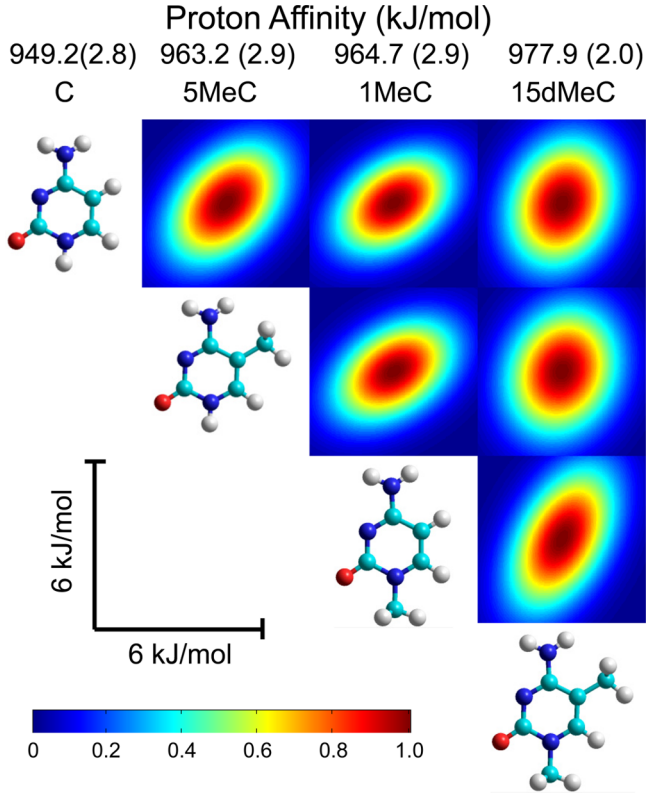


Figure 6. Maximal likelihood estimate of the proton affinities (PAs) of C, 5MeC, 1MeC, and 15dMeC and plots of the six 2-D cross sections of the uncertainty in the determination for each nucleobase pair. Values determined using the TCID relative N3 PAs listed in Table 2 and absolute PAs of C and 1MeC taken from refs 22 and 23. The color bar at the lower left maps likelihoods, or probability densities (scaled to 1 at the peak), to the colors used in the 2D cross section images.

Influence of Methylation on the N3 PA. As shown in Figure 5b, methylation at either the 1- or 5-position leads to an increase in the N3 PA of cytosine as expected on the basis of the electronic properties of the methyl substituent. The electron-donating methyl substituent increases the electron density within the aromatic ring, and hence results in stabilization of the positive charge associated with the excess proton. 1-Methylation of cytosine produces a slightly larger effect on the N3 PA than 5-methylation. This is because the methyl substituent donates more electron density to the aromatic ring when bound to the electronegative N1 atom than the C5 atom. The TCID measured N3 PAs of cytosine and the methylated cytosines follow the order 15dMeC > 1MeC > 5MeC > C, consistent with the trend predicted from apparent thresholds of the protonated heteronucleobase pairs. The correlation between the polarizabilities of *x*C and the TCID measured absolute N3 PAs is illustrated in Figure 7a. A linear regression fit through all of the data is also shown in the figure. The absolute N3 PAs of *x*C clearly increase with increasing polarizability of *x*C.

Influence of Methylation on the BPEs. The TCID measured and calculated BPEs at 0 K of the seven protonated (*x*C)H$^+$(*y*C) nucleobase pairs measured here and three previously reported are listed in Table 1. The measured BPEs of the permethylated protonated nucleobase pairs of cytosine are greater than that of the (C)H$^+$(C) protonated nucleobase pair;[26] however, the increase in the BPE upon 1-methylation is smaller than the uncertainties in these measurements. Thus,

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

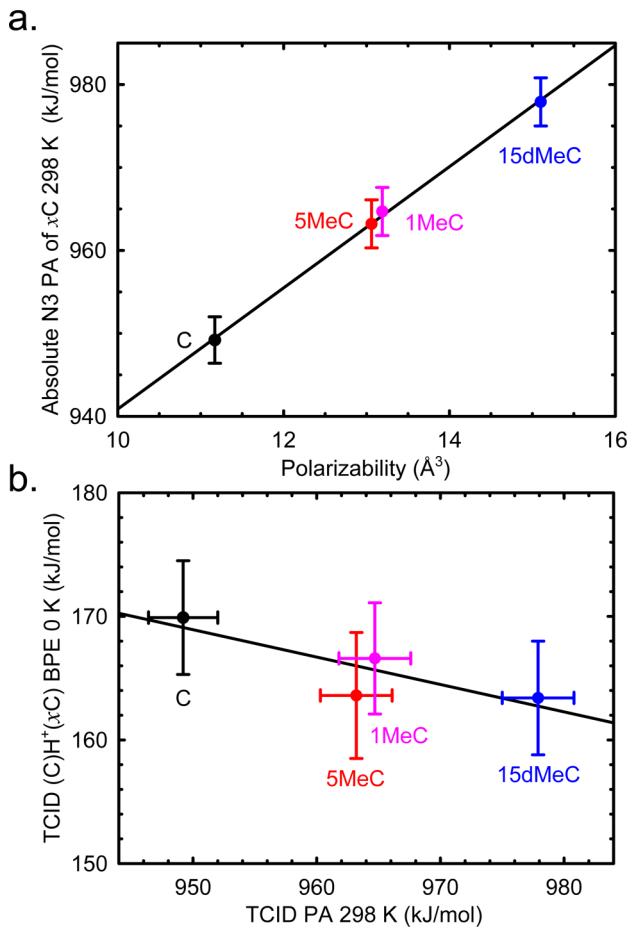


Figure 7. TCID measured absolute N3 PA of *x*C at 298 K (in kJ/mol) versus calculated polarizability volumes of *x*C, where *x*C = C, 5MeC, 1MeC, and 15dMeC. The black line is a linear regression fit to the data (part a). TCID measured BPEs of $(C)H^+(xC)$ at 0 K (in kJ/mol) versus measured relative N3 PAs of *x*C and C, where *x*C = C, 5MeC, 1MeC, and 15dMeC (part b).

permethylation is expected to provide additional albeit minor stabilization of DNA *i*-motif conformations. However, theory suggests that 1-permethylation of cytosine leads to a small decrease in the BPE. In the case of protonated heteronucleobase pairs, all single modifications result in a decrease in the BPE,[26] indicating that all modifications of only one of the nucleobases of the protonated nucleobase pair slightly weaken the base-pairing interactions and lead to very minor destabilization of DNA *i*-motif conformations. The correlation between the measured BPEs and relative N3 PAs of C and *x*C is shown in Figure 7b. Clearly, the BPEs of the protonated heteronucleobase pairs are inversely correlated with the absolute difference in the relative N3 PAs of C and *x*C due to the unequal sharing of the excess proton in these protonated heteronucleobase pairs. This behavior suggests that the effect of modifications on the BPEs should directly correlate with their effect on the N3 PA.

Implications for the Stability of DNA *i*-Motif Conformations. DNA *i*-motif conformations are mainly stabilized by the hydrogen-bonding interactions in the protonated nucleobase pairs of cytosine. Previous TCID studies of protonated homonucleobase pairs of cytosine and modified cytosines found that 5-permethylation of cytosine increases the BPE of the protonated nucleobase pair, and would therefore tend to stabilize DNA *i*-motif conformations.[26] This result also suggests that hypermethylation of $(CCG)_n{\cdot}(CGG)_n$ trinucleotide repeats, the cause of Fragile X syndrome, occurs to further stabilize DNA *i*-motif conformations. In contrast, the current TCID results indicate that 1-permethylation of cytosine residues has almost no effect on the strength of the base-pairing interactions, and thus should have little or no effect on the stability of DNA *i*-motif conformations, whereas theory suggests that 1-methylation of both cytosine residues leads to a decrease in BPE and thus should destabilize DNA *i*-motif conformations. The BPE of the $(15dMeC)H^+(15dMeC)$ protonated homonucleobase pair is slighter higher than that of the $(C)H^+(C)$ protonated homonucleobase pair, indicating that permethylation of both cytosine residues at the N1 and C5 positions slightly increases the BPE and hence should slightly stabilize DNA *i*-motif conformations. By extension, these results also suggest that the BPE of the protonated nucleobase pair of 2′-deoxycytidine (dCyd) should be approximately equal to that of C. However, polarizability effects also play a role such that this conclusion must be experimentally (and theoretically) verified and is the subject of future investigations. Methylation of a single cytosine residue at any position weakens the base-pairing interactions. However, the BPEs of all $(xC)H^+(yC)$ protonated heteronucleobase pairs significantly exceed those of canonical G·C and neutral C·C base pairs, suggesting that DNA *i*-motif conformations are still favored over Watson−Crick base pairing upon modification. Thus, although methylation of only one of the cytosine residues at the N1, C5 or N1 and C5 positions tends to decrease the strength of the base-pairing interactions in the protonated nucleobase pairs of cytosine, the effects are sufficiently small that *i*-motif conformations should be stable upon such modifications. Only in the case of hypermethylation at the C5 or N1 and C5 positions are the BPEs increased, and this leads to the diseased state associated with the Fragile X syndrome.[8,9] Even though the change in the BPE induced by methylation is not large for a single protonated nucleobase pair, the accumulated effect can be dramatic in diseased state trinucleotide repeats associated with the Fragile X syndrome where more than 230 trinucleotides and hundreds of methylated protonated nucleobase pairs may be present. Because methylation at different positions may lead to an increase or decrease in the BPE, the influence of methylation will be seen in the number of trinucleotide repeats required to cause structural conversion from canonical Watson-Crick base-pairing to DNA *i*-motif conformations.

To fully elucidate the effects of modifications on the stability of DNA *i*-motif conformations, other factors that may stabilize/destabilize these noncanonical DNA structures including π-stacking interactions, steric effects associated with the sugar moiety, and the folding of the nucleic acid strands must also be considered. Further studies to characterize the evolution of these base-pairing interactions as the size of the model systems increases to protonated nucleobase pairs of the analogous 2′-deoxycytidine nucleosides, nucleotides, and $(CCG)_n$ trinucleotide repeats that are associated with Fragile X syndrome are currently underway. Present studies have determined that the B3LYP level of theory accurately describes the base-pairing interactions in the protonated nucleobase pairs investigated here and therefore may be suitable for examining larger and more biologically relevant model systems. Present studies have also shown that the effects of the bulky electron-donating methyl group differ depending on the position of substitution. 1-Methylation, a mimic for the sugar moiety, has almost no

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

effect on the BPE of the protonated nucleobase pair and hence should exert almost no effect on the stability of DNA *i*-motif conformations. In contrast, epigenetic 5-methylation of cytosine exerts a larger more stabilizing influence and would tend to stabilize DNA *i*-motif conformations, thereby requiring fewer trinucleotide repeats for structural conversion of canonical base paired DNA to *i*-motif conformations. Information obtained in this work including the structures, TCID behavior, and trends in the BPEs of these protonated nucleobase pairs should also facilitate CID experiments and data analyses for studies of these larger and more biologically relevant model systems.

CONCLUSIONS

Cytosine methylation, one of the most common epigenetic modifications, can regulate gene expression by altering the structure and stability of DNA or DNA−protein interactions. In order to elucidate the effects of cytosine methylation on the base-pairing interactions responsible for stabilizing DNA *i*-motif conformations and the proton affinities of the nucleobases, the energy-resolved CID behavior of seven protonated nucleobase pairs, (1MeC)H^+(1MeC), (15dMeC)H^+(15dMeC), (1MeC)-H^+(C), (15dMeC)H^+(C), (1MeC)H^+(5MeC), (15dMeC)-H^+(5MeC), and (15dMeC)H^+(1MeC), is examined using TCID techniques. The only dissociation pathway observed for the protonated homonucleobase pairs corresponds to cleavage of the three hydrogen-bonding interactions that stabilize these species, resulting in elimination of the neutral nucleobase. For the protonated heteronucleobase pairs, two dissociation pathways involving production of the two protonated nucleobases occur in parallel, and compete with one another. PESs and TS were calculated to determine the tautomerization barriers for dissociation of the ground-state conformations of the protonated nucleobase pairs to produce O2-protonated nucleobase products ($\mathbf{I}^+$). Theoretical results confirm that the tautomerization barriers are greater than the dissociation energy for production of the N3-protonated nucleobases ($\mathbf{II}^+$), indicating that tautomerization cannot occur upon dissociation at or near threshold energies. The BPEs correspond to the thresholds for CID reactions that produce the N3-protonated nucleobases and are determined using a loose PSL TS model that carefully considers the effects of the kinetic and internal energy distributions of the (xC)H^+(yC) and Xe reactants, multiple collisions with Xe, and the lifetime of the activated (xC)H^+(yC) protonated nucleobase pair. Competitive threshold analyses of the two dissociation pathways that occur in parallel for the protonated heteronucleobase pairs provide the relative N3 PAs of cytosine and the methylated cytosine. Theoretical calculations were performed at the B3LYP and MP2(full) levels of theory using the 6-311+G(2d,2p) and def2-TZVPPD basis sets to obtain theoretical estimates for the BPEs of the (xC)H^+(yC) complexes and the N3 PAs of xC. Excellent agreement is found between experimental and B3LYP calculated BPEs, whereas the MP2(full) theory systematically underestimates the BPEs, even when BSSE corrections are not included in the computed BPEs. Excellent agreement is also achieved for the measured and B3LYP/def2-TZVPPD calculated relative and absolute N3 PAs of cytosine and methylated cytosines. These results suggest that the B3LYP/def2-TZVPPD level of theory is suitable for providing reliable energetics for related systems that bind via multiple hydrogen bonds. Methylation clearly influences the base-pairing interactions in the protonated nucleobase pairs. In the case of homodimers, 5-hypermethylation is found to increase the BPE,[26] whereas 1-hypermethylation is found to exert almost no effect on the BPE. Hence, 1,5-dimethylation of both cytosines results in an intermediate increase in the BPE. These results suggest that DNA *i*-motif conformations should be stabilized under 5-hypermethylation conditions. In the case of the heterodimers, methylation of a single cytosine at the N1, C5 or N1 and C5 positions weakens the BPE, and therefore would tend to destabilize DNA *i*-motif conformations. The magnitude of the decrease in the BPE is found to directly correlate with the difference in the N3 PA induced by methylation. However, the BPEs of all methylated protonated nucleobase pairs are still much greater than those of canonical G·C and neutral C·C base pairs, suggesting that the effects of methylation are not sufficient to destroy DNA *i*-motif conformations but may alter the number of trinucleotide repeats required to induce structural conversion from canonical base-pairing to DNA *i*-motif conformations. Methylation also affects the N3 PA of cytosine. In contrast to its effects on the BPEs, methylation of cytosine increases the N3 PA regardless of the position of substitution. The N3 PAs of cytosine and the methylated cytosines follow the order 15dMeC (979.9 ± 2.9 kJ/mol) > 1MeC (964.7 ± 2.9 kJ/mol) > 5MeC (963.2 ± 2.9 kJ/mol) > C (949.9 ± 2.8 kJ/mol), indicating that N1-methylation has a greater influence on the N3 PA than C5-methylation, and the effects of N1, C5-dimethylation are roughly additive.

ASSOCIATED CONTENT

Supporting Information

A description of the procedures employed for data handling, thermochemical analysis, and thermal corrections is provided. Tables of B3LYP/6-31G* vibrational frequencies, average vibrational energies at 298 K, and rotational constants of the neutral and protonated nucleobases as well as the protonated nucleobase pairs, B3LYP/def-TZVPPD relative stabilities of the three low-energy tautomeric conformations of xC and H^+(xC), relative energies of species along the PES for dissociation of seven (xC)H^+(yC) dimers, (1MeC)H^+(1MeC), (15dMeC)H^+(15dMeC), (1MeC)H^+(C), (15dMeC)H^+(C), (1MeC)H^+(5MeC), (15dMeC)H^+(5MeC), and (15dMeC)-H^+(1MeC), fitting parameters of eqs S1 and S2, and enthalpies and free energies of the base-pairing interactions at 298 K. Figures showing cross sections for CID of five (xC)H^+(yC) dimers, the low-energy stable tautomeric conformations of xC and H^+(xC), PESs determined at the B3LYP/def-TZVPPD level of theory for dissociation of five (xC)H^+(yC) dimers, and thermochemical analyses of zero-pressure extrapolated CID cross sections of the (xC)H^+(yC) dimers, (15dMeC)-H^+(15dMeC), (15dMeC)H^+(C), (1MeC)H^+(5MeC), (15dMeC)H^+(5MeC), and (15dMeC)H^+(1MeC). Figures comparing TCID measured and B3LYP and MP2(full) calculated BPEs using the def2-TZVPPD and 6-311+G(2d,2p) basis sets. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support for this work was provided by the National Science Foundation, Grant CHE-1409420. We thank Wayne

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

State University C&IT for computer time and support. B.Y. gratefully acknowledges support from Wayne State University Thomas C. Rumble Graduate and Summer Dissertation Fellowships. Thanks are also due to Dr. T. H. Morton for providing the 1-methylated cytosines.

REFERENCES

(1) Berger, S. L.; Kouzarides, T.; Shiekhattar, R.; Shilatifard, A. An Operational Definition of Epigenetics. *Genes Dev.* **2009**, *23*, 781−783.
(2) Bird, A. DNA Methylation Patterns and Epigenetic Memory. *Genes Dev.* **2002**, *16*, 6−21.
(3) Cooper, D. N.; Youssoufian, H. The CpG Dinucleotide and Human Genetic Disease. *Hum. Genet.* **1988**, *78*, 151−155.
(4) Rideout, W. M., III; Coetzee, G. A.; Olumi, A. F.; Jones, P. A. 5-Methylcytosine as an Endogenous Mutagen in the Human LDL Receptor and p53 Genes. *Science* **1990**, *249*, 1288−1290.
(5) McLenan, Y.; Polussa, J.; Tassone, F.; Hagerman, R. Fragile X Syndrome. *Curr. Genomics* **2011**, *12*, 216−224.
(6) Garber, K. B.; Visootsak, J.; Warren, S. T. Fragile X Syndrome. *Eur. J. Hum. Genet.* **2008**, *16*, 666−672.
(7) Turner, G.; Webb, T.; Wake, S.; Robinson, H. Prevalence of Fragile X Syndrome. *Am. J. Med. Genet.* **1996**, *64*, 196−197.
(8) Sutcliffe, J. S.; Nelson, D. L.; Zhang, F.; Pieretti, M.; Caskey, C. T.; Saxe, D.; Warren, W. T. DNA Methylation Represses *FMR-1* Transcription in Fragile X Syndrome. *Hum. Mol. Genet.* **1992**, *1*, 397−400.
(9) Oberlé, I.; Rousseau, F.; Heitz, D.; Kretz, C.; Devys, D.; Hanauer, A.; Boué, J.; Bertheas, M. F.; Mandel, J. L. Instability of a 550-Base Pair DNA Segment and Abnormal Methylation in Fragile X Syndrome. *Science* **1991**, *252*, 1097−1102.
(10) Pieretti, M.; Zhang, F.; Fu, Y.-H.; Warren, S. T.; Oostra, B. A.; Caskey, C. T.; Nelson, D. L. Absence of Expression of the *FMR-1* Gene in Fragile X Syndrome. *Cell* **1991**, *66*, 817−822.
(11) Darlow, J. M.; Leach, D. R. F. Secondary Structure in d(CGG) and d(CCG) Repeat Tracts. *J. Mol. Biol.* **1998**, *275*, 3−16.
(12) Mitas, M. Trinucleotide Repeats Associated with Human Disease. *Nucleic Acids Res.* **1997**, *25*, 2245−2253.
(13) Gehring, K.; Leroy, J.-L.; Guéron, M. A. Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs. *Nature* **1993**, *363*, 561−565.
(14) Fotjík, P.; Vorlícková, M. The Fragile X Chromosome (GCC) Repeat Folds into a DNA Tetraplex at Neutral pH. *Nucleic Acids Res.* **2001**, *29*, 4684−4690.
(15) Rosu, F.; Gabelica, V.; Joly, L.; Grégoire, G.; Pauw, E. D. Zwitterionic *i*-Motif Structures are Preserved in DNA Negatively Charged Ions Produced by Electrospray Mass Spectrometry. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13448−13454.
(16) Oomens, J.; Moehlig, A. R.; Morton, T. H. 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.
(17) Yang, B.; Wu, R. R.; Berden, G.; Oomens, J.; Rodgers, M. T. 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*, 14191−11006.
(18) Bhavsar-Jog, Y. P.; Dornshuld, E. V.; Brooks, T. A.; Tschumper, G. S.; Wadkins, R. M. Epigenetic Modification, Dehydration, and Molecular Crowding Effects on the Thermodynamics of i-Motif Structure Formation from C-Rich DNA. *Biochemistry* **2014**, *53*, 1586−1594.
(19) Little, C. D.; Nau, M. M.; Carney, D. N.; Gazdar, A. F.; Minna, J. D. Amplification and Expression of the c-myc Oncogene in Human Lung Cancer Cell Lines. *Nature* **1983**, *306*, 194−196.
(20) Munzel, P.; Marx, D.; Kochel, H.; Schauer, A.; Bock, K. W. Genomic Alterations of the c-myc Protooncogene in Relation to the Overexpression of c-erbB2 and Ki-67 in Human Breast and Cervix Carcinomas. *J. Cancer Res. Clin. Oncol.* **1991**, *117*, 603−607.
(21) Clark, H. M.; Yano, T.; Otsuki, T.; Jaffe, E. S.; Shibata, D.; Raffeld, M. Mutations in the Coding Region of c-myc in AIDS-Associated and other Aggressive Lymphomas. *Cancer Res.* **1994**, *54*, 3383−3386.
(22) Hunter, E. P.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413−656.
(23) NIST Chemistry Webbook, http://webbook.nist.gov/chemistry/.
(24) Liu, M.; Li, T.; Amegayibor, S.; Cardoso, D. S.; Fu, Y.; Lee, J. K. Gas-Phase Thermochemical Properties of Pyrimidine Nucleobases. *J. Org. Chem.* **2008**, *73*, 9283−9291.
(25) Yang, B.; Rodgers, M. T. Base-Pairing 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.
(26) Yang, B.; Wu, R. R.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Homodimers Determined by Guided Ion Beam Tandem Mass Spectrometry: Application to Cytosine and 5-Substituted Cytosines. *Anal. Chem.* **2013**, *85*, 11000−11006.
(27) Rodgers, M. T. Substituent Effects in the Binding of Alkali Metal Ions to Pyridines Studied by Threshold Collision-Induced Dissociation and Ab Initio Theory: The Methylpyridines. *J. Phys. Chem. A* **2001**, *105*, 2374−2383.
(28) Teloy, E.; Gerlich, D. Integral Cross Sections for Ion−Molecule Reactions I. The Guided Ion Beam Technique. *Chem. Phys.* **1974**, *4*, 417−427.
(29) Gerlich, D. Inhomogeneous RF Fields: A Versatile Tool for the Study of Processes with Slow Ions. Diplomarbeit, University of Freiburg, Federal Republic of Germany, 1971.
(30) Gerlich, D. In *State-Selected and State-to-State Ion−Molecule Reaction Dynamics, Part I, Experiment*; Ng, C.-Y., Baer, M., Eds.; Advances in Chemical Physics; J. Wiley: New York, 1992; Vol. *82*, pp 1−176.
(31) Dalleska, N. F.; Honma, K.; Armentrout, P. B. Stepwise Solvation Enthalpies of Protonated Water Clusters: Collision-Induced Dissociation as an Alternative to Equilibrium Studies. *J. Am. Chem. Soc.* **1993**, *115*, 12125−12131.
(32) Aristov, N.; Armentrout, P. B. Collision-Induced Dissociation of Vanadium Monoxide Ion. *J. Phys. Chem.* **1986**, *90*, 5135−5140.
(33) Hales, D. A.; Armentrout, P. B. Effect of Internal Excitation on the Collision-Induced Dissociation and Reactivity of Co_2^+. *J. Cluster Sci.* **1990**, *1*, 127−142.
(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009. See the Supporting Information for the full reference.
(35) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(36) Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17*, 1571−1586.
(37) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, L.; Windus, T. L. Basis Set Exchange: A Community Database for Computational Science. *J. Chem. Inf. Model.* **2007**, *47*, 1045−1052.
(38) Smith, S. M.; Markevitch, A. N.; Romanov, D. A.; Li, X.; Levis, R. J.; Schlegel, H. B. Static and Dynamic Polarizabilities of Conjugated Molecules and Their Cations. *J. Phys. Chem. A* **2004**, *108*, 11063−11072.
(39) Peng, C.; Schlegel, H. B. Combining Synchronous Transit and Quasi-Newton Methods for Finding Transition States. *Isr. J. Chem.* **1993**, *33*, 449−454.
(40) *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Foresman, J. B., Frisch, Æ., Eds.; Gaussian, Inc.: Pittsburgh, PA, 1996; p 64.

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868

(41) Boys, S. F.; Bernardi, R. Calculation of Small Molecular Interactions by Differences of Separate Total Energies - Some Procedures with Reduced Errors. *Mol. Phys.* **1979**, *19*, 553−566.
(42) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. State of the Art in Counterpoise Theory. *Chem. Rev.* **1994**, *94*, 1873−1885.
(43) Muntean, F.; Armentrout, P. B. Guided Ion Beam Study of Collision-Induced Dissociation Dynamics: Integral and Differential Cross Sections. *J. Chem. Phys.* **2001**, *115*, 1213−1228.
(44) Beyer, T. S.; Swinehart, D. S. Number of Multiply-Restricted Partitions [A1]. *Commun. ACM* **1973**, *16*, 379.
(45) Stein, S. E.; Rabinovitch, B. S. Accurate Evaluation of Internal Energy Level Sums and Densities Including Anharmonic Oscillators and Hindered Rotors. *J. Chem. Phys.* **1973**, *58*, 2438−2445.
(46) Stein, S. E.; Rabinovitch, B. S. On the Use of Exact State Counting Methods in RRKM Rate Calculations. *Chem. Phys. Lett.* **1977**, *49*, 183−188.
(47) Khan, F. A.; Clemmer, D. E.; Schultz, R. H.; Armentrout, P. B. Sequential Bond Energies of $Cr(CO)_x^+$, $x = 1-6$. *J. Phys. Chem.* **1993**, *97*, 7978−7987.
(48) Rodgers, M. T.; Ervin, K. M.; Armentrout, P. B. Statistical Modeling of Collision-Induced Dissociation Thresholds. *J. Chem. Phys.* **1997**, *106*, 4499−4508.
(49) Chesnavich, W. J.; Bowers, M. T. Theory of Translationally Driven Reactions. *J. Phys. Chem.* **1979**, *83*, 900−905.
(50) Chen, Y.; Rodgers, M. T. Structural and Energetic Effects in the Molecular Recognition of Acetylated Amino Acids by 18-Crown-6. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 2020−2030.
(51) Austin, C. A.; Chen, Y.; Rodgers, M. T. Alkali Metal Cation-Cyclen Complexes: Effects of Alkali Metal Cation Size on the Structure and Binding Energy. *Int. J. Mass Spectrom.* **2012**, *330*, 27−34.
(52) Armentrout, P. B.; Yang, B.; Rodgers, M. T. Metal Cation Dependence of Interactions with Amino Acids: Bond Energies of Rb^+ and Cs^+ to Met, Phe, Tyr, and Trp. *J. Phys. Chem. B* **2013**, *117*, 3771−3781.
(53) Nose, H.; Chen, Y.; Rodgers, M. T. Energy-Resolved Collision-Induced Dissociation Studies of 1,10-Phenanthroline Complexes of the Late First-Row Divalent Transition-Metal Cations: Determination of the Third Sequential Binding Energies. *J. Phys. Chem. A* **2013**, *117*, 4316−4330.
(54) Nose, H.; Rodgers, M. T. Energy-Resolved Collision-Induced Dissociation Studies of 2,2′-Bipyridine Complexes of the Late First-Row Divalent Transition-Metal Cations: Determination of the Third Sequential Binding Energies. *ChemPlusChem* **2013**, *78*, 1109−1123.
(55) Sim, F.; St-Amant, A.; Pápai, I.; Salahub, D. R. Gaussian Density Functional Calculations on Hydrogen-Bonded Systems. *J. Am. Chem. Soc.* **1992**, *114*, 4391−4400.
(56) Laasonen, K.; Parrinello, M.; Car, R.; Lee, C.; Vanderbilt, D. Structure of Small Water Clusters Using Gradient-Corrected Density Functional Theory. *Chem. Phys. Lett.* **1993**, *207*, 208−213.
(57) Kim, K.; Jordan, K. D. Comparison of Density Functional and MP2 Calculations on the Water Monomer and Dimer. *J. Phys. Chem.* **1994**, *98*, 10089−10094.
(58) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. Properties of Hydrogen-Bonded Complexes Obtained from the B3LYP Functional with 6-31G(d,p) and 6-31+G(d, p) Basis Sets: Comparison with MP2/6-31+G(d,p) Results and Experimental Data. *J. Phys. Chem.* **1995**, *99*, 10705−10707.
(59) Novoa, J. J.; Sosa, C. Evaluation of the Density Functional Approximation on the Computation of Hydrogen Bond Interactions. *J. Phys. Chem.* **1995**, *99*, 15837−15845.
(60) Hobza, P.; Šponer, J.; Reschel, T. Density Functional Theory and Molecular Clusters. *J. Comput. Chem.* **1995**, *16*, 1315−1325.
(61) Suhai, S. Density Functional Studies of the Hydrogen-Bonded Network in an Infinite Water Polymer. *J. Phys. Chem.* **1995**, *99*, 1172−1181.
(62) Valdés, H.; Klusák, V.; Pitoňák, M.; Exner, O.; Starý, I.; Hobza, P.; Rulisek, L. Evaluation of the Intramolecular Basis Set Superposition Error in the Calculations of Large Molecules: [*n*]Helicenes and Phe-Gly-Phe Tripeptide. *J. Comput. Chem.* **2008**, *29*, 861−870.

DOI: 10.1021/acs.jpcb.5b00035
J. Phys. Chem. B 2015, 119, 1857−1868