Influence of 5-Halogenation on the Base Pairing Energies of Protonated Cytidine Nucleoside Analogue Base Pairs: Implications for the Stabilities of Synthetic *i*-Motif Structures for DNA Nanotechnology Applications

M. T. Rodgers*, Yakubu S. Seidu, and E. Israel
Department of Chemistry, Wayne State University, Detroit, MI, 48202, USA

ABSTRACT:

DNA nanotechnology has been employed to develop devices based on *i*-motif structures. The protonated cytosine-cytosine base pairs that stabilize *i*-motif conformations are favored under slightly acidic conditions. This unique property has enabled development of the first DNA molecular motor driven by pH changes. The ability to alter the stability and pH transition range of such DNA molecular motors is desirable. Understanding how *i*-motif structures are influenced by modifications, and which modifications enhance stability and/or affect the pH characteristics, are therefore of great interest. Here, the influence of 5-halogenation of the cytosine nucleobases on the base pairing of protonated cytidine nucleoside analogue base pairs are examined using complimentary threshold collision-induced dissociation techniques and computational methods. The nucleoside analogues examined here include the 5-halogenated forms of the canonical DNA and RNA cytidine nucleosides. Comparisons among these systems and to the analogous canonical base pairs previously examined enable the influence of 5-halogenation and the 2′-hydroxy substituent on the base pairing to be elucidated. 5-Halogenation of the cytosine nucleobases is found to enhance the strength of base pairing of DNA base pairs, and generally weakens the base pairing for RNA base pairs. Trends in the strength of base pairing indicate that both inductive and polarizability effects influence the strength of base pairing. Overall, present results suggest that 5-halogenation, and in particular, 5-fluorination and 5-iodination provide effective means of stabilizing DNA *i*-motif conformations for applications in nanotechnology, whereas only 5-iodination is effective for stabilizing RNA *i*-motif conformations but the enhancement in stability is less significant.

INTRODUCTION

The desire to design and manufacture artificial nucleic acid structures for various technological applications has driven the field of DNA nanotechnology since it was first conceptualized by Seeman in the early 1980s.[1,2] The high specificity of base pairing that enables nucleic acids to preserve and translate genetic information in biological systems,[3] also enables them to be employed as high specificity engineering materials for nanotechnology. Rational design of DNA sequences with complementary and noncomplementary regions has been employed to selectively assemble a variety of two- and three-dimensional structures and arbitrary shapes.[4-7] Assembly of complementary regions via tile-by-tile assembly[8] or folding via DNA origami approaches[9-12] has been based primarily on canonical Watson-Crick base-paired DNA duplexes.[3] The transition from one-pot solution synthesis to approaches based on solid-support methodologies[13] has also expanded the topologies of the structures that can be fabricated. The variety of structures than can be produced continues to grow as new assembly approaches are imagined and realized. Approaches that enable dynamic control of structural transformations and assembly are of particular interest as DNA nanostructures that can be subjected to tuning of their structures and materials properties in response to controlled external stimuli can be used to drive nanomechanical devices[14-,,,,19] and thus find useful applications in a variety of areas including: nanoelectronics, nanomedicine, nanophotonics, nanorobotics, nanosensors, and biocomputing.[20-25]

Noncanonical structures such as G-quadruplexes and the *i*-motif have been of specific interest as reconfigurable elements for DNA nanotechnology. In particular, DNA nanotechnology has been employed to develop artificial molecular motors and devices based on *i*-motif structures.[26] Nucleic acid sequences rich in cytosine (Cyt) have the potential to fold into an *i*-motif in which the four-stranded core is held together by very strong protonated cytosine-cytosine $(Cyt)H^+(Cyt)$ base-pairing interactions.[27] Differences in the nature of the noncovalent interactions that stabilize noncanonical *i*-motif structures vs. canonical DNA duplexes also alter the conditions needed for their self-assembly, which in turn provides the ability to manipulate DNA nanostructures in ways that would not be possible based solely on DNA duplex assembly.

Specifically, the ability to assemble *i*-motif structures intramolecularly from four different Cyt tracts within a single strand, or intermolecularly from the association of two, three or four separate strands provides unique opportunities for dynamic control of structural transformations. Formation of an *i*-motif requires N3-protonation of the Cyt residues in the core region, which has been observed under acidic conditions[28] and at physiological pH.[29-30] This unique property has enabled the development of the first DNA molecular motor driven by pH changes.[31] The ability to alter the properties of this DNA molecular motor, e.g., the pH transition range, is desirable such that understanding how the stability of *i*-motif structures are influenced by modifications, and in particular, which modifications can be employed to enhance stability and/or effect the pH characteristics of the *i*-motif are of great interest. In order to expand the applicability of *i*-motif conformations, the effects of nucleobase and sugar modifications on *i*-motif stability have been examined.[32] Notably, 5-methylation of the cytosine residues has been shown to slightly enhance the stability and increase the pK_a of the *i*-motif.[33,34]

The impact of 5-modification of Cyt residues on the structure and function of nucleic acids has long been appreciated and studied.[35,36] However, information on how such modifications influence hydrogen-bonding interactions among noncanonical DNA structures remains limited.[33,37-,,40] Given the variety of roles that *i*-motif conformations are beginning to play in applications for DNA nanotechnology,[40] a greater appreciation for how other modifications impact the structure and stabilities of DNA (and RNA) *i*-motif architectures is desirable. We began our investigations with the simplest model for the core stabilizing interactions of *i*-motif structures, the protonated cytosine base pair, (Cyt)H^+(Cyt). Using complementary threshold collision-induced dissociation (TCID) techniques and computational methods, the base-pairing energy (BPE) of the (Cyt)H^+(Cyt) base pair was measured as 169.9 ± 4.6 kJ/mol in excellent agreement with the theoretically predicted value of 168.9 kJ/mol.[41] Notably, the BPE of the (Cyt)H^+(Cyt) base pair greatly exceeds those of the canonical Watson-Crick guanine-cytosine (Gua)·(Cyt) and noncanonical neutral cytosine-cytosine (Cyt)·(Cyt) base pairs, with theoretically predicted BPEs of 96.6 kJ/mol and 68.0 kJ/mol, respectively. The much stronger base-pairing interactions in

(Cyt)H^+(Cyt) base pairs are important contributors to the stability of *i*-motif conformations. This initial study as well as a series of parallel follow-on investigations introduced increasing complexity to the (Cyt)H^+(Cyt) base-pair model for the *i*-motif by examining the effects of a series of modifications of the Cyt nucleobase.[41-,,44] Consistent with previous findings that 5-methylation stabilizes noncanonical protonated cytosine nucleobase analogue base pairs[45] and *i*-motif conformations,[46] the BPE was found to increase upon 5-methylation. In contrast, the BPE was found to decrease upon 5-fluorination, 5-bromination, and 5-iodination. Again these results are consistent with previous findings that 5-fluorination destabilizes noncanonical protonated cytosine nucleobase analogue base pairs.[45] However, the BPEs of the 5-halogenated protonated cytosine-cytosine base pairs also significantly exceed the BPEs of the (Gua)·(Cyt) and (Cyt)·(Cyt) base pairs indicating that *i*-motif conformations should be stable to 5-halogenation. The model systems investigated were extended to nucleosides with the examination of protonated nucleoside base pairs of the canonical DNA and RNA cytidine nucleosides, 2′-deoxycytidine (dCyd) and cytidine (Cyd) as well as several modified nucleosides including 2′,3′-dideoxycytidine (ddCyd), 5-methyl-2′-deoxycytidine (m^5dCyd), and 5-methylcytidine (m^5Cyd).[47,48] Comparisons among these systems and the protonated nucleobase pairs previously examined thus enable the influence of both 5-methylation and the sugar moiety on the BPE to be elucidated. The measured BPEs of the protonated cytidine nucleoside analogue base pairs indicate that the 2′- and 3′-hydroxy substituents have very little impact on the base-pairing interactions, whereas 5-methylation was found to enhance the BPE of dCyd and Cyd. However, the enhancement in the BPE upon 5-methylation was found to be more than twice as large for the DNA analogue than the RNA analogue. These results suggest that the influence of the 2′-hydroxy substituents may be more significant when the nucleobase is modified at the 5-position.

To further elucidate the influence of 5-modifications on the properties of Cyt residues, the influence of 5-halogenation on the base-pairing interactions of protonated cytidine nucleoside analogue base pairs, (*x*Cyd)H^+(*x*Cyd), are examined here by complimentary threshold collision-induced dissociation techniques and computational methods. The synergy of the experimental

measurements, theoretical calculations, and statistical data analyses provides accurate and validated thermochemical data and expanding insight into the role of modifications on the strength of noncanonical binding relevant to synthetic *i*-motif structures. The *x*Cyd nucleosides examined here include the 5-halogenated forms of the canonical DNA and RNA cytidine nucleosides, 5-fluoro-2′-deoxycytidine (fl^{5}dCyd), 5-fluorocytidine (fl^{5}Cyd), 5-chloro-2′-deoxycytidine (cl^{5}dCyd), 5-chlorocytidine (cl^{5}Cyd), 5-bromo-2′-deoxycytidine (br^{5}dCyd), 5-bromocytidine (br^{5}Cyd), 5-iodo-2′-deoxycytidine (io^{5}dCyd), and 5-iodocytidine (io^{5}Cyd). Comparisons among these model base pairs and to the analogous canonical DNA and RNA cytidine nucleoside base pairs and their 5-methylated analogues previously studied enable the influence of 5-halogenation vs. 5-methylation and the 2'-deoxyribose vs. ribose sugar moieties on the base-pairing interactions to be determined. Results of this study provide insights that can be employed to guide the design of DNA sequences that enable dynamic control of structural transformations based on pH dependent *i*-motif formation that can be used to drive nanomechanical devices.

EXPERIMENTAL AND COMPUTATIONAL METHODS

Instrumentation and Experimental Procedures. The cross sections for CID of eight protonated cytidine nucleoside analogue base pairs, (*x*Cyd)H$^+$(*x*Cyd), are measured as a function of kinetic energy using a custom-built guided ion beam tandem mass spectrometer that has been described in detail previously.[49,50] The (*x*Cyd)H$^+$(*x*Cyd) base pairs are generated by electrospray ionization (ESI) from a 0.5–1.0 mM solution of the nucleoside and 1% acetic acid in an approximately 50%:50% MeOH:H_2O mixture. The solution is pumped at a flow rate of 0.1–0.3 μL/min through a 35-gauge 304 stainless steel ESI needle operating at a voltage of 1.5–2.0 kV. The droplets emanating from the spray are sampled into the vacuum region of the mass spectrometer through an inlet capillary, biased at 20–35 V, and heated to a temperature of ~100 ºC to facilitate ion desolvation. The ions exiting the inlet capillary are focused into a radiofrequency ion funnel (rf IF). A linear dc gradient across the rf IF is used to facilitate ion transport into the rf hexapole ion guide, which traps the ions in the radial direction with an amplitude of ~250 V_{pp}. As

the ions drift through the hexapole they undergo >10^4 collisions with the ambient gases, which results in thermalization to room temperature.[51]

The thermalized ions exiting the hexapole are accelerated and focused into a magnetic sector momentum analyzer, where the precursor (*x*Cyd)H^+(*x*Cyd) base pair is selected in the first stage of mass analysis. The (*x*Cyd)H^+(*x*Cyd) base pairs are subsequently decelerated in an exponential retarder, focused, and injected into an rf octopole ion guide, where the ions are radially trapped with an amplitude of ~170 V_{pp}.[52-,54] The octopole passes through a static collision gas cell where Xe is present at low pressure. The pressure dependence of the CID behavior is examined at three nominal pressures, ~0.05, 0.10, and 0.20 mTorr. At these pressures, single-collision conditions prevail as the (*x*Cyd)H^+(*x*Cyd) base pairs undergo CID with Xe as they drift along the octopole and through the collision cell.[55,56] Because Xe is heavy, monoatomic, chemically unreactive, and highly polarizable, it is the collision gas employed to induce dissociation to ensure efficient kinetic-to-internal energy transfer in the (*x*Cyd)H^+(*x*Cyd) base pair upon collision.[57,58] The axial translational energy of the precursor (*x*Cyd)H^+(*x*Cyd) base pair is controlled by the application of a dc voltage to the octopole. The octopole dc voltage is scanned through the range of laboratory energies over which the CID behavior is of interest. In the experiments performed here, the CID cross sections were measured over the range of laboratory energies corresponding to energies in the center-of-mass frame of 0–7.5 eV. After collision, the CID products and any undissociated (*x*Cyd)H^+(*x*Cyd) base pairs drift to the end of the octopole where the ions are focused into a quadrupole mass filter for the second stage of mass analysis, and detected using a Daly detector and standard pulse-counting techniques.

Theoretical Procedures. The experimental measurements are supported and enhanced by structural and energetic information and molecular parameters provided by complementary electronic structure calculations. The harmonic vibrational frequencies and rotational constants of the (*x*Cyd)H^+(*x*Cyd) base pairs and their CID products, the neutral *x*Cyd and protonated H^+(*x*Cyd) nucleosides, and the isotopic molecular polarizabilities of the neutral *x*Cyd nucleosides are used for thermochemical modeling of the measured CID cross sections. The energies of the ground

conformations of the (*x*Cyd)H^+(*x*Cyd) and their CID products, H^+(*x*Cyd) and *x*Cyd, are used to predict their BPEs for comparison to the measured values. This synergism between experiment and theory enables more accurate threshold determinations with concomitant validation of both experimental and theoretical results.

Simulated Annealing Protocol. Structural characterization of the (*x*Cyd)H^+(*x*Cyd) base pairs and the neutral *x*Cyd and protonated H^+(*x*Cyd) nucleosides begins with molecular mechanics simulated annealing using the Amber 3 force field and HyperChem software.[59] A three-stage simulated annealing process is employed. Each cycle begins at 0 K, the temperature is linearly increased to 1000 K, the simulation temperature, over a period of 0.3 ps. Conformational space is then sampled for 0.2 ps, and the temperature is then linearly decreased to 0 K over 0.3 ps. A minimum of 300 cycles of heating and cooling are typically performed for each initial structure examined to ensure sufficient sampling of conformational space. All unique conformers found during the simulated annealing procedures that are within 30 kJ/mol of the most stable conformer found are subjected to further interrogation.

Geometry Optimization, Frequency Analysis, and Single-Point Energy Calculations. Density functional theory calculations are performed to examine the unique low-energy conformers identified with the simulated annealing and energy filtering procedures described above; these calculations are performed using Gaussian 16.[60] Geometry optimizations are performed at the B3LYP/6-311+G(d,p) level of theory. Frequency analyses are then performed to confirm that the optimized geometries found are stable structures, to provide molecular parameters needed to compute thermal corrections, and for the thermochemical analysis of the measured CID cross sections. Single point energies of the B3LYP/6-311+G(d,p) optimized geometries are computed at the B3LYP/6-311+G(2d,2p), B3LYP/def2-TVZPPD, B3P86/6-311+G(2d,2p), and M06-2X/6-311+G(2d,2p) levels of theory. Zero-point energy corrections are applied to the single point energies to yield 0 K thermodynamic values. Basis set superposition error corrections are also applied to the computed BPEs using the counterpoise method.[61,62]

Three important structural parameters that are used to characterize various stable conformations of nucleosides as previously described[63] are extracted from the B3LYP/6-311+G(d,p) optimized geometries. First among these parameters is the pseudorotation phase angle (*P*), which describes the sugar puckering and is determined from the five dihedral angles of the sugar ring: ∠C4′O4′C1′C2′, ∠O4′C1′C2′C3′, ∠C1′C2′C3′C4′, ∠C2′C3′C4′O4′, and ∠C3′C4′O4′C1′). The historical designations for the sugar puckering described in terms of envelope designations, C*n*'-endo and C*n*'-exo, are most commonly employed to describe the sugar puckering. Here *n* denotes the number of the atom that deviates most from the plane of the sugar ring, endo indicates that the deviation is toward, whereas exo indicates that the direction of deviation is away, from the 5′-hydroxy substituent. The sugar puckering is alternatively described in terms of envelope (E) and twist (T) designations, where twist puckering describes the more commonly observed structures in which two atoms of the sugar ring pucker in opposite directions to the plane defined by the other three ring atoms. Endo configurations are designated using superscripts, whereas exo puckering is denoted using subscripts. The atom number of the major deviator (*n*') appears before the E or T designation, whereas the atom number of the minor deviator is given after the T designation. The orientation of the nucleobase relative to the sugar, the glycosidic bond angle, is described by the ∠O4′C1′N1C2 dihedral angle. The glycosidic bond angle is designated as *anti* for values between 90º and 270º or *syn* for values in the range from –90º to 90º. The orientation of the 5′-hydroxy substituent relative to the sugar is described by the ∠O4′C4′C5′O5′ dihedral angle. The 5′-hydroxy orientation is defined as gauche$^+$ when the O5′–H bond lies between the O4′ and C3′ atoms (corresponding to ∠O4′C4′C5′O5′ dihedral angles in the range between 120º and 240º), gauche$^-$ when the O5′–H bond points away from the O4′ atom (dihedral angles between 240º and 360º), or trans when the O5′–H bond points away from C3′ atom (dihedral angles between 0º and 120º).

Isotropic Molecular Polarizabilities of the *x*Cyd Nucleosides. Among the parameters required for thermochemical analysis of cross sections for CID of noncovalently bound complexes such as the (*x*Cyd)H$^+$(*x*Cyd) base pairs examined here are the isotropic molecular polarizabilities

(α) of the neutral CID products, which are the neutral *x*Cyd nucleosides. The isotropic molecular polarizabilities of the B3LYP/6-311+G(d,p) ground conformers of the neutral *x*Cyd nucleosides are computed at the PBE0/6-311+G(2d,2p) level of theory. Polarizabilities of polyatomic molecules calculated using the PBE0 functional have been found to exhibit very good agreement with experimentally measured values.[64]

Electrostatic Potential Maps. Electrostatic potential (ESP) maps of the ground conformations of the (*x*Cyd)H^+(*x*Cyd) base pairs are computed at the B3LYP/6-311+G(d,p) level of theory and color-mapped onto an isosurface of 0.06 a.u. of the total SCF electron density using GaussView6.[60] The ESP maps and Mülliken charges enable the distribution of charges localized on the atoms of the Cyt residues and sugar moieties that serve as hydrogen-bond donor and acceptors of the base-pairing interactions as well as those associated with sugar-sugar and sugar-nucleobase stabilizing interactions to be visualized. Importantly, comparisons among the ESP maps for these systems elucidate the effects of 5-halogenation (and the trends as a function of the size of the 5-halogen substituent) on the stability of the (*x*Cyd)H^+(*x*Cyd) base pairs.

Noncovalent Interaction Maps. Three-dimensional maps of the noncovalent interaction (NCI) regions in the ground conformers of the (*x*Cyd)H^+(*x*Cyd) base pairs are determined using the combined NCI and quantum theory of atoms in molecules (QTAIM) formalisms[65-,,68] using Multi-WFN 3.7.[69] The NCI maps are superimposed onto structures using the Visual Molecular Dynamics (VMD) program.[70] The NCI maps enable identification and characterization of the strength of the noncovalent interactions in these (*x*Cyd)H^+(*x*Cyd) base pairs with the associated bond paths identified by the QTAIM analysis. Attractive interactions (hydrogen bonds, van der Waals) and repulsive interactions (steric clashes) are differentiated using a conventional scalar coloring scheme where very intense colors are associated with regions of higher electron density and correspond to stronger noncovalent interactions. Blue is used for attractive (stabilizing) interactions, whereas red is used to show repulsive (destabilizing) interactions; green and yellow represent delocalized weak attractive and repulsive noncovalent interactions, respectively. The

bond paths associated with noncovalent interactions are shown as gold lines between the interacting atoms.

Data Handling Procedures. The precursor and product ion intensities, measured as a function of collision energy, are converted to absolute energy-dependent CID cross sections using Beer's law as described previously.[71] The sum of the uncertainties in the absolute cross sections measured are ~ 20%, and arises from approximately equal contributions from random errors in pressure measurement and the collision cell length.

Conversion from the laboratory frame ion kinetic energies, E_{lab}, to the relative collision energy in the center-of-mass frame, E_{cm}, is performed using the stationary target assumption given by $E_{cm} = E_{lab}\, m\, /(m + M)$, where m is the mass of the neutral collision gas, Xe, and M is the mass of the (xCyd)H$^+$(xCyd) base pair. All energies reported here are in the *cm* frame unless otherwise noted. A retarding potential analysis is performed using the octopole ion guide to measure the precursor ion beam absolute zero of energy and the kinetic energy distribution as previously described.[71] The first derivative of the normalized ion beam intensity is fitted to a Gaussian distribution. The center and full width at half-maximum (fwhm) of the nearly Gaussian fit gives the origin for the laboratory frame energy and the width of the kinetic energy distribution of the ion beam, which is typically in the range from 0.2 to 0.4 eV (lab), respectively.

The shapes, and in particular, the threshold regions of measured CID cross sections are sensitive to pressure effects arising from multiple ion-neutral collisions. The pressure dependence is straightforwardly eliminated from the measured CID cross sections by performing CID experiments at multiple Xe pressures (~ 0.05, 0.1 and 0.2 mTorr) that each correspond to nominally single-collision conditions, and linearly extrapolating to zero pressure of the Xe reactant. Cross sections subjected to thermochemical analysis therefore represent rigorously single-collision conditions.[72,73]

Thermochemical Analysis. To ensure that accurate thermochemical information is extracted from modeling of CID cross sections requires that the kinetic and internal energy distributions of the (xCyd)H$^+$(xCyd) and Xe reactants,[73,74] the effects of multiple ion-neutral

collisions,[72] and lifetime effects of the activated (*x*Cyd)H$^+$(*x*Cyd) base pair,[75,76] which may exceed the time-of-flight of the experiment, are properly accounted for in the analysis.

The zero-pressure-extrapolated CID cross sections of the (*x*Cyd)H$^+$(*x*Cyd) base pairs are subjected to thermochemical analysis using an empirical threshold function of the form given by eq 1:

$$\sigma(E) = \sigma_0 \sum_i g_i (E + E_i - E_0)^n / E \quad (1)$$

where σ_0 is an energy independent scaling factor, E is the relative translational energy of the (*x*Cyd)H$^+$(*x*Cyd) and Xe reactants, E_0, is the threshold for dissociation at 0 K, and n is an adjustable fitting parameter that describes the efficiency of translational-to-internal energy transfer upon collision.[77] The summation is over the ro-vibrational states of the (*x*Cyd)H$^+$(*x*Cyd) base pair with excitation energies, E_i, and relative populations, g_i, where $\Sigma g_i = 1$. The relative populations, g_i, are calculated based on a 300 K Maxwell–Boltzmann distribution. The number and densities of ro-vibrational states are directly counted using the Beyer–Swinehart–Stein–Rabinovitch algorithm and employed to evaluate the internal energy distribution and the unimolecular rate constant for dissociation of the (*x*Cyd)H$^+$(*x*Cyd) base pair.[78-80]

The (*x*Cyd)H$^+$(*x*Cyd) base pairs examined here are sufficiently large and possess many vibrational degrees of freedom (171 for the DNA and 177 for the RNA (*x*Cyd)H$^+$(*x*Cyd) base pair analogues) such that even when they possess sufficient internal energy to undergo dissociation, there is a measurable probability that the average time needed for dissociation to occur may exceed the time scale of the experiment, which is ~100 μs on our custom-built GIBMS instrument. This leads to a shift to higher translational energies in the observed threshold for dissociation, which is commonly referred to as a "kinetic" shift. The lifetime of the energized complex is explicitly accounted for in the analysis of the CID cross sections by incorporating Rice–Ramsperger–Kassel–Marcus (RRKM) statistical theory for unimolecular dissociation into eq 1, as described by eq 2.[75,76] This RRKM treatment has been validated is previous studies where the thresholds determined for large energized complexes without inclusion of the RRKM formalism in the modeling procedures

leads to fits that are not able to reproduce curvature in the onset or threshold regions of the CID cross sections and produce anomalously high threshold values.[76]

$$\sigma_j(E) = \frac{n\sigma_0}{E}\sum_i g_i \int_{E_0-E_i}^{E} (E-\varepsilon)^{n-1} P_D dE \quad (2)$$

In the model of eq 2, ε is the energy deposited into the internal degrees of freedom of the (*x*Cyd)H$^+$(*x*Cyd) base pair upon collision with the neutral Xe reactant, such that its internal energy increases to $E^* = \varepsilon + E_i$. The probability of dissociation, P_D, is expressed in terms of the RRKM unimolecular dissociation rate coefficient, $k(E^*)$, as $P_D = (1 - \exp\{-k(E^*)T\})$. The RRKM unimolecular dissociation rate coefficient is calculated using eq 3.

$$k(E^*) = sN^\dagger(E^* - E_0) / h\rho(E^*) \quad (3)$$

where s is the reaction degeneracy, $N^\dagger(E^* - E_0)$ is the sum of the ro-vibrational states of the transition state (TS), h is Planck's constant, and $\rho(E^*)$ is the density of ro-vibrational states for the energized (*x*Cyd)H$^+$(*x*Cyd) base pair at E^*, the internal energy available. Whenever the rate of decomposition is much faster than the average experimental time window, eq 2 simplifies to eq 1.

The ro-vibrational frequencies of the energized (*x*Cyd)H$^+$(*x*Cyd) base pair (EM) and transition state (TS) for dissociation are incorporated in the calculation of the RRKM unimolecular rate constants. Theoretical calculations of the energized (*x*Cyd)H$^+$(*x*Cyd) base pair provide the required molecular parameters for the EM. CID of the (*x*Cyd)H$^+$(*x*Cyd) base pair results in noncovalent dissociation of the base-pairing interactions to produce the protonated H$^+$(*x*Cyd) and neutral *x*Cyd nucleosides. The threshold energy determined from modeling the measured CID cross section thus provides the BPE of the (*x*Cyd)H$^+$(*x*Cyd) base pair. Noncovalent bond cleavage processes such as those observed here occur via a loose phase space limit (PSL)[76] TS, in which the TS is located at the centrifugal barrier for dissociation. As a result, there is no reverse activation barrier in excess of the endothermicity of dissociation, and the molecular parameters that describe such a TS are simply those of the H$^+$(*x*Cyd) and *x*Cyd products. Combining the RRKM statistical treatment of the lifetime of the dissociating ions with the PSL formalism enables accurate threshold energies to be determined, which are denoted as E_0(PSL) to differentiate these values from the

kinetically shifted threshold energies, E_0, extracted from analyses that exclude the RRKM lifetime analysis.

Threshold energies are determined by fitting the threshold regions of the CID cross sections to the model function of eq 2. The CID cross section model of eq 2 takes into account effects of the internal energy of the (*x*Cyd)H$^+$(*x*Cyd) base pair, Doppler broadening by Xe, and the relative kinetic energy distribution of the (*x*Cyd)H$^+$(*x*Cyd) base pair and Xe reactants by convoluting the model cross sections with these energy distributions. The convoluted cross section is then compared with the experimental cross section over as large a range of energies as possible while still accurately reproducing the threshold region and simultaneously optimizing the fitting parameters, σ_0, E_0 or E_0(PSL), and n, to extract the best least-squares fit of the data. The statistical uncertainties associated with the σ_0, E_0 or E_0(PSL), and n fitting parameters are conservatively estimated from variations in these parameters determined across different data sets, and by modeling the zero-pressure cross sections using vibrational frequencies for the TS and EM that have been scaled by ±10%, and by scaling the assumed experimental post-collision flight distance up and down by a factor of 2 (i.e., 15.0 cm, 30.0, and 7.5 cm). These values correspond to the assumption that the CID event occurs at the center of the collision cell, and the entrance to the octopole, and half way between the collision cell and the exit of the octopole. Because all sources of energy are accounted for in these analyses, the measured thresholds, E_0(PSL), correspond to the BPE at 0 K for the (*x*Cyd)H$^+$(*x*Cyd) base pair.

RESULTS

Cross Sections for Collision-Induced Dissociation. Experimental CID cross sections for the interaction of the (fl^{5}dCyd)H$^+$(fl^{5}dCyd) and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs with Xe over the range of collision energies extending from ~0 to 6 eV are compared in Figure 1. A similar comparison is shown for all eight (*x*Cyd)H$^+$(*x*Cyd) base pairs in Figure S1 of the Supporting Information. The major CID pathway observed involves loss of an intact neutral *x*Cyd nucleoside as shown in the CID reactions 4.

$$(x\text{Cyd})H^+(x\text{Cyd}) + Xe \rightarrow H^+(x\text{Cyd}) + x\text{Cyd} + Xe \quad (4)$$

Reaction 4, an endothermic process, involves heterolytic cleavage of the intermolecular hydrogen bonds associated with the base-pairing interactions in these $(x\text{Cyd})H^+(x\text{Cyd})$ base pairs. At elevated collision energies, the primary $H^+(x\text{Cyd})$ product is formed with enough internal energy that it undergoes sequential dissociation. Two sequential dissociation pathways, both involving glycosidic bond cleavage occur in competition. In all cases glycosidic bond cleavage with the excess proton retained by the departing nucleobase is preferred, reaction 5, whereas retention of the excess proton by the sugar moiety is less favourable by more than an order of magnitude), reaction 6.

$$H^+(x\text{Cyd}) \rightarrow H^+(x\text{Cyt}) + (x\text{Cyd}–x\text{Cyt}) \quad (5)$$

$$H^+(x\text{Cyd}) \rightarrow H^+(x\text{Cyd}–x\text{Cyt}) + x\text{Cyt} \quad (6)$$

The primary and sequential dissociation behavior observed for these $(x\text{Cyd})H^+(x\text{Cyd})$ base pairs parallels that observed previously for protonated base pairs of cytosine nucleobase and cytidine nucleoside analogues[41-44,47,48,81] and that of isolated protonated cytidine nucleoside analogues[82-,,,,87] using GIBMS and IRMPD approaches.

Threshold Analysis. The threshold regions of the CID cross sections for reaction 4 were analyzed by modeling with and without the incorporation of lifetime effects via RRKM using the empirical threshold laws of eq 2 and eq 1, respectively. Modeling of reaction 4 is based on a loose PSL TS.[76] The PSL TS model has been shown to provide the most accurate correction for kinetic shifts observed for CID reactions of noncovalently-bound complexes.[41-44,47,48,88-,,,,,,,96] The molecular parameters (vibrational frequencies and rotational constants) that describe the energized $(x\text{Cyd})H^+(x\text{Cyd})$ base pair and corresponding PSL TSs used for the threshold analyses are listed in Tables S1 and Table S2. The threshold energies, E_0(PSL), along with the modeling parameters obtained from these analyses and observed kinetics shifts are summarized in Table 1. Fits to the CID cross sections of the $(fl^5dCyd)H^+(fl^5dCyd)$ and $(io^5dCyd)H^+(io^5dCyd)$ base pairs are illustrated in Figure 2. A similar comparison for all eight $(x\text{Cyd})H^+(x\text{Cyd})$ base pairs is shown in Figure S2. As can be seen in Figure 2 and Figure S2, the experimental cross sections of reaction 4

are reproduced with high fidelity throughout the threshold regions using the loose PSL TS model. These analyses yield E_0(PSL) threshold values of 1.86 ± 0.05, 1.75 ± 0.05, 1.79 ± 0.05, and 1.87 ± 0.05 eV for the (*x*Cyd)H$^+$(*x*Cyd) base pairs of the DNA analogues, *x*Cyd = fl^{5}dCyd, cl^{5}dCyd, br^{5}dCyd, and io^{5}dCyd. The E_0(PSL) threshold values determined for the base pairs of the RNA analogues are lower, 1.66 ± 0.04, 1.68 ± 0.04, 1.70 ± 0.05, and 1.75 ± 0.05 eV for *x*Cyd = fl^{5}Cyd, cl^{5}Cyd, br^{5}Cyd, and io^{5}Cyd. Entropies of activation at 1000 K, $\Delta S^{\ddagger}$(PSL), calculated from the molecular constants used for RRKM lifetime modeling are listed in Table 1. The entropies of activation at 1000 K are large and positive and vary from 82 to 109 J·mol^{-1}K^{-1}, consistent with expectations for cleavage of multiple noncovalent interactions in the CID process and occurring via a loose PSL TS.[88-96] The extent of kinetic shifting associated with the finite experimental time window is estimated by modeling the CID cross sections without accounting for lifetime effects. Without inclusion of the RRKM formalism in the modeling, the threshold values determined increase by 1.74, 1.72, 1.86, and 1.72 eV for the base pairs of the DNA analogues, *x*Cyd = fl^{5}dCyd, cl^{5}dCyd, br^{5}dCyd, and io^{5}dCyd, and by 1.52, 1.64, 1.61, and 1.97 eV for the base pairs of the RNA analogues, *x*Cyd = fl^{5}Cyd, cl^{5}Cyd, br^{5}Cyd, and io^{5}Cyd, respectively. These significant and relatively large kinetic shifts are consistent with the size and strength of binding in these (*x*Cyd)H$^+$(*x*Cyd) base pairs. The stronger binding in the base pairs of the DNA analogues leads to greater kinetic shifting than produced by the six additional vibrational degrees of freedom available to the base pairs of the RNA analogues.

Optimized Geometries of the Ground and Stable Low-Energy Conformers of the (*x*Cyd)H$^+$(*x*Cyd) Base Pairs. Ground and stable low-energy structures of the protonated nucleoside base pairs, (*x*Cyd)H$^+$(*x*Cyd), protonated nucleosides, H$^+$(*x*Cyd), and neutral nucleosides, *x*Cyd, were calculated as described in the Theoretical Procedures section. The ground conformations of the (fl^{5}dCyd)H$^+$(fl^{5}dCyd) and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs are compared to the canonical (dCyd)H$^+$(dCyd) base pair in Figure 3, whereas Table S3 compares the geometric parameters of these ground conformations. A similar comparison of the ground conformations of the (*x*Cyd)H$^+$(*x*Cyd) base pairs of the canonical DNA and RNA, and the 5-methylated and 5-

halogenated cytidine nucleoside analogues is provided in Figure S3. All low-energy conformers (within 10 kJ/mol Gibbs energy) of the ground conformers of these (*x*Cyd)H$^+$(*x*Cyd) base pairs are compared in Figure S4. As can be seen in Figure 3 and Figures S3 and S4, the ground and low-energy conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs are bound by three hydrogen-bonding interactions. The central hydrogen bond, N3H$^+$···N3, is ionic due to N3-protonation of one of the Cyt residues, whereas the two terminal hydrogen bonds,[84] N4H···O2 and O2···HN4, are conventional neutral hydrogen bonds. The N4H···O2 hydrogen bond in which the protonated nucleoside serves as the hydrogen-bond donor (hereafter referred to as the upper-terminal hydrogen bond) is much shorter than the O2···HN4 hydrogen bond in which the protonated nucleoside acts as the hydrogen-bond acceptor (hereafter referred to as the lower-terminal hydrogen bond).

The presence of the 2'- and 3'-hydroxy substituents in the (*x*Cyd)H$^+$(*x*Cyd) base pairs of the RNA nucleoside analogues, enables the formation of an O2′H···O3′ sugar-sugar intramolecular hydrogen bond in each nucleoside of the base pair. Addition of the 2'-hydroxy substituents to the 5-halogenated DNA base pairs to produce the analogous RNA base pairs generally leads to a slight increase in the length of all three hydrogen-bonding interactions, with the magnitude of the changes increasing with the size of the halogen substituent for the upper- terminal and central-ionic hydrogen bonds and decreasing in magnitude for the lower-terminal hydrogen bond. The influence of 5-halogenation varies with the size of the substituent. 5-Fluorination leads to a decrease in the length of the central-ionic hydrogen bond, whereas the larger halogens produce a slight lengthening in this hydrogen bond. 5-Halogenation also leads to a minor decrease in the length of the terminal hydrogen bonds, which increase with the size of the substituent. 5-Halogenation also leads to a slight lengthening in the noncanonical O5′···HC6 hydrogen-bonding interactions that also increases with the size of the substituent.

The protonated and neutral *x*Cyd nucleosides in the (*x*Cyd)H$^+$(*x*Cyd) base pairs adopt an antiparallel configuration, as generally found in double-stranded nucleic acid structures. The sugar puckering of the ground conformer of the canonical (dCyd)H$^+$(dCyd) base pair is **C3'endo**·C3'-

endo, whereas its 5-halogenated analogues all exhibit **C2'-endo**·C2'-endo sugar puckering. Here, the sugar puckering of the protonated H^+(*x*Cyd) nucleoside is indicated first in boldface font, followed by that of the neutral *x*Cyd nucleoside in standard font. In contrast, the preferred sugar puckering for the canonical (Cyd)H^+(Cyd) base pair and the br^{5}Cyd and io^{5}Cyd analogues is **C2'-endo**·C2'-endo, whereas its fl^{5}Cyd and cl^{5}Cyd analogues exhibit a very slight preference for **C2'-endo**·C3'-endo sugar puckering (see Figure 3 and Figure S3).

Optimized Geometries of the Ground and Stable Low-Energy Structures of the Neutral *x*Cyd and Protonated H^+(*x*Cyd) Nucleosides. Parallel density functional theory calculations were performed for the neutral *x*Cyd and protonated H^+(*x*Cyd) nucleosides. The stable low-energy conformers of these species along with their sugar puckerings, nucleobase and 5'-hydroxy orientations, and relative Gibbs energies at 298 K are given in Figures S5 and S6, respectively. The geometric parameters of the ground conformers of the neutral nucleosides are listed in Table S4, whereas the geometric parameters of the ground N3- and O2-protonated conformers of the nucleosides are summarized in Tables S5 and S6, respectively.

As can be seen in Figure S5, the ground conformations of dCyd and its 5-halogenated analogues all exhibit C2'-endo sugar puckering, a *syn* nucleobase orientation, and a gauche$^+$ 5-hydroxy orientation. The 5-methylated analogue, m^{5}dCyd, was found to exhibit C3'-endo sugar puckering, a *syn* nucleobase orientation, and a gauche$^+$ 5'-hydroxy orientation. The *syn* nucleobase orientation is favored over *anti* as it is stabilized by a O5′H···O2 hydrogen-bonding interaction. The most stable *anti* conformation found for dCyd is 5.0 kJ/mol less stable, whereas the most stable *anti* conformations of the 5-substituteed dCyd analogues are more stable and within 2.7 kJ/mol of the ground conformer. In contrast, the ground conformations of Cyd and its 5'-halogenated analogues exhibit C2′-endo sugar puckering, *anti* nucleobase orientations, and gauche$^+$ 5'-hydroxy orientations. Again, the 5-methylated analogue, m^{5}Cyd, differs slightly as it exhibits C2′-endo sugar puckering, a *trans* nucleobase orientation, and a gauche$^+$ 5'-hydroxy orientation. In all cases, the *anti* nucleobase orientation is favored over *syn* as it is stabilized by O3′H···O2′H···O2 dual hydrogen-bonding interactions.

5-Halogenation is found to alter the preferred site of protonation, see Figure S6. N3 protonation is favored over O2 for dCyd and Cyd,[48] whereas the 5-halogenated analogues favor O2 protonation over N3. The preference for O2 protonation is larger for the RNA than the DNA analogues and falls off as the size of the halogen substituent increases. In contrast, the preference for N3 protonation over O2 is preserved but reduced in magnitude for the 5-methylated analogues. The ground conformations of the protonated nucleosides, all exhibit C2'-endo or C3'-endo sugar puckering, *anti* nucleobase orientations, and gauche$^+$ 5'-hydroxy orientations. C2'-endo sugar puckering is favored for Cyd and the 5-halogenated analogues of dCyd and Cyd, whereas C3'-endo sugar puckering is favored for dCyd, m^{5}dCyd, and m^{5}Cyd. However, the difference in stability for C2'-endo vs. C3'-endo sugar puckered conformers is quite small and less than 1 kJ/mol for the dCyd analogues and less than 3 kJ/mol for the Cyd analogues. In all cases, the *anti* nucleobase orientations are stabilized by a noncanonical O5'···HC6 hydrogen-bonding interaction, with additional stabilization of the Cyd and the 5-halogenated analogues provided by an O2'H···O3' hydrogen-bonding interaction, and a O3'H···O2' hydrogen-bonding interaction for m^{5}Cyd.

Key Structural Parameters of the *x*Cyd Nucleosides. As summarized above, the sugar puckering, nucleobase orientation, and the 5'-hydroxy orientations provide key information relating to the structure and conformational flexibility available to the nucleosides. A more detailed examination of these parameters is warranted as the designations defined for each of these parameters span a range of values as described in the Theoretical Procedures section. These key structural parameters of the ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs are compared in the polar plots of Figure 4; a similar comparison of these parameters for all stable conformers computed within 10 kJ/mol of the ground conformers is provided in Figure S8. Among the ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs only C2'-endo and C3'-endo conformations are observed. The 2'-hydroxy and 5-halogen substituents do not significantly influence the preferred pseudorotation phase angles, glycosidic bond angles, or 5'-hydroxy orientations in the (*x*Cyd)H$^+$(*x*Cyd) base pairs. The (*x*Cyd)H$^+$(*x*Cyd) base pairs exclusively adopt *anti* conformations (as this is required for base pairing) with glycosidic bond angles of ~230º when the sugar exhibits

C2′-endo puckering and ~195º when the sugar exhibits C3′-endo puckering. 5-Halogenation leads to a slight increase (~1–3º) in the glycosidic bond angles of the protonated nucleoside of the base pair. In all cases, the 5'-hydroxy orientations are gauche$^+$, as this orientation provides stabilization via formation of noncanonical C6H···O5′ hydrogen-bonding interactions in both nucleosides of the base pair. Only modest variation in the pseudorotation phase angles is found among the low-energy conformers compared in Figure S8. Notably, a greater number of conformers exhibit twisted rather than envelope C2′-endo and C3′-endo sugar puckering, and a few conformers exhibit C3′-exo sugar puckering in the protonated nucleoside of the base pairs. The low-energy conformers exclusively adopt *anti* conformations; however, a broader distribution of glycosidic bond angles that vary between 150–250º are found with most conformers exhibiting glycosidic bond angles of ~230º, ~195º, or ~155º. While all of the ground and many of the low-energy conformers exhibit gauche$^+$ 5'-hydroxy orientations, conformers with gauche$^-$ and trans orientations are also found among the low-energy conformers compared in Figure S8. These conformers lack the noncanonical C6H···O5′ hydrogen-bonding interactions; their relative stabilities suggest that these interactions provide ~6–8 kJ/mol to the overall stabilization of the DNA base pairs and ~4–8 kJ/mol for the RNA base pairs.

These three key structural parameters for the low-energy conformers (those within 10 kJ/mol of the ground conformer) of the neutral and protonated nucleosides are similarly summarized in the polar plots of Figures S9–S11. The polar plots for the neutral *x*Cyd nucleosides show that the stable low-energy conformers computed exhibit pseudorotation phase angles in the range between 6 and 192º. However, there is a strong preference for either C2'-endo or C3'-endo sugar puckering, with a greater number of conformers exhibiting C2'-endo puckering. Both *syn* and *anti* nucleobase orientations are represented among the low-energy conformers, with a greater number of conformers exhibiting *anti* nucleobase orientations even for the dCyd analogues where an energetic preference for a *syn* nucleobase orientation is found. Among the neutral *x*Cyd nucleosides, an energetic preference for a gauche$^+$ 5'-hydroxy orientation is found as this enables the C6H···O5′ hydrogen-bonding interaction. However, almost equal populations (~25% each) of

the low-energy conformers of the neutral nucleosides adopt gauche$^-$ or trans orientations. The polar plots for the protonated H$^+$(*x*Cyd) nucleosides show that both O2- and N3-protonation reduce the conformational flexibility of the nucleoside such that much less variation in the pseudorotation phase angles, nucleobase orientations, and 5'-hydroxy orientations are found among the low-energy conformations, particularly for the N3-protonated forms. The polar plots for the protonated *x*Cyd nucleosides show that the stable low-energy conformers computed exhibit pseudorotation phase angles in the range between 12 and 200º. There is again a strong preference for either C2′-endo or C3′-endo sugar puckering, with a greater number of conformers exhibiting C2′-endo puckering. However, a few conformers exhibiting C3′-exo puckering are found. Both *syn* and *anti* nucleobase orientations are represented among the low-energy conformers for H$^+$(dCyd), H$^+$(fl^{5}dCyd) and all the protonated Cyd analogues, with a greater number of conformers exhibiting *anti* nucleobase orientations, whereas only *anti* nucleobase orientations are found among the low-energy conformers of H$^+$(cl^{5}dCyd), H$^+$(br^{5}dCyd), H$^+$(io^{5}dCyd), and H$^+$(m^{5}dCyd). Among the N3- and O2-protonated *x*Cyd nucleosides, an energetic preference for a gauche$^+$ 5′-hydroxy orientation is again found as this enables the C6H···O5′ hydrogen-bonding interaction. No gauche$^-$ or trans orientations are found among the low-energy N3-protonated conformers but are found for the O2-protonated forms. These conformers lack the noncanonical C6H···O5′ hydrogen-bonding interactions; their relative stabilities suggest that these interactions provide ~11–14 kJ/mol to the overall stabilization of the dCyd analogues and ~12–17 kJ/mol for the Cyd analogues, roughly twice that found for the analogous protonated base pairs.

ESP Maps of the Ground Conformers of the (*x*Cyd)H$^+$(*x*Cyd) Base Pairs: Further insight into the structural and energetic effects of 5-halogenation on the base-pairing interactions in protonated cytidine nucleoside analogue base pairs, (*x*Cyd)H$^+$(*x*Cyd), is garnered from ESP maps. ESP maps and atomic charges obtained from Mülliken population analysis are superimposed on the B3LYP/6-311+G(d,p) ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs and compared in Figure S12. The Mülliken charges are shown for the atoms participating in the base-pairing interactions, the excess proton, the N3, O2, and N4H atoms of the Cyt residues as well as those of

the C5 atom and the 5-substituent, H or F, Cl, Br or I. The ESP maps of all of the (*x*Cyd)H^+(*x*Cyd) base pairs are highly parallel. As the most basic sites in cytidine nucleosides,[83] the N3 and O2 atoms have negative charges and appear red. As the hydrogen-bond donors of the N3H^+···N3 and N4H···O2 base-pairing interactions, the excess proton and the N4H atoms have positive charges and appear blue. Variation in the inductive and polarizability effects of the 5-substituents produces significant variations in the polarization of the C–X bond. There is obvious asymmetry in the ESP maps, readily differentiating the protonated H^+(*x*Cyd) nucleoside (shown on the left) from the neutral nucleoside *x*Cyd (shown on the right) of the base pair. At room temperature these base pairs should have enough internal energy to allow the excess proton to oscillate between the N3 atoms of both nucleosides of the base pair, simultaneously interconverting the upper and lower terminal (or short and long) N4H···O2 hydrogen bonds between these base pairs such that the actual structure and ESP map might be better represented as an average of the conformations. Among the hydrogen-bond donors, the excess proton exhibits the largest atomic charge followed by the N4H atom of the protonated nucleoside, and then the N4H atom of the neutral nucleoside. Among the hydrogen-bond acceptors, the N3 atom of the protonated nucleoside is most basic, followed by the O2 atom of the neutral nucleoside, then the N3 atom of the neutral nucleoside, and the O2 atom of the protonated nucleoside is the least basic. The larger positive charge on the N4H atom of the protonated nucleoside and larger negative charge on the O2 atom of the neutral nucleoside make them better hydrogen-bond donors and acceptors, whereas the lower positive charge on the N4H atom of the neutral nucleoside and lower negative charge on the O2 atom of protonated nucleoside make them poorer hydrogen-bond donors and acceptors consistent with formation of a shorter and a longer hydrogen bond (upper terminal N4H···O2 vs. lower terminal O2···HN4 hydrogen bonds of Figure 3 and Figure S3).

Effects of the 2'-Hydroxy Substituent. Comparison of the ESP maps for the DNA vs. RNA cytidine nucleoside analogue base pairs enables the specific influence of the 2'-hydroxy substituent to be elucidated. The Mülliken charges on the C5 atoms exhibit the largest differences, which vary with the C5 substituent. The Mülliken charges on the N3 and O2 atoms exhibit the

next largest differences. Importantly, the Mülliken charges are larger for the DNA than RNA nucleoside base pairs, consistent with the stronger base pairing in the DNA base pairs.

Effects of the 5-Substituents. Comparison of the ESP maps for the DNA and RNA cytidine nucleoside analogue base pairs as a function of the 5-substituent enables their influences to be elucidated. As both inductive and polarizability effects contribute, the trends are not quite simple. 5-Substitution generally leads to an increase in charge on the O2 and N4H atoms of both nucleosides and on the N3 atom of the protonated nucleoside, and a small decrease of the charge on the N3 atom of the neutral nucleoside of each base pair. The charges on the C5–X atoms become more negative for 5-fluoro substitution, and less negative (or even positive) for the other substituents. Importantly, the C–X bond exhibits the greatest polarization for the canonical and 5-methyl analogues, and the direction of the bond polarization differs for the canonical and 5-fluoro analogues from that of the other base pairs.

Noncovalent Interaction Maps of the (*x*Cyd)H$^+$(*x*Cyd) Base Pairs. To visualize the noncovalent interactions that determine structure and provide stabilization to the (*x*Cyd)H$^+$(*x*Cyd) base pairs and how these interactions are influenced by 5-halogenation, NCI maps are superimposed on the structures of the ground conformers along with bonding paths obtained from QTAIM analysis in Figure 5 for the (dCyd)H$^+$(dCyd), (fl^{5}dCyd)H$^+$(fl^{5}dCyd), and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs. A similar comparison for all the (*x*Cyd)H$^+$(*x*Cyd) base pairs is shown in Figure S13. The presence of intermolecular hydrogen-bonding interactions between the Cyt residues, i.e., the base-pairing interactions, along with additional intramolecular interactions between the Cyt residues and sugar moieties and within the sugar moiety are apparent in these figures. The base-pairing interactions are seen as two small blue and a third green disk-shaped region between the Cyt residues. Blue isosurfaces in the NCI plots indicate regions of relatively high electron density between the interacting atoms and stronger stabilizing interactions for the upper-terminal N4H···O2 and central-ionic N3H$^+$···N3 hydrogen-bonding interactions, whereas the green isosurface indicates lower electron density and a weaker stabilizing interaction for the bottom-terminal O2···HN4 hydrogen bond. The NCI plots also reveal the presence of weak

intramolecular van der Waal attractions between the Cyt and sugar moieties of each nucleoside. These also appear as green isosurfaces, and account for the formation of the noncanonical C6H···O5′ hydrogen-bonding interactions. NCI analyses further highlight weak O2′H···O3′ hydrogen-bonding interactions in the sugar moiety for the (Cyd)H$^+$(Cyd) base pair and its 5-halogenated analogues and noncanonical C2′H···O5′ hydrogen-bonding interactions in the structures exhibiting C2′-endo puckering (see Figure S13). 5-Halogenation of the Cyt residues introduces additional unique weak N4H···X intramolecular interactions that become more stabilizing as the size of the halogen substituent increases. While a small disk is visible in the NCI maps of the 5-halogenated (*x*Cyd)H$^+$(*x*Cyd) base pairs, a bond path associated with these N4H···X intramolecular interactions is only predicted by the QTAIM analysis for the 5-iodinated base pairs. Red isosurfaces arising from steric clashes among the ring atoms are seen in the center of the aromatic ring structure of the Cyt residues. The structural effects of the 2′-hydroxy and 5-halogen substituents on the inter- and intramolecular interactions were further elucidated by examining the intensities of the blue isosurfaces between the Cyt residues as shown in Figure 5 and Figure S13. In several of the (*x*Cyd)H$^+$(*x*Cyd) base pairs, a more intense isosurface is evident as indicated by an opening in the center of the upper blue isosurfaces qualitatively suggesting an increase in electron density. This ‘hole’ is very faint in the canonical (dCyd)H$^+$(dCyd) and (cl^{5}dCyd)H$^+$(cl^{5}dCyd) base pairs, and becomes increasingly noticeable for the (*x*Cyd)H$^+$(*x*Cyd) base pairs of br^{5}dCyd, fl^{5}dCyd, and io^{5}dCyd, respectively. In contrast, no “hole” is observed for the (*x*Cyd)H$^+$(*x*Cyd) base pairs of Cyd, fl^{5}Cyd, cl^{5}dCyd, and br^{5}dCyd, and is faint for the base pairs of io^{5}Cyd and m^{5}Cyd, indicative of depletion of electron density by the 2′-hydroxy substituent. The bond paths determined from topological analysis of Bader’s theory of atoms in molecules[57] are shown as solid yellow lines in Figure 5 (and Figure S13). A bond path is found to interconnect three intermolecular base-pairing interactions: N3H$^+$···N3, N4H···O2, and O2···HN4, and two intramolecular nucleobase-sugar and sugar-sugar hydrogen-bonding interactions in each nucleoside, C6H···O5′, and O2′H···O3′. Although NCI analysis indicates a weak green isosurface for the C2′H···O2 and N4H···X interactions, no bond path was found interconnecting the

interacting atoms except in the 5-iodinated base pairs where the large and polarizable I atom leads to sufficient stabilization that a bond path is found for the N4H···I interactions. The corresponding 2D NCI scatter maps for the ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs as shown in Figure S14 are plotted as the RDG versus the electron density. Three unique spikes are distinguishable in the 2D scatter maps, which are specified by the color-filled isosurface. The black horizontal line indicates the RDG plotted in Figure 5 (and Figure S13), such that the segments just crossing the peaks represent the RDG isosurface. The blue filled regions correspond to the stronger N3H$^+$···N3, N4H···O2, and O2···N4H hydrogen-bonding interactions. The weak noncanonical C6H···O5′ hydrogen-bonding interactions are shown as green regions, whereas the red regions refer to interactions within the Cyt and sugar rings where there is strong steric repulsion.

DISCUSSION

Comparison of Experimental and Theoretical (*x*Cyd)H$^+$(*x*Cyd) BPEs. The TCID measured and theoretically predicted BPEs of the protonated (*x*Cyd)H$^+$(*x*Cyd) base pairs are summarized in Table 2. In addition to the values determined here for the 5-halogenated cytidine nucleoside analogues (*x*Cyd = fl^{5}dCyd, cl^{5}dCyd, br^{5}dCyd, io^{5}dCyd, fl^{5}Cyd, cl^{5}Cyd, br^{5}Cyd, and io5Cyd), values previously reported for the canonical cytidine nucleosides (*x*Cyd = dCyd and Cyd) and their 5-methylated analogues (*x*Cyd = m^{5}dCyd and m^{5}Cyd) are also included.[48] Several levels of theory (B3LYP, B3P86, and M06-2X) are compared to assess their relative abilities to accurately describe the strength of the base-pairing interactions and the influence of the various modifications on the binding. The B3LYP results are compared with experiment in Figure 6; the performance of the 6-311+G(2d,2p) and def2-TZVPPD basis sets is evaluated. As can be seen in the figure, the B3LYP predicted BPEs exhibit excellent agreement with the TCID measured values. The mean absolute deviations (MADs) between the theoretically predicted and experimentally measured BPEs are 0.3 ± 0.2 and 2.2 ± 0.5 kJ/mol, respectively. These values are well within the average experimental uncertainty (AEU) in the TCID determinations, 4.6 ± 0.5 kJ/mol. Thus, B3LYP provides an excellent description of the noncovalent interactions that

stabilize the (*x*Cyd)H$^+$(*x*Cyd) base pairs with the 6-311+G(2d,2p) basis set performing marginally better than the def2-TVZPPD basis set. The B3P86 and M06-2X results are compared with experiment in Figure S13. For these theoretical approaches, only the 6-311+G(2d,2p) basis set is examined. As can be seen in the figure, both B3P86 and M06-2X systematically overestimate the BPEs. The MAD for B3P86 (8.8 ± 3.3 kJ/mol) is much poorer than found for B3LYP with either basis set, but better than found using M06-2X (13.0 ± 3.7 kJ/mol). The systematic shift in the predicted BPEs exceeds the AEU in the measured values indicating that these approaches are not able to describe the noncovalent interactions that contribute to the stability of these base pairs as effectively as the B3LYP functional. Notably, highly parallel performance is found with each of the theoretical approaches for the (*x*Cyd)H$^+$(*x*Cyd) base pairs of the canonical, 5-methylated,[48] and 5-halogenated cytidine nucleoside analogues as well as for protonated cytosine base pairs[41-44] suggesting that the B3LYP approach may serve as a quantitatively accurate guide for characterizing the strength of base pairing in other modified cytidine nucleoside analogue base pairs not yet examined experimentally, whereas B3P86 and M06-2X should provide accurate trends, but overpredicted BPEs.

Conversion from 0 to 298 K Thermochemistry. Although the TCID experiments reported here were performed at room temperature; the thresholds energies extracted from thermochemical analysis of the CID cross sections remove experimental broadening to provide 0 K bond enthalpies. To facilitate comparison to room temperature values typically reported in the literature, the 0 K BPEs determined here are converted to enthalpies and Gibbs energies at 298 K using standard statistical thermodynamics formulas and the molecular parameters (vibrational frequencies and rotational constants) of the B3LYP/6-311+G(d,p) optimized geometries of the ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs and their dissociation products, H$^+$(*x*Cyd) and *x*Cyd. The vibrational frequencies and average internal energies are listed in Table S1, whereas Table S2 lists the rotational constants. Listed in Table 3 are 0 and 298 K enthalpy, Gibbs energy, and enthalpic and entropic corrections for all eight 5-halogenated (*x*Cyd)H$^+$(*x*Cyd) base pairs examined in this work. The uncertainties in the enthalpies and Gibbs energies reported in Table 3

are estimated based on variation in the computed values arising from scaling the harmonic vibrational frequencies by ±10%.

Influence of the 2′-Hydroxy Substituents on the Base Pairing Interactions. The TCID measured and theoretically predicted BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs are summarized in Table 2. The TCID measured BPEs are also plotted as a function of the *x*Cyd nucleoside in Figure 7. In all cases, the presence of the 2′-hydroxy substituents, which differentiate the DNA and RNA cytidine nucleoside analogues weaken the base-pairing interactions. Weakening of the base-pairing interaction by the 2′-hydroxy substituents was previously reported for the base pairs involving the canonical cytidine nucleosides, dCyd and Cyd, where the BPE is reduced by a mere 0.8 ± 7.0 kJ/mol (TCID), 2.3 kJ/mol (B3LYP), and 0.1–0.6 kJ/mol (other levels of theory). The impact of the 2'-hydroxy substituents on the strength of base pairing is found to be much more dramatic when the cytosine nucleobases are modified at the 5-position. A decrease of 5.8 ± 7.6 kJ/mol (TCID), 6.4 kJ/mol (B3LYP), and 4.0–6.6 kJ/mol (other levels of theory) was found for the (*x*Cyd)H$^+$(*x*Cyd) base pairs involving the 5-methylated cytidine nucleoside analogues, m^{5}dCyd and m^{5}Cyd.[48] Here, the influence of the 2′-hydroxy substituents on the base pairing in the 5-halogenated (*x*Cyd)H$^+$(*x*Cyd) base pairs is found to be even more pronounced and to vary with the identity of the 5-halogen substituent. The largest difference is found for the 5-fluorinated cytidine nucleoside analogues, fl^{5}dCyd and fl^{5}Cyd, where the 2′-hydroxy substituents lower the BPE of the (*x*Cyd)H$^+$(*x*Cyd) base pairs by 18.7 ± 6.3 kJ/mol (TCID), 19.1 kJ/mol (B3LYP), and 13.6-20.2 kJ/mol (other levels of theory). The smallest difference is found for the 5-chlorinated cytidine nucleoside analogues, cl^{5}dCyd and cl^{5}Cyd, where the 2′-hydroxy substituents lower the BPE of the (*x*Cyd)H$^+$(*x*Cyd) base pairs by 8.6 ± 6.2 kJ/mol (TCID), 6.6 kJ/mol (B3LYP), and 7.5–15.2 kJ/mol (other levels of theory). As the size of the 5-halogen substituents increases from Cl to Br to I, the influence of the 2′-substituents on the base pairing increases. For the 5-brominated cytidine nucleoside analogues, br^{5}dCyd and br^{5}Cyd, the 2′-hydroxy substituents lower the BPE of the (*x*Cyd)H$^+$(*x*Cyd) base pairs by 9.4 ± 7.4 kJ/mol (TCID), 9.6 kJ/mol (B3LYP), and 8.4–10.8 kJ/mol (other levels of theory). For the 5-iodinated cytidine nucleoside analogues, io^{5}dCyd and io^{5}Cyd,

the 2′-hydroxy substituents lower the BPE of the (*x*Cyd)H$^+$(*x*Cyd) base pairs by 11.5 ± 6.2 kJ/mol (TCID), 11.1 kJ/mol (B3LYP), and 7.0–11.2 kJ/mol (other levels of theory). These trends in the BPEs are consistent with differences in the ESP maps of the (*x*Cyd)H$^+$(*x*Cyd) base pairs (see Figure S12). Overall, the trends in the measured and computed BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs indicate that the 2′-hydroxy substituents have very little impact on the strength of base-pairing in the canonical cytidine nucleosides, but that their influence is markedly altered by 5-modifications.

Influence of 5-Halogenation on the Base-Pairing Interactions. The TCID measured and theoretically predicted BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs are summarized in Table 2. The TCID measured BPEs are also plotted as a function of the *x*Cyd nucleoside in Figure 7. As discussed in the previous section, the influence of the 2′-hydroxy substituents on the BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs varies with the size of the 5-halogen substituent such that it is most instructive to examine the trends individually for the DNA and RNA cytidine nucleoside analogues. 5-Halogenation of the (dCyd)H$^+$(dCyd) base pair is found to enhance the BPE. The largest enhancement is the binding is found upon 5-fluorination with an 11.4 ± 7.2 kJ/mol increase in the BPE. 5-Chlorination has almost no effect with an increase in the BPE of 0.8 ± 7.0 kJ/mol. As the size of the 5-halogen substituents increases from Cl to Br to I, the enhancement in the BPE follows, with increases of 5.0 ± 7.2 kJ/mol and 12.6 ± 6.8 kJ/mol measured for 5-bromination and 5-iodination, respectively. B3LYP predict similar trends with enhancements in the BPE upon 5-halogenation of the DNA base pair of 10.6, –0.6, 3.9, and 11.3 kJ/mol, respectively. In contrast, 5-halogenation of the (Cyd)H$^+$(Cyd) base pair generally weakens the strength of the base-pairing interaction. 5-Fluorination and 5-chlorination produce similar weakening of the BPE, 6.5 ± 6.1 kJ/mol (B3LYP) and 7.0 ± 6.2 kJ/mol (B3LYP), respectively. 5-Bromination results in a smaller decrease in the strength of the base-pairing interaction, with a decrease of 3.6 ± 7.2 kJ/mol. The large polarizability of iodine alters this trend such that 5-iodination slightly increases the BPE, by 1.9 ± 6.2 kJ/mol. Again, B3LYP predicts similar trends with weakening in the BPE upon 5-halogenation of the RNA base pair of 6.2, 5.0, and 3.4 kJ/mol for F, Cl and Br, and a modest enhancement in the BPE for I of 2.5 kJ/mol. Overall, the trends in the measured and computed

BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs indicate that the 5-halogen substituents exert an influence on the base-pairing interactions that differs markedly for the protonated DNA vs. RNA cytidine nucleoside analogue base pairs and that further depend on the identity of the halogen. In contrast, 5-methylation was found to enhance the BPE for both the protonated DNA and RNA cytidine nucleoside base pairs, by 8.9 ± 8.0 and 3.9 ± 6.5 kJ/mol, respectively. The different influences of 5-halogenation vs. 5-methylation on the BPEs of the (*x*Cyd)H$^+$(*x*Cyd) base pairs suggest that both inductive and polarizability effects of the 5-substituents are influencing the base-pairing interactions.

Influence of Polarizability on the BPEs of (*x*Cyd)H$^+$(*x*Cyd) Base Pairs. The measured BPEs of the 5-halogenated (*x*Cyd)H$^+$(*x*Cyd) base pairs reported here along with those previously reported for the base pairs involving the canonical cytidine nucleosides, dCyd and Cyd, and their 5-methylated analogues, m^{5}dCyd and m^{5}Cyd,[48] are compared to the PBE0/6-311+G(2d,2p) predicted isotropic molecular polarizabilities of the neutral *x*Cyd nucleosides in Figure 8. As can be seen in the figure no simple correlation is found across these systems. In all cases, the BPE of the DNA base pair exceeds that of the analogous RNA base pair despite the increase in polarizability derived from the 2′-hydroxy substituent. 5-Halogenation is found to enhance the BPE of the DNA base pair for all four halogens, whereas 5-halogenation of the RNA base pair generally weakens the strength of the base-pairing interaction. Roughly parallel linear correlations between the BPEs of the DNA and RNA (*x*Cyd)H$^+$(*x*Cyd) base pairs and polarizabilities of the cytidine nucleoside analogues is found for the 5-chlorinated, 5-brominated, and 5-iodinated cytidine nucleoside analogues suggesting that polarizability is an important contributor to the binding in these base pairs. In contrast, the canonical, 5-fluorinated, and 5-methylated cytidine nucleoside analogues deviate from this trend in that their BPEs are greater than expected based on the correlations found for the larger halogens. Among the DNA analogues, deviations from this correlation are largest for fl^{5}dCyd, followed by m^{5}dCyd, and the least for dCyd. The opposite trend is observed among the RNA analogues where the largest deviation is found for Cyd, followed by

m^{5}Cyd, and least for fl^{5}Cyd. These trends clearly indicated that inductive effects are also important contributors to the binding in these base pairs.

Inductive Effects on the BPEs of (*x*Cyd)H$^+$(*x*Cyd) Base Pairs. The measured BPEs of the 5-halogenated (*x*Cyd)H$^+$(*x*Cyd) base pairs reported here along with those previously reported for the base pairs involving the canonical cytidine nucleosides, dCyd and Cyd, and their 5-methylated analogues, m^{5}dCyd and m^{5}Cyd,[48] are compared to the projections of the dipole moments of the neutral *x*Cyd nucleosides along the direction of the base-pairing interactions in Figure 9. As can be seen in the figure a weak linear correlation is found across these systems. However, polarizability effects are clearly important as several of the systems significantly deviate from this trend, and notably, the 5-halogenated DNA base pairs (and separately the 5-halogenated RNA base pairs) generally exhibit similar dipole moment projections yet the BPEs vary considerably. Most notably, the presence of the 2′-hydroxy substituents in the RNA base pairs decreases the magnitude of the projection of the dipole moment along the direction of the base-pairing interactions, and results in a concomitant decrease in the strength of base pairing.

Implications of 5-Modification of Cytosine Residues on the Stability of Synthetic *i*-Motif Conformations for DNA Nanotechnology. To broaden the scope of usage of nucleic acid *i*-motif conformations in DNA nanotechnology applications, specific knowledge and chemical insight gained from studies that elucidate the influence of modifications on the properties, structures, and energetics of base pairing in protonated cytidine nucleobase or nucleoside analogue base pairs, the core stabilizing functional subunits of the *i*-motif architecture, are needed. The influence of 5-modifications of Cyt on the strength of base pairing in protonated cytosine nucleobase pairs has previously been examined by our research group.[41-44] Both the TCID measured and theoretically predicted BPEs indicate that 5-methylation of the Cyt residues of the (Cyt)H$^+$(Cyt) base pair results in strengthening of the base-pairing interactions, whereas 5-halogenation weakens the base pairing. These results suggest that 5-methylation increases the stability of nucleic acid *i*-motif conformations, and 5-halogenation decreases their stability. However, the 2′-deoxyribose and ribose moieties have been shown to exhibit a differential effect

on the strength of base pairing, as found for 5-methylation of the Cyt residues on the strength of base pairing in protonated DNA vs. RNA cytosine nucleoside base pairs.[48] Notably, the strength of base pairing is very similar for the (dCyd)H^+(dCyd) and (Cyd)H^+(Cyd) base pairs, 167.7 ± 5.3 kJ/mol (TCID) and 167.4–179.2 kJ/mol (theory) vs. 166.9 ± 4.5 kJ/mol (TCID) and 166.4–179.1 kJ/mol (theory), respectively. However, while the increase in stability resulting from 5-methylation is only modest, it is more than twice as large for the DNA than RNA protonated cytidine base pair, 8.9 ± 8.0 kJ/mol (TCID) and 7.5–12.0 kJ/mol (theory) vs. 3.9 ± 6.5 kJ/mol (TCID) and 3.4–6.7 kJ/mol (theory). These results suggest that canonical DNA *i*-motif conformations should be more stable than analogous RNA *i*-motif conformations consistent with solution-phase reports where DNA was found to form more stable *i*-motif conformations than its RNA counterparts.[97,98] These results further suggest that 5-methylation of the Cyt residues provides greater stabilization to DNA than RNA *i*-motif conformations.

Consistent with findings for 5-methylation, present results find that the influence of 5-halogenation of the Cyt residues differs for the DNA vs. RNA cytidine nucleoside analogue base pairs. 5-Halogenation is found to increase the stability of the (dCyd)H^+(dCyd) base pair. The increase in stability is greatest for 5-fluorination, which enhances the BPE by 11.4 ± 7.1 kJ/mol (TCID) and 9.9–10.6 kJ/mol (theory). The increase in stability is the least for 5-chlorination, 0.8 ± 7.0 kJ/mol (TCID) and –0.6–2.1 kJ/mol (theory) and increases with increasing size of the 5-halogen substituent, by 5.0 ± 7.2 kJ/mol and 3.4–8.3 (theory) for 5-bromination, and 12.6 ± 6.8 kJ/mol and 10.3–14.8 kJ/mol (theory) for 5-iodination. In contrast, 5-halogenation is found to decrease the stability of the (Cyd)H^+(Cyd) base pair. The influence of 5-fluorination is again the most significant and the BPE decreases by 6.5 ± 6.1 kJ/mol (TCID) and 3.4–9.5 (theory). The influence of 5-halogenation on the BPE falls off with increasing size of the 5-halogen substituent, with decreases of 4.8 ± 6.0 kJ/mol (TCID) and 4.9–14.2 kJ/mol (theory) for 5-chlorination, and 3.6 ± 7.2 kJ/mol and 2.4–5.8 (theory) for 5-bromination. However, a slight increase in stability is found for 5-iodination, which increases the BPE by 1.9 ± 6.3 kJ/mol and –0.3–7.9 kJ/mol (theory).

Overall, these studies suggest that 5-modifications, and in particular 5-methylation, 5-fluorination, and 5-iodination, should be effective means of enhancing the stability of DNA *i*-motif conformations. In contrast, 5-modifications generally reduce the stability of RNA *i*-motif conformations except for 5-iodination and 5-methylation, which produce a minor enhancement in the strength of base pairing. Trends in the measured BPEs for the (*x*Cyd)H$^+$(*x*Cyd) nucleoside base pairs examined thus far indicate that both dipole and polarizability effects of 5-modifications contribute to their influence on the strength of base pairing. Further, the 2′-hydroxy substituents are found to play a significant role in the differential influence of 5-modifications.

The influence of 5-methylation on *i*-motif stability predicted based on our model studies of protonated cytosine nucleobase and cytidine nucleoside analogue base pairs is consistent with findings of solution phase studies where 5-methylation was also found to increase the stability of DNA *i*-motif conformations.[33,34] However, studies of the influence of 5-halogenation on *i*-motif stability have not yet been reported. Studies of base pairing in protonated cytosine nucleobase pairs suggested that 5-halogenation would destabilize *i*-motif conformations, whereas present results suggest that the sugar moieties alter the influence of 5-modifications, and in particular, the 2′-deoxyribose sugar produces more dramatic effects. In addition to stabilizing *i*-motif conformations, 5-methylation was also found to increase the pK_a of *i*-motif conformations and thus shift the pH dependent conditions that govern its formation.[99] Although the influence of 5-halogenation on the pK_a for DNA and RNA cytidine nucleoside base pair formation was not explicitly examined here, our calculations find that 5-modifications do alter the proton affinities (PAs) of dCyd and Cyd (see Table 4). Further, the influences of 5-halogenation vs. 5-methylation shift the PA in opposite directions; 5-halogenation reduces the PA and 5-methylation increases the PA. These results suggest that 5-halogenation should also alters the pH dependent conditions that govern *i*-motif conformations and likely in a direction opposite to that found for 5-methylation. Overall, our findings suggest that 5-modifications of Cyt provide both nature and DNA nanotechnology a strong handle for manipulating the stability of protonated cytidine nucleoside analogue base pairs and DNA *i*-motif conformations.

CONCLUSIONS

DNA nanotechnology has been employed to develop artificial molecular motors and functional devices based on *i*-motif structures. The protonated cytosine-cytosine base pairs responsible for the stabilization of *i*-motif conformations are favored under mildly acidic conditions, enabling the development of the first DNA molecular motor driven by pH changes. The ability to alter the properties of such DNA molecular motors, e.g., the stability and pH transition range, is desirable such that understanding how the stability of *i*-motif structures are influenced by modifications, and in particular, which modifications can be employed to enhance stability and/or effect the pH characteristics of the *i*-motif are of interest. Model studies of protonated cytosine nucleobase and cytidine nucleoside analogue base pairs have established that 5-methylation of Cyt residues strengthens the base-pairing interactions.[42-,47,48] Consistent with these findings, solution-phase melting of *i*-motif conformations have confirmed that 5-methylation enhances the stability of *i*-motif conformations and shifts the pK_a for its formation,[33,34] as desired for DNA nanotechnology applications. Model studies have also shown that 5-halogenation slightly weakens the base-pairing interactions in protonated cytosine nucleobase pairs.[41,42,44] To further elucidate the influence of 5-modifications of Cyt residues on its properties and base-pairing interactions, the influence of 5-halogenation of protonated cytidine nucleoside analogues base pairs, (*x*Cyd)H^+(*x*Cyd), are examined here by complimentary threshold collision-induced dissociation techniques and computational methods. The 5-halogenated forms of the canonical DNA and RNA cytidine nucleosides, 5-fluoro-2′-deoxycytidine (fl^{5}dCyd), 5-fluorocytidine (fl^{5}Cyd), 5-chloro-2′-deoxycytidine (cl^{5}dCyd), 5-chlorocytidine (cl^{5}Cyd), 5-bromo-2′-deoxycytidine (br^{5}dCyd), 5-bromocytidine (br^{5}Cyd), 5-iodo-2′-deoxycytidine (io^{5}dCyd), and 5-iodocytidine (io^{5}Cyd), are included to examine trends as a function of the identity of the halogen substituent as well as the impact of the 2'-hydroxy substituents on the binding.

Structures of the (*x*Cyd)H^+(*x*Cyd) base pairs and their dissociation products, H^+(*x*Cyd) and *x*Cyd along with theoretical estimates for their BPEs based on the ground conformations of these species, were determined from theoretical calculations. The performance of several density

function approaches and both Pople- and Dunning-type triple-zeta basis sets were examined. Excellent agreement was found between the TCID measured and BPEs calculated at the B3LYP level of theory with either People or Dunning basis sets, suggesting that the B3LYP functional describes the intrinsic properties of the noncovalent interactions in these (*x*Cyd)H^+(*x*Cyd) base pairs that dominate the stabilization of *i*-motif conformations. In contrast, the B3P86 and M062X functionals were found to systematically overestimate the strength of the base pairing.

The complementarity of the experimental measurements, theoretical calculations, and statistical data analyses provides accurate and validated thermochemical data and expanding insight into the role of modifications on the strength of noncanonical binding relevant to synthetic *i*-motif structures. Comparisons among these model base pairs and to the analogous canonical DNA and RNA cytidine nucleoside base pairs and their 5-methylated analogues previously studied enable the influence of 5-halogenation (vs. 5-methylation) and the 2′-deoxyribose vs. ribose sugar moieties on the base-pairing interactions to be determined. 5-Halogenation of the cytosine nucleobases is found to enhance the strength of the base-pairing interactions of the DNA base pairs, and generally weakens the base-pairing interactions of the RNA base pairs. Trends in the base-pairing energies as a function of the size of the halogen substituent and the presence or absence of the 2′-hydroxy substituents indicate that both inductive and polarizability effects are at play. Overall, combined previous and present results suggest that 5-modifications, and in particular, 5-methylation, 5-fluorination and 5-iodination should be effective means of stabilizing DNA *i*-motif conformations for applications in nanotechnology, whereas only 5-iodination is weakly stabilizing for RNA *i*-motif conformations. Further, the influence of 5-halogenation on the proton affinities of the cytidine nucleoside analogues suggests that modification of Cyt residues at the 5-position may also enable tuning of the pH response of *i*-motif conformations. Results of this study provide insights that can be employed to guide the design of DNA sequences that enable dynamic control of structural transformations based on pH dependent *i*-motif formation that can be used to drive nanomechanical devices.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs/acs.org/doi/10.1021/jasms.***

Tables of average internal energies, vibrational frequencies, rotational constants, and geometric parameters, of the B3LYP/6-311+G(d,p) ground conformers of the (*x*Cyd)H^+(*x*Cyd) base pairs and the neutral and N3- and O2-protonated forms of the *x*Cyd nucleosides. Figures comparing the energy-dependent CID cross sections for interaction of the (*x*Cyd)H^+(*x*Cyd) base pairs with Xe, and fits to the CID cross sections using eq 3; stable low-energy B3LYP/6-311+G(d,p) conformations of the (*x*Cyd)H^+(*x*Cyd) base pairs and the neutral and N3- and O2-protonated forms of the *x*Cyd nucleosides, and relative B3LYP/6-311+G(2d,2p) Gibbs energies at 298 K; PES for dissociation of the (*x*Cyd)H^+(*x*Cyd) base pairs to produce the N3- and O2-protonated forms of the *x*Cyd nucleosides; polar plots of the pseudorotation phase angles, 5′-hydroxy orientations, and glycosidic bond angles of the low-energy conformations of the (*x*Cyd)H^+(*x*Cyd) base pairs and the neutral and N3- and O2-protonated forms of the *x*Cyd nucleosides; ESP and NCI maps superimposed on the ground conformations of the (*x*Cyd)H^+(*x*Cyd) base pairs and the corresponding reduced electron density gradient scatter maps; comparison of the B3P86 and M06-2X predicted versus TCID measured BPEs of the (*x*Cyd)H^+(*x*Cyd) base pairs (PDF).

AUTHOR INFORMATION

Corresponding Author

M. T. Rodgers – *Department of Chemistry, Wayne State University, Detroit, Michigan, 48202, United States;* orcid.org/0000-0002-5614-0948; Phone (313) 577-2431; Email: mrodgers@chem.wayne.edu

Authors

Yakubu S. Seidu – *Department of Chemistry, Wayne State University, Detroit, Michigan, 48202, United States;* orcid.org/0000-0002-8347-7263

E. Israel - *Department of Chemistry, Wayne State University, Detroit, Michigan, 48202, United States;* orcid.org/0000-0002-5563-9720

ACKNOWLEDGEMENTS

Financial support for this work was provided by the National Science Foundation, grant CHE-1709789. Computational resources were provided by Wayne State University C&IT. The authors acknowledge support from Wayne State University via a Thomas C. Rumble Graduate Research Fellowship for YSS.

REFERENCES

1. Seeman, N. C. Nucleic Acid Junctions: Building Blocks for Genetic Engineering in Three Dimensions. Biomolecular Stereodynamics. Ed. R. H. Sarma (New York: Adenine Press) 1981, pp. 269–77.
2. Seeman, N. C. Nucleic Acid Junctions and Lattices. *J. Theor. Biol*. **1982**, *99*, 237–247.
3. Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
4. Chen. J.; Seeman, N. C. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature* **1991**, *350*, 631–633.
5. Zhang, Y.; Seeman, N. C. The Construction of a DNA Truncated Octahedron. *J. Am. Chem. Soc.* **1994**, *116*, 1661–1669
6. Rothermund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 1–82.
7. Iinuma, R.; Yonggang, K. E.; Jungman, R.; Schlichthaerle, T.; Woehrstein, J.B.; Yin, P. Polyhedra Self-Assembled from DNA Tripods and Characterized with 3D DNA-PAINT. *Science* **2014**, *344*, 65–69.
8. Lin, C.; Liu, Y.; Rinker, S.; Yan, H. DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures. *ChemPhysChem.* **2006**, *7*, 1641–1647.
9. Hong, F.; Zhang, F.; Liu, Y.; Yan, H. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem. Rev.* **2017**, *117*, 12584–12640.
10. Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297–302.
11. Sacca, B.; Niemeyer, C. M. DNA origami: The Art of Folding DNA. *Angew. Chem. Int. Ed.* **2011**, *51*, 58–66.
12. Tapio, K.; Bald, I. The Potential of DNA Origami to Build Multifunctional Materials. *Multifunct. Mater.* **2020**, *3*, 032001.
13. Zhang, Y.; Seeman, N. C. A Solid-Support Methodology for the Construction of Geometrical Objects from DNA. *J. Am. Chem. Soc.* **1992**, *114*, 2656–2663.
14. Zhu, D.; Pei, H.; Yao G.; Wang, L.; Su, S.; Chao, J.; Wang, L.; Aldalbahi, A.; Song, S.; Shi, J.; Hu J.; Fan, C.; Zuo, Z. A Surface-Confined Proton-Driven DNA Pump Using a Dynamic 3D DNA Scaffold. *Adv. Mater.* **2016**, *28*, 6860–6865.
15. Kim, S. H.; Kim, K.-R.; Ahn, D.-R.; Lee, J. E.; Yang, E. G.; Kim, S. Y. Reversible Regulation of Enzyme Activity by pH-Responsive Encapsulation in DNA Nanocages. *ACS Nano* **2017**, *11*, 9352–9359.
16. Majikes, J. M.; Ferraz, L. C. C.; LaBean, T. H. pH-Driven Actuation of DNA Origami via Parallel *i*-Motif Sequences in Solution and on Surfaces. *Bioconjug. Chem.* **2017**, *28*, 1821–1825.
17. Peng, P.; Du, Y.; Zheng, J.; Wang, H.; Li, T. Reconfigurable Bioinspired Framework Nucleic Acid Nanoplatform Dynamically Manipulated in Living Cells for Subcellular Imaging. *Angew. Chem. Int. Ed.* **2019**, *58*, 1648–1653.
18. Zhou, Y.-J.; Wan, Y.-H.; Nie, C.-P.; Zhang, J.; Chen, T.-T.; Chu, X. Molecular Switching of a Self-Assembled 3D DNA Nanomachine for Spatiotemporal pH Mapping in Living Cells. *Anal. Chem.* **2019**, *91*, 10366–10370.
19. Deluca, M.; Shi, Z.; Castro, C. E.; Arya, G. Dynamic DNA Nanotechnology.: Toward Functional Nanoscale Devices. *Nanoscale Horiz.* **2020**, *5*, 182–201.

20. Douglas, S. M.; Bachelet, I.; Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012**, *335*, 831–834.
21. Linko, V.; Ora, A.; Kostiainen, M. A. DNA Nanostructures as Smart Drug-Delivery Vehicles and Molecular Devices. *Trends Biotechnol.* **2015**, *33*, 586–594.
22. Choi Y.; Kotthoff L.; Olejko L.; Resch-Genger U.; Bald, I. DNA Origami-Based Förster Resonance Energy-Transfer Nanoarrays and Their Application as Ratiometric Sensors. *ACS Appl. Mater. Interfaces* **2018**, *10*, 23295–23302.
23. Ramezani, J.; Dietz, H. Building Machines with DNA Molecules. *Nat. Rev. Genet.* **2020**, *21*, 5–26.
24. Liu, X.; Lu, C.H.; Willner, I. Switchable Reconfiguration of Nucleic Acid Nanostructures by Stimuli-Responsive DNA Machines. *Acc. Chem. Res.* **2014**, *47*, 1673–1680.
25. Yao, G.; Li, J.; Chen, X.; Liu, Z.; Wang, F.; Qu, Z.; Ge, Z.; Narayanan, R. P.; Williams, D.; Pei, H.; Zuo, X.; Wang, L.; Yan, H.; Feringa, B. L.; Fan, C. Programming Nanoparticle Valence Bonds with Single-Stranded DNA Encoders. *Nat. Mater.* **2020**, *19*, 781–788.
26. Dong, Y.; Yang, Z.; Liu, D. DNA Nanotechnology Based on *i*-Motif Structures. *Accts. Chem. Res.* **2014**, *47*, 1853–1860.
27. Saha, S.; Krishnan, Y. pH Sensitive DNA Devices. In *DNA Congugates and Sensors*; Fox, K.R.; Brown, T.; Eds; *RSC Biomolecular Sciences* **2012**, 166–189.
28. Day, H. A.; Pavlou, P.; Waller, Z. A. E. *i*-Motif DNA: Structure, Stability and Targeting with Ligands. *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418.
29. Zeraati, M.; Langley, D. B.; Schofield, P.; Moye, A. L.; Rouet, R.; Hughes, W. E.; Bryan, T. M.; Dinger, M. E.; Christ, D. *i*-Motif DNA Structures are Formed in the Nuclei of Human Cells. *Nat. Chem.* **2018**, *10*, 631–637.
30. Tang, W.; Niu, K.; Yu, G.; Jin, Y.; Zhang, X.; Peng, Y.; Chen, S.; Deng, H.; Li, S.; Wang, J. In Vivo Visualization of the *i*-Motif DNA Secondary Structure in the Bombyx Mori Testis. *Epigenetics Chromatin* **2020**, *13*, 1–12.
31. Shu, W.; Liu, D.; Watari, M.; Riener, C. K.; Strunz, T.; Welland, M. E.; Balasubramanian, S.; McKendry, R. A. DNA Molecular Motor Driven Micromechanical Cantilever Arrays. *J. Am. Chem. Soc.* **2005**, *127*, 17054–17060.
32. McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent Progress in Non-native Nucleic Acid Modifications. *Chem. Soc. Rev*. **2021**, ***50***, 5126–5164.
33. Xu, B.; Devi, G.; Shao, F. Regulation of Telomeric *i*-Motif Stability by 5-Methylcytosine and 5-Hydroxymethylcytosine Modification. *Org. Biomol. Chem.* **2015**, *13*, 5646–5651.
34. Školáková, P.; Badri, Z.; Foldynová-Trantírková, S.; Ryneš, J.; Šponer, J.; Fojtová, M.; Fajkus, J.; Marek, R.; Vorlíčková, M.; Mergny, J.-L. Composite 5-Methylations of Cytosines Modulate *i*-Motif Stability in a Sequence-Specific Manner: Implications for DNA Nanotechnology and Epigenetic Regulation of Plant Telomeric DNA. *Biochim. Biophys. Acta Gen. Subj.* **2020**, *1864*, 129651.
35. Pérez, A.; Castellazzi, C. L.; Battistini, F.; Collinet, K.; Flores, O.; Deniz, O.; Ruiz, M. L.; Torrents, D.; Eritja, R.; Soler-López, M. Impact of Methylation on the Physical Properties of DNA. *Biophys. J.* **2012**, *102*, 2140–2148.
36. Breiling, A.; Lyko, F. Epigenetic Regulatory Functions of DNA Modifications: 5-Methylcytosine and Beyond. *Epigenetics Chromatin* **2015**, *8*, 1–9.
37. Karabıyık, H.; Sevinçek, R.; Karabıyık, H. π-Cooperativity Effect on the Base Stacking Interactions in DNA: Is There a Novel Stabilization Factor Coupled with Base Pairing H-Bonds? *Phys. Chem. Chem. Phys.* **2014**, *16*, 15527–15538.

38. Lieblein, A. L.; Krämer, M.; Dreuw, A.; Fürtig, B.; Schwalbe, H. The Nature of Hydrogen Bonds in Cytidine···H^+···Cytidine DNA Base Pairs. *Angew. Chem. Int. Ed.* **2012**, *51*, 4067–4070.
39. Berger, I.; Egli, M.; Rich, A. Inter-Strand CH···O Hydrogen Bonds Stabilizing Four-Stranded Intercalated Molecules: Stereoelectronic Effects of O4′ in Cytosine-Rich DNA. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12116–12121.
40. Mitas, M. Trinucleotide Repeats Associated with Human Disease. *Nucleic Acids Res.* **1997**, *25*, 2245–2253.
41. 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.
42. 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.
43. Yang, B.; Moehlig, A. R.; Frieler, C. E.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Dimers and Proton Affinities of 1-Methylated Cytosines: Model Systems of the Effects of the Sugar Moiety on the Stability of DNA *i*-Motif Conformations. *J. Phys. Chem. B* **2015**, *119*, 1857–1868.
44. Yang, B.; Wu, R. R.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Dimers and Proton Affinities of 1-Methyl-5-Halocytosines: Implications for the Effects of Halogenation on the Stability of the DNA *i*-Motif. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1469–1482.
45. Moehlig, A. R.; Djernes, K. E.; Krishnan, V. M.; Hooley, R. J. Cytosine Derivatives Form Hemiprotonated Dimers in Solution and the Gas Phase. *Org. Lett.* **2012**, *14*, 2560–2563.
46. Abou Assi, H.; Lin, Y. C.; Serrano, I.; González, C.; Damha, M. J. Probing Synergistic Effects of DNA Methylation and 2'-β-Fluorination on *i*-Motif Stability. *Chem. Eur. J.* **2018**, *24*, 471–477.
47. Yang, B.; Rodgers, M. T. Base-Pairing Energies of Protonated Nucleoside Base Pairs of dCyd and m^5dCyd: Implications for the Stability of DNA *i*-Motif Conformations. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1394–1403.
48. Seidu, Y. S.; Roy, H. A.; Rodgers, M. T. Influence of 5-Methylation and the 2′ and 3′-Hydroxy Substituents on the Base-Pairing Energies of Protonated Cytidine Nucleoside Analogue Base Pairs: Implications for the Stabilities of *i*-Motif Structures. *J. Phys. Chem. A* **2021**, *125*, 5939–5955.
49. 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.
50. Chen, Y.; Rodgers, M.T. Structural and Energetic Effects in the Molecular Recognition of Protonated Peptidomimetic Bases by 18-Crown-6. *J. Am. Chem. Soc.* **2011**, *134*, 2313–2324.
51. Coates, R. A.; Armentrout, P. B. Thermochemical Investigations of Hydrated Nickel Dication Complexes by Threshold Collision-Induced Dissociation and Theory. *J. Phys. Chem. A* **2017**, *121*, 3629–3646.
52. Teloy, E.; Gerlich, D. Integral Cross Sections for Ion-Molecule Reactions I. The Guided Ion Beam Technique. Chem. Phys. **1974**, *4*, 417–427.
53. 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**.
54. Gerlich, D. State-Selected and State-to-State Ion-Molecule Reaction Dynamics, Part I,

Experiment, Ng, C.-Y.; Baer, M. Eds.; *Adv. Chem. Phys*. **1992**, *82*, 1–176.
55. 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.
56. 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.
57. Aristov, N.; Armentrout, P. B. Collision-Induced Dissociation of Vanadium Monoxide Ion. *J. Phys. Chem.* **1986**, *90*, 5135–5140.
58. Dalleska, N. F.; Honma, K.; Sunderlin, L. S.; Armentrout, P. B. Solvation of Transition Metal Ions by Water. Sequential Binding Energies of $M^+(H_2O)_x$ (x = 1–4) for M = Ti to Cu Determined by Collision-Induced Dissociation. *J. Am. Chem. Soc.* **1994**, *116*, 3519–3528.
59. Wolinski, K.; Hinton, J.; Wishart, D.; Sykes, B.; Richards, F.; Pastone, A.; Saudek, V.; Ellis, P.; Maciel, G.; McIver, J.: HyperChem Computational Chemistry Software Package, Version 5.0, Hypercube *Inc*: Gainsville, FL, **1997**.
60. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford, CT, **2009**.
61. Boys, S. F., Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553–566.
62. 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.
63. Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. New Description Using the Concept of Pseudorotation. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.
64. 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.
65. Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.
66. Contreras-García, J.; Yang, W.; Johnson E. R. Analysis of Hydrogen-Bond Interaction Potentials from the Electron Density: Integration of Noncovalent Interaction Regions. *J. Phys. Chem. A* **2011**, *115*, 12983–12990.
67. Boto, R. A.; Peccati, F.; Laplaza, R.; Quan, C.; Carbone, A.; Piquemal, J.-P.; Maday, Y.; Contreras-García, J. NCIPLOT4: Fast, Robust, and Quantitative Analysis of Noncovalent Interactions. *J. Chem. Theory Comput*. **2020**, *16*, 4150–4158.

68. Bader, R. F. W. Atoms in Molecules: A Quantum Theory. Clarendon Press, Oxford (1990)
69. Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592.
70. Humphrey, W.; Dalke, A.; Schulten, K.: VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
71. Ervin, K. M.; Armentrout, P. B. Translational Energy Dependence of $Ar^+ + XY \rightarrow ArX^+ + Y$ (XY= H_2, D_2, HD) from Thermal to 30 eV cm. *J. Chem. Phys.* **1985**, *83*, 166–189.
72. Hales, D. A.; Lian, L.; Armentrout, P. B. Collision-Induced Dissociation of Nb_n^+ (n = 2 - 11): Bond Energies and Dissociation Pathways. *Int. J. Mass Spectrom. Ion Processes* **1990**, *102*, 269-301.
73. Schultz, R. H.; Crellin, K. C.; Armentrout, P. B. Sequential Bond Energies of Iron Carbonyl $Fe(CO)_x^+$ (x = 1–5): Systematic Effects on Collision-Induced Dissociation Measurements. *J. Am. Chem. Soc.* **1991**, *113*, 8590–8601.
74. Chantry, P. Doppler Broadening in Beam Experiments. *J. Chem. Phys.* **1971**, *55*, 2746–2759.
75. 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.
76. Rodgers, M. T.; Ervin, K. M.; Armentrout, P. B. Statistical Modeling of Collision-Induced Dissociation Thresholds. *J. Chem. Phys.* **1997**, *106*, 4499–4508.
77. 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.
78. 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.
79. Beyer, T.; Swinehart, D. Algorithm 448: Number of Multiply-Restricted Partitions. *Commun. ACM*, **1973**, *16*, 379.
80. 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.
81. 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 Base-Pairing Interactions. *J. Phys. Chem. B* **2013**, *117*, 14191–14201.
82. Wu, R. R.; Rodgers, M. T. O2 Protonation Controls Threshold Behavior for N-Glycosidic Bond Cleavage of Protonated Cytosine Nucleosides. *J. Phys. Chem. B* **2016,** *120*, 4803–4811.
83. Wu, R. R.; Yang, B.; Berden, G.; Oomens, J.; Rodgers, M. T. N3 and O2 Protonated Tautomeric Conformations of 2'-Deoxycytidine and Cytidine Coexist in the Gas Phase. *J. Phys. Chem. B* **2015**, *119*, 5773–5784.
84. He, C. C.; Hamlow, L. A.; Devereaux, Z. J.; Zhu, Y.; Nei, Y.-w.; Fan, L.; McNary, C.; Maître, P.; Steinmetz, V.; Schindler, B.; Armentrout, P. B.; Rodgers, M. T. Structural and Energetic Effects of 2′-Ribose Methylation of Protonated Pyrimidine Nucleosides. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2318–2334.
85. Devereaux, Z. J.; Roy, H. A.; He, C. C.; Zhu, Y.; Cunningham, N. A.; Hamlow, L. A.; Berden, G.; Oomens, J.; Rodgers, M. T. Influence of 2′-Fluoro Modification on Glycosidic Bond Stabilities and Gas-Phase Ion Structures of Protonated Pyrimidine Nucleosides. *J. Fluorine Chem.* **2019**, *219*, 10–22.
86. Hamlow, L. A.; He, C. C.; Devereaux, Z. J.; Roy, H. A.; Cunningham, N. A.; Soley, E. O.; Berden, G.; Oomens, J.; Rodgers, M. T. Gas-Phase Structures of Protonated Arabino

Nucleosides. *Int. J. Mass Spectrom.* **2019**, *438*, 124–134.
87. Hamlow, L. A.; Devereaux, Z. J.; Roy, H. A.; Cunningham, N. A.; Berden, G.; Oomens, J.; Rodgers, M. T. Impact of the 2′- and 3′-Sugar Hydroxyl Moieties on Gas-Phase Nucleoside Structure. *J. Am. Soc. Mass Spectrom*. **2019**, *30*, 832–845.
88. Rannulu, N. S.; Rodgers, M. T. Noncovalent Interactions of Cu^+ with N-Donor Ligands (Pyridine, 4,4-Dipyridyl, 2,2-Dipyridyl, and 1,10-Phenanthroline): Collision-Induced Dissociation and Theoretical Studies. *J. Phys. Chem. A* **2007**, *111*, 3465–3479.
89. Hallowita, N.; Carl, D. R.; Armentrout, P. B.; Rodgers, M. T. Dipole Effects on Cation-π Interactions: Absolute Bond Dissociation Energies of Complexes of Alkali Metal Cations to N-Methylaniline and N,N-Dimethylaniline. *J. Phys. Chem. A* **2008**, *112*, 7996–8008.
90. Vitale, G.; Valina, A. B.; Huang, H.; Amunugama, R.; Rodgers, M. T. Solvation of Copper Ions by Acetonitrile. Structures and Sequential Binding Energies of $Cu^+(CH_3CN)_x$, x =1–5 from Collision-Induced Dissociation and Theoretical Studies. *J. Phys. Chem. A* **2001**, *105*, 11351–11364.
91. Ruan, C.; Yang, Z.; Rodgers, M. T. Influence of the d Orbital Occupation on the Nature and Strength of Copper Cation–π Interactions: Threshold Collision-Induced Dissociation and Theoretical Studies. *Phys. Chem. Chem. Phys*. **2007**, *9*, 5902–5918.
92. Rannulu, N. S.; Amunugama, R.; Yang, Z.; Rodgers, M. T. Influence of s and d Orbital Occupation on the Binding of Metal Ions to Imidazole. *J. Phys. Chem. A* **2004**, *108*, 6385–6396.
93. Ruan, C.; Rodgers, M. T. Cation-π Interactions: Structures and Energetics of Complexation of Na^+ and K^+ with the Aromatic Amino Acids, Phenylalanine, Tyrosine and Tryptophan. *J. Am. Chem. Soc.* **2004***, 126*, 14600–14610.
94. Huang, H.; Rodgers, M. T. Sigma versus Pi Interactions in Alkali Metal Ion Binding Affinities of Azoles: Threshold Collision-Induced Dissociation and Ab Initio Theory Studies. *J. Phys. Chem. A* **2002**, *106*, 4277–4289.
95. Rodgers, M. T.; Armentrout, P. B. Absolute Alkali Metal Ion Binding Affinities of Several Azoles Determined by Threshold Collision-Induced Dissociation. *Int. J. Mass Spectrom*. **1999**, *187*, 359–380.
96. Ruan, C.; Yang, Z.; Hallowita, N.; Rodgers, M. T. Cation-π Interactions with a Model for the Side Chain of Tryptophan: Structures and Absolute Binding Energies of Alkali Metal Cation-Indole Complexes. *J. Phys. Chem. A* **2005**, *109*, 11539–11550.
97. Lacroix, L.; Mergny, J.-L.; Leroy, J.-L.; Hélène, C. Inability of RNA to Form the *i*-Motif: Implications for Triplex Formation. *Biochemistry* **1996**, *35*, 8715–8722.
98. Poeta, L.; Drongitis, D.; Verrillo, L.; Miano, M. G. DNA Hypermethylation and Unstable Repeat Diseases: A Paradigm of Transcriptional Silencing to Decipher the Basis of Pathogenic Mechanisms. Genes **2020**, *11*, 684.
99. Bhavsar-Jog, Y. P.; Van Dornshuld, E.; 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.

Table 1. Threshold Dissociation Energies at 0 K, Entropies of Activation at 1000 K, Fitting Parameters of Eq 2, and Kinetics Shifts of Protonated Cytidine Nucleoside Analogue Base Pairs.[a]

(*x*Cyd)H$^+$(*x*Cyd)	E_0(PSL)[b] (eV)	$\Delta S^{\ddagger}$(PSL)[b] (J·mol^{-1}K^{-1})	σ[b]	n[b]	E_0[c] (eV)	Kinetic Shift[d] (eV)
(fl^{5}dCyd)H$^+$(fl^{5}dCyd)	1.86 (0.05)	98 (3)	96 (5)	0.8 (0.05)	3.60 (0.08)	1.74
(cl^{5}dCyd)H$^+$(cl^{5}dCyd)	1.75 (0.05)	82 (3)	86 (5)	0.8 (0.03)	3.47 (0.07)	1.72
(br^{5}dCyd)H$^+$(br^{5}dCyd)	1.79 (0.05)	82 (3)	169 (6)	0.8 (0.09)	3.65 (0.08)	1.86
(io^{5}dCyd)H$^+$(io^{5}dCyd)	1.87 (0.05)	109 (3)	108 (7)	0.9 (0.05)	3.59 (0.09)	1.72
(fl^{5}Cyd)H$^+$(fl^{5}Cyd)	1.66 (0.04)	99 (3)	58 (2)	0.7 (0.03)	3.18 (0.07)	1.52
(cl^{5}Cyd)H$^+$(cl^{5}Cyd)	1.68 (0.04)	89 (3)	216 (6)	0.7 (0.04)	3.32 (0.08)	1.64
(br^{5}Cyd)H$^+$(br^{5}Cyd)	1.70 (0.06)	94 (3)	164 (9)	0.8 (0.06)	3.31 (0.09)	1.61
(io^{5}Cyd)H$^+$(io^{5}Cyd)	1.75 (0.05)	91 (3)	159 (10)	0.9 (0.08)	3.54 (0.09)	1.79

[a] Present results, uncertainties are listed in parentheses. [b] Average values obtained based on fits using a loose PSL TS. [c] No RRKM lifetime analysis included.

Table 2. Base-Pairing Energies of Protonated Cytidine Nucleoside Analogue Base Pairs at 0 K in kJ/mol[a].

(*x*Cyd)H$^+$(*x*Cyd)	TCID	Level of Theory[b]			
		B3LYP[c]	B3LYP[d]	B3P86[e]	M06-2X[f]
(fl^{5}dCyd)H$^+$(fl^{5}dCyd)	179.1 (4.8)	179.3	177.5	188.3	189.1
(cl^{5}dCyd)H$^+$(cl^{5}dCyd)	168.5 (4.5)	168.1	167.0	177.6	181.3
(br^{5}dCyd)H$^+$(br^{5}dCyd)	172.7 (4.9)	172.6	170.8	182.5	187.5
(io^{5}dCyd)H$^+$(io^{5}dCyd)	180.3 (4.3)	180.0	177.7	190.5	194.0
(fl^{5}Cyd)H$^+$(fl^{5}Cyd)	160.4 (4.1)	160.2	157.3	169.6	175.7
(cl^{5}Cyd)H$^+$(cl^{5}Cyd)	162.1 (4.0)	161.5	159.5	163.1	168.0
(br^{5}Cyd)H$^+$(br^{5}Cyd)	163.3 (5.6)	163.0	161.0	174.1	176.7
(io^{5}Cyd)H$^+$(io^{5}Cyd)	168.8 (4.4)	168.9	166.5	180.0	187.0
(dCyd)H$^+$(dCyd)[g]	167.7 (5.3)	168.7	167.4	177.7	179.2
(m^{5}dCyd)H$^+$(m^{5}dCyd)[g]	176.6 (6.0)	176.2	176.5	188.0	191.2
(Cyd)H$^+$(Cyd)[g]	166.9 (4.5)	166.4	166.8	177.3	179.1
(m^{5}Cyd)H$^+$(m^{5}Cyd)[g]	170.8 (4.7)	169.8	171.6	184.0	184.6
(Cyt)H$^+$(Cyt)[h]	169.9 (4.6)	168.9	169.2		
(fl^{5}Cyt)H$^+$(fl^{5}Cyt)[h]	162.7 (3.8)	165.7	167.2		
(br^{5}Cyt)H$^+$(br^{5}Cyt)[h]	168.5 (4.9)	161.5	162.8		
(io^{5}Cyt)H$^+$(io^{5}Cyt)[h]	163.2 (4.7)	161.1	162.8		
(m^{5}Cyt)H$^+$(m^{5}Cyt)[h]	177.4 (5.3)	173.4	173.3		
AEU/MAD[i]	4.6 (0.5)	0.3 (0.2)	2.2 (0.5)	8.8 (3.3)	13.0 (3.7)

[a]Present results except as noted, uncertainties are listed in parentheses. [b]Theoretical BPEs are computed based on structures optimized at the B3LYP/6-311+G(d,p) level of theory including ZPE and BSSE corrections. [c]Calculated at the B3LYP/6-311+G(2d,2p) level of theory. [d]Calculated at the B3LYP/def2-TVZPPD level of theory. [e]Calculated at the B3P86/6-311+G(2d,2p) level of theory. [f]Calculated at the M062X/6-311+G(2d,2p) level of theory. [g]Values taken from reference 48. [h]Values taken from reference 41. [i]Average experimental uncertainty (AEU) of the BPEs determined in the present study. Mean absolute deviation (MAD) between the measured and calculated BPEs determined in the present study.

Table 3. Enthalpies and Gibbs Energies of Base-Pairing of Protonated Cytidine Nucleoside Analogue Base Pairs at 0 and 298 K in kJ/mol[a]

(*x*Cyd)H$^+$(*x*Cyd)	ΔH_0	ΔH_0[b]	$\Delta H_{298}-\Delta H_0$[b]	ΔH_{298}	ΔH_{298}[b]	$T\Delta S_{298}$[b]	ΔG_{298}	ΔG_{298}[b]
(fl^{5}dCyd)H$^+$(fl^{5}dCyd)	179.1 (4.8)	179.3	0.3 (0.1)	179.4 (4.8)	179.6	54.7 (1.5)	124.7 (5.0)	124.9
(cl^{5}dCyd)H$^+$(cl^{5}dCyd)	168.5 (4.5)	168.1	0.5 (0.1)	169.0 (4.5)	168.6	50.2 (1.5)	118.8 (4.7)	118.4
(br^{5}dCyd)H$^+$(br^{5}dCyd)	172.7 (4.9)	172.6	0.3 (0.1)	173.0 (4.9)	172.9	50.3 (1.5)	122.7 (5.1)	122.6
(io^{5}dCyd)H$^+$(io^{5}dCyd)	180.3 (4.3)	180.0	1.0 (0.1)	181.3 (4.3)	181.0	57.4 (1.3)	123.9 (4.5)	123.6
(fl^{5}Cyd)H$^+$(fl^{5}Cyd)	160.4 (4.1)	160.2	0.6 (0.1)	161.0 (4.1)	160.8	54.5 (1.4)	106.5 (4.3)	106.3
(cl^{5}Cyd)H$^+$(cl^{5}Cyd)	162.1 (4.3)	161.5	0.1 (0.1)	162.2 (4.3)	161.6	52.4 (1.6)	109.8 (4.3)	109.2
(br^{5}Cyd)H$^+$(br^{5}Cyd)	163.3 (5.6)	163.0	0.2 (0.1)	163.5 (5.6)	163.2	50.3 (1.5)	113.2 (5.8)	112.9
(io^{5}Cyd)H$^+$(io^{5}Cyd)	168.8 (4.4)	168.9	0.2 (01)	169.0 (4.4)	169.1	52.7 (1.5)	116.3 (4.6)	116.4

[a]Present results, uncertainties are listed in parentheses. Note that the BPE is defined as the enthalpy of base pairing at 0 K.[b]Calculated at the B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) level of theory with frequencies scaled by 0.9804.

Table 4. Isotropic Molecular Polarizabilities, Dipole Moments and Enthalpies and Gibbs Energies of Proton Binding to N3 and O2 of Cytidine Nucleoside Analogues at 0 and 298 K (in kJ/mol).[a]

*x*Cyd	Polarizability (Å^3)	Dipole Moment (D)	ΔH_0(N3) (kJ/mol)	ΔH_{298}(N3) (kJ/mol)	ΔG_{298}(N3) (kJ/mol)	ΔH_0(O2) (kJ/mol)	ΔH_{298}(O2) (kJ/mol)	ΔG_{298}(O2) (kJ/mol)
fl^{5}dCyd	22.0	4.65	979.4	985.9	950.5	986.3	993.5	956.9
cl^{5}dCyd	23.7	4.59	974.4	981.6	942.4	977.3	985.2	944.9
br^{5}dCyd	25.2	4.67	982.0	988.8	951.5	984.1	991.6	953.2
io^{5}dCyd	27.7	4.99	994.9	1001.6	965.9	995.6	1002.9	966.1
fl^{5}Cyd	22.6	5.58	950.7	956.5	923.0	960.6	966.8	932.2
cl^{5}Cyd	24.7	4.91	965.6	972.1	936.7	971.5	978.4	941.9
br^{5}Cyd	25.9	6.49	968.6	975.1	939.7	973.5	980.6	944.0
io^{5}Cyd	27.8	6.74	979.7	986.8	949.5	983.1	990.7	952.3
dCyd[b]	21.6	7.26	997.5	1004.3	962.3	993.2	1000.3	957.5
m^{5}dCyd[b]	23.6	7.38	1001.9	1008.1	972.1	1001.7	1008.6	971.5
Cyd[b]	22.4	7.23	990.5	996.7	963.7	992.2	998.4	965.4
m^{5}Cyd[b]	24.2	7.69	993.1	999.2	964.7	991.8	998.2	963.7

[a]Values computed at the B3LYP/6-311+G(2d,2p)// B3LYP/6-311+G(d,p) level of theory including ZPE and BSSE corrections with frequencies scaled by 0.9804. [b]Values taken from reference 48.

FIGURE CAPTIONS

Figure 1. Cross sections for CID of the (fl^{5}dCyd)H$^+$(fl^{5}dCyd) and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs 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. Data are shown for a Xe pressure of ~0.2 mTorr.

Figure 2. Zero-pressure-extrapolated cross sections for CID of the (fl^{5}dCyd)H$^+$(fl^{5}dCyd) and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs 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 a and b. The solid lines show the best fits to the data using the model of eq 2 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 base pairs with an internal temperature of 0 K.

Figure 3. B3LYP/6-311+G(d,p) optimized geometries of the ground conformers of the (dCyd)H$^+$(dCyd), (fl^{5}dCyd)H$^+$(fl^{5}dCyd), and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs. The base pairs are oriented such that the protonated nucleoside is shown on the left. The nucleobase orientation, 5-hydroxy orientation, and sugar puckering of the protonated nucleoside is indicated first in boldface font followed by those of the neutral nucleoside in standard font. The hydrogen-bond lengths are also shown.

Figure 4. Pseudorotation phase angles (*P*), glycosidic bond angles, and 5-hydroxy orientations of the ground conformers of the (*x*Cyd)H$^+$(*x*Cyd) base pairs. Values for the *x*^{5}dCyd analogues are indicated with open symbols, whereas values for the *x*Cyd analogues are shown as closed symbols. The protonated nucleoside of each base pair is differentiated from the neutral nucleoside by a + sign over the symbol.

Figure 5. Noncovalent interaction maps superimposed on the B3LYP/6-311+G(d,p) optimized geometries of the ground conformers of the (dCyd)H$^+$(dCyd), (fl^{5}dCyd)H$^+$(fl^{5}dCyd), and (io^{5}dCyd)H$^+$(io^{5}dCyd) base pairs. The NCI maps are shown at an isosurface of 0.5 a.u. of the B3LYP/6-311+G(2d,2p) reduced electron density gradients. Regions exhibiting strong attractive interactions such as ionic or strong neutral hydrogen-bonding interactions appear blue. Weaker attractive interactions such as London dispersion or noncanonical hydrogen-bonding interactions appear green. Highly repulsive interactions appear red. Additional shading is used to highlight interactions between N4H and the 5-halogen substituents.

Figure 6. B3LYP calculated vs. TCID measured BPEs of the (*x*Cyd)H^+(*x*Cyd) base pairs at 0 K (in kJ/mol), where *x*Cyd = fl^5dCyd, cl^5dCyd, br^5dCyd, io^5dCyd, fl^5Cyd, cl^5Cyd, br^5Cyd, and io^5Cyd. Results for the 6-311+G(2d,2p) and def2-TZVPPD basis sets are compared. Calculated values include ZPE and BSSE corrections. The diagonal line indicates values for which calculated and measured BDEs are in perfect agreement.

Figure 7. TCID measured BPEs of the (*x*Cyd)H^+(*x*Cyd) base pairs at 0 K (in kJ/mol) versus the neutral *x*Cyd nucleosides: m^5dCyd, m^5Cyd, dCyd, Cyd, fl^5dCyd, fl^5Cyd, cl^5dCyd, cl^5Cyd, br^5dCyd, br^5Cyd, io^5dCyd, and io^5Cyd. Data for the base pairs involving dCyd, Cyd, m^5dCyd and m^5Cyd is taken from reference 48.

Figure 8. TCID measured BPEs of the (*x*Cyd)H^+(*x*Cyd) base pairs at 0 K (in kJ/mol) versus the PBE0/6-311+G(2d,2p) computed isotropic molecular polarizability volumes (α in $Å^3$) of the neutral *x*Cyd nucleosides. The solid line is a linear regression fit to the data for the base pairs involving cl^5Cyd, br^5Cyd, and io^5Cyd, whereas the dashed line is a linear regression fit to the data for the base pairs involving cl^5dCyd, br^5dCyd, and io^5dCyd. Data for the base pairs involving dCyd, Cyd, m^5dCyd and m^5Cyd is taken from reference 48.

Figure 9. TCID measured BPEs of the (*x*Cyd)H^+(*x*Cyd) base pairs at 0 K (in kJ/mol) versus the projection of the dipole moment (in D) of the neutral *x*Cyd nucleosides along the direction of the base pairing. The solid line is a linear regression fit to the data. Data for the base pairs involving dCyd, Cyd, m^5dCyd and m^5Cyd is taken from reference 48.

Figure 1.

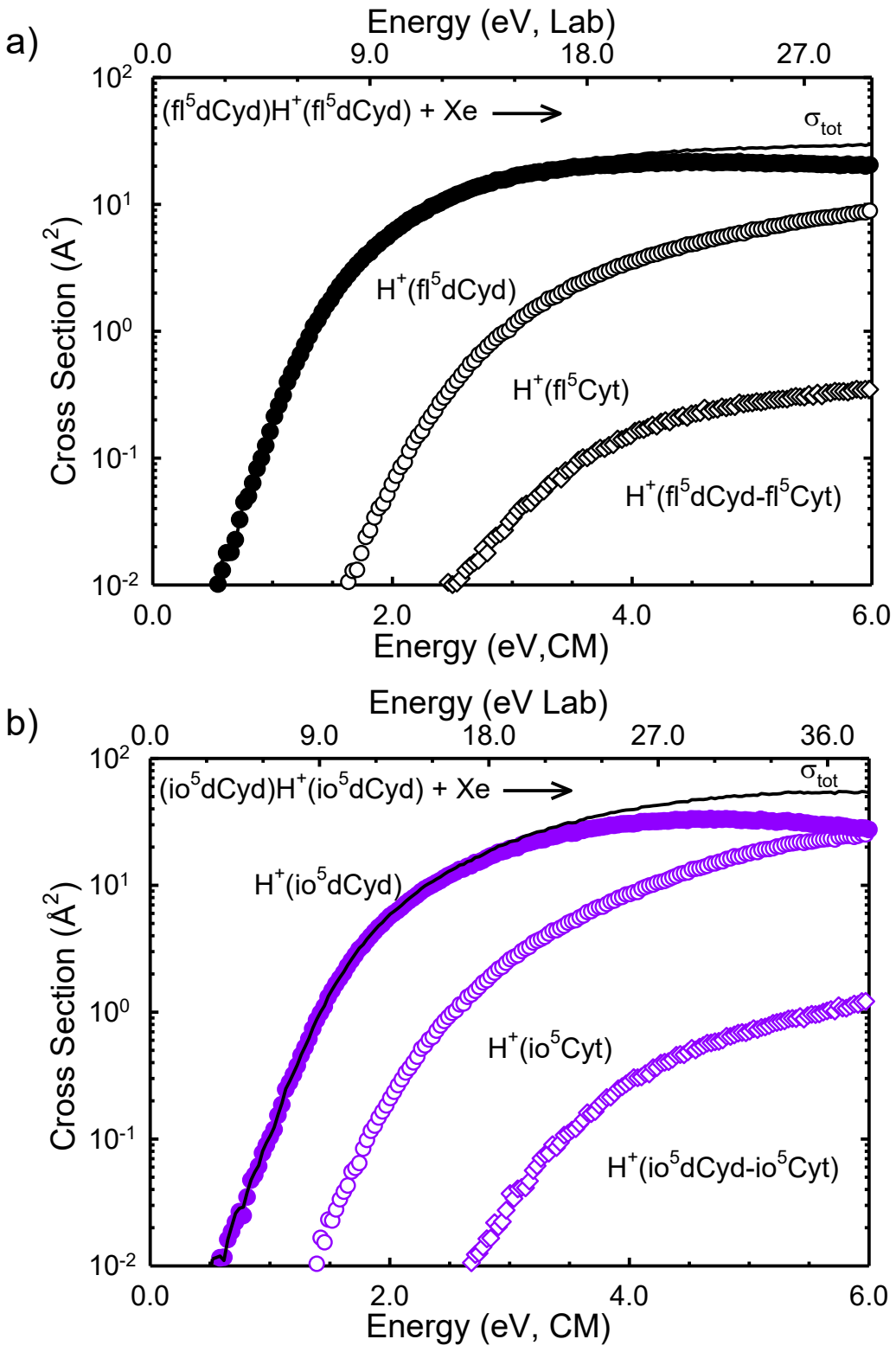

a)
Energy (eV, Lab)
0.0
9.0
18.0
27.0
(fl5dCyd)H+(fl5dCyd) + Xe
σtot
H+(fl5dCyd)
H+(fl5Cyt)
H+(fl5dCyd-fl5Cyt)
Cross Section (Å2)
10-2
10-1
100
101
102
0.0
2.0
4.0
6.0
Energy (eV,CM)
b)
Energy (eV Lab)
0.0
9.0
18.0
27.0
36.0
(io5dCyd)H+(io5dCyd) + Xe
σtot
H+(io5dCyd)
H+(io5Cyt)
H+(io5dCyd-io5Cyt)
Cross Section (Å2)
10-2
10-1
100
101
102
0.0
2.0
4.0
6.0
Energy (eV, CM)

Figure 2.

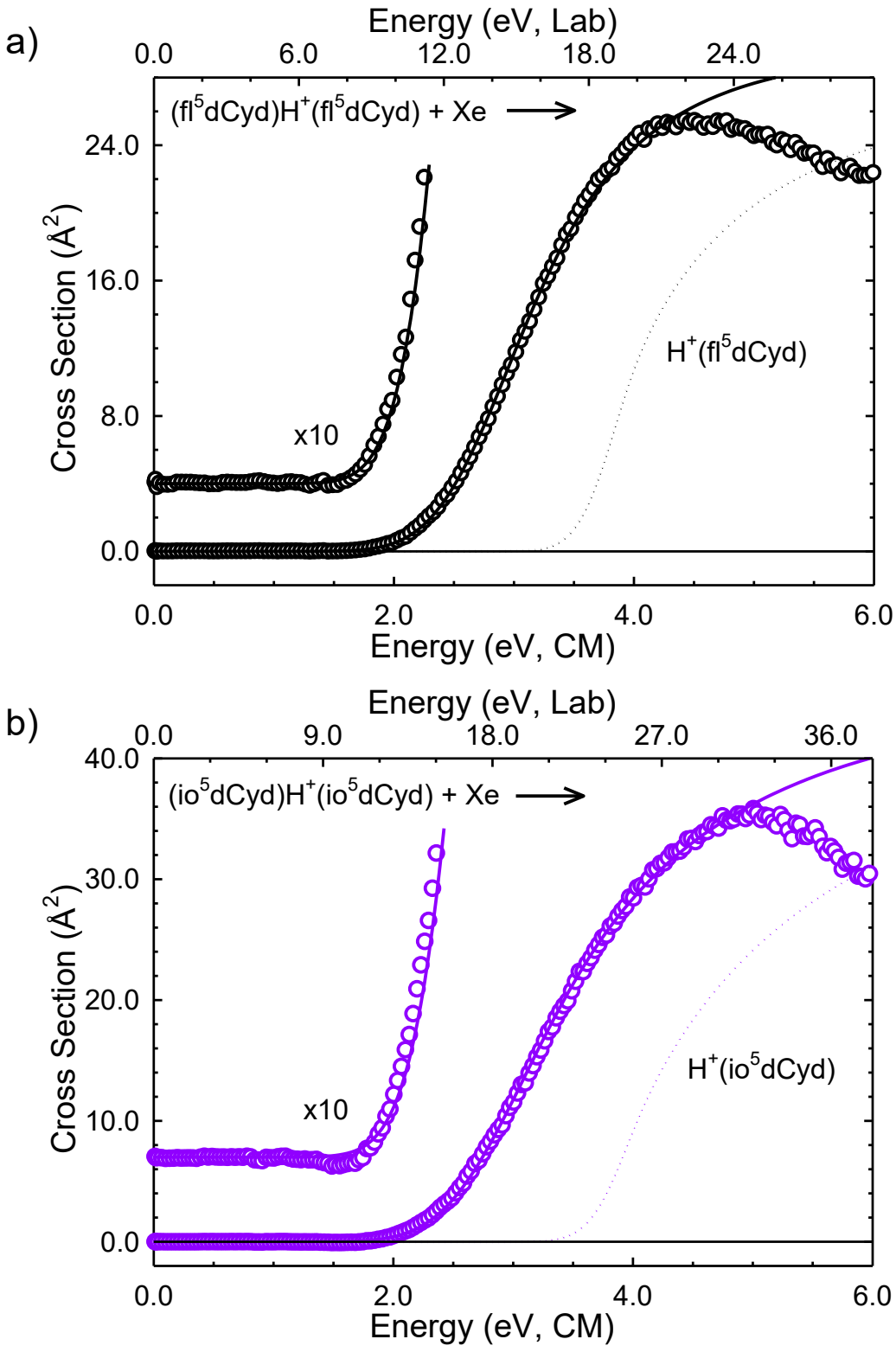
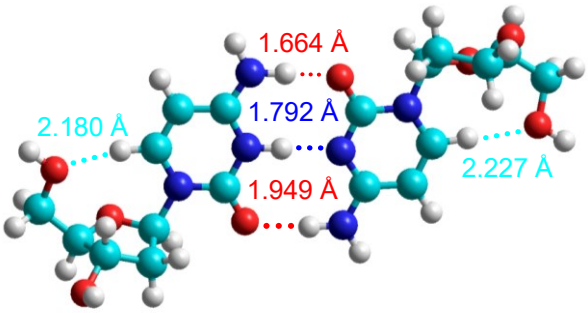

a)
Energy (eV, Lab)
0.0
6.0
12.0
18.0
24.0
(fl5dCyd)H+(fl5dCyd) + Xe
Cross Section (Å2)
24.0
16.0
8.0
0.0
x10
H+(fl5dCyd)
0.0
2.0
4.0
6.0
Energy (eV, CM)
b)
Energy (eV, Lab)
0.0
9.0
18.0
27.0
36.0
(io5dCyd)H+(io5dCyd) + Xe
Cross Section (Å2)
40.0
30.0
20.0
10.0
0.0
x10
H+(io5dCyd)
0.0
2.0
4.0
6.0
Energy (eV, CM)

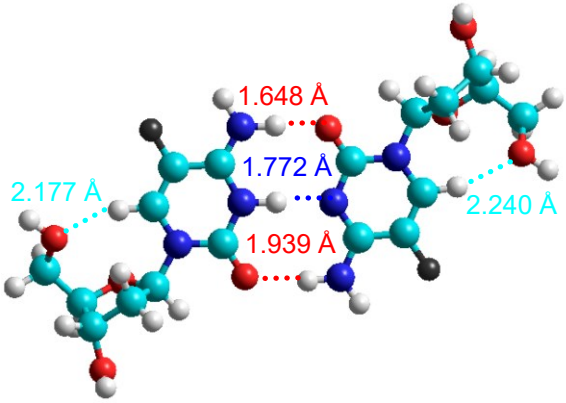
Figure 3.



(dCyd)H$^+$(dCyd)

***anti*, gauche$^+$**·*anti*, gauche$^+$

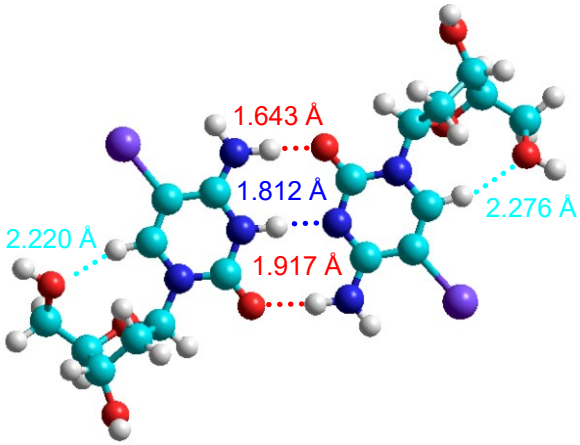
C3′-endo·C3′-endo (**^{3}T$_2$**· ^{3}T$_2$)



(fl^{5}dCyd)H$^+$(fl^{5}dCyd)

***anti*,gauche$^+$**·*anti*,gauche$^+$

C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)

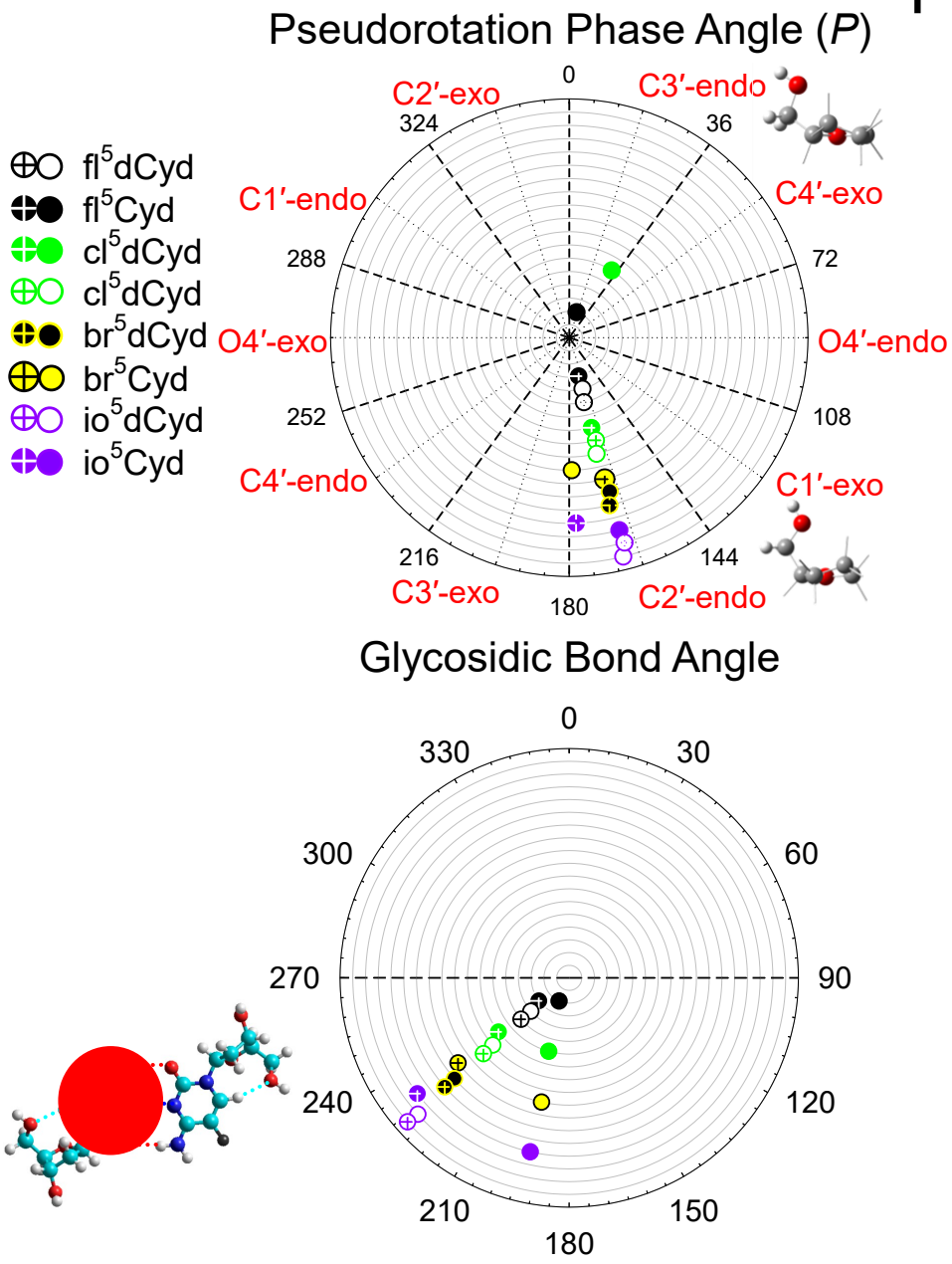


(io^{5}dCyd)H$^+$(io^{5}dCyd)

***anti*,gauche$^+$**·*anti*,gauche$^+$

C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)

Figure 4.


Pseudorotation Phase Angle (P)
fl5dCyd
fl5Cyd
cl5dCyd
cl5dCyd
br5dCyd
br5Cyd
io5dCyd
io5Cyd
0
36
72
108
144
180
216
252
288
324
C2′-exo
C3′-endo
C4′-exo
O4′-endo
C1′-exo
C2′-endo
C3′-exo
C4′-endo
C1′-endo
O4′-exo

Glycosidic Bond Angle
0
30
60
90
120
150
180
210
240
270
300
330

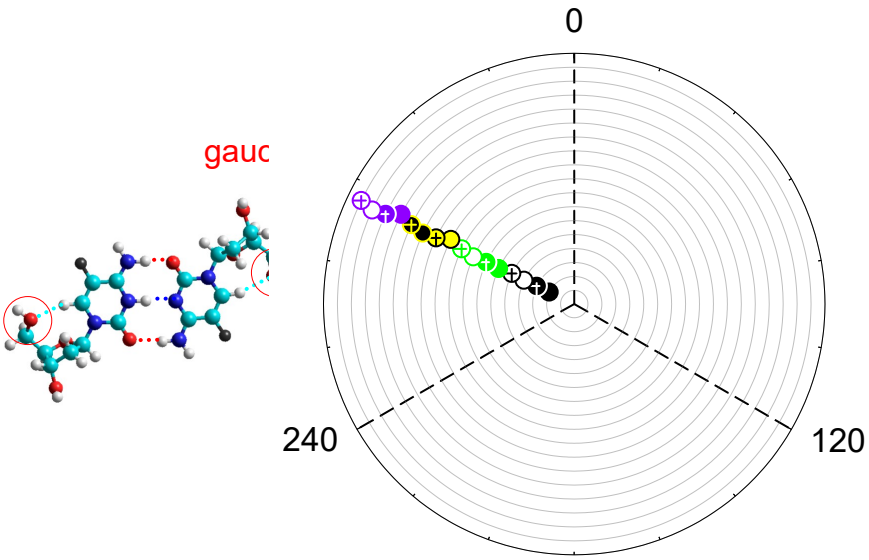
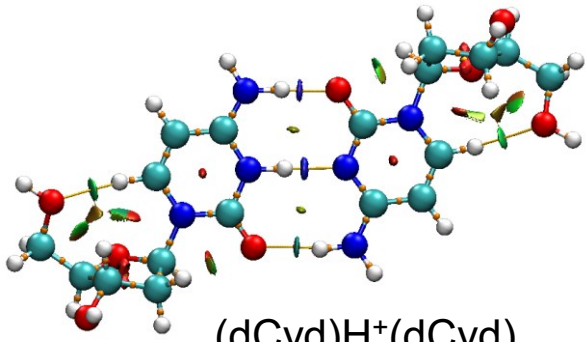
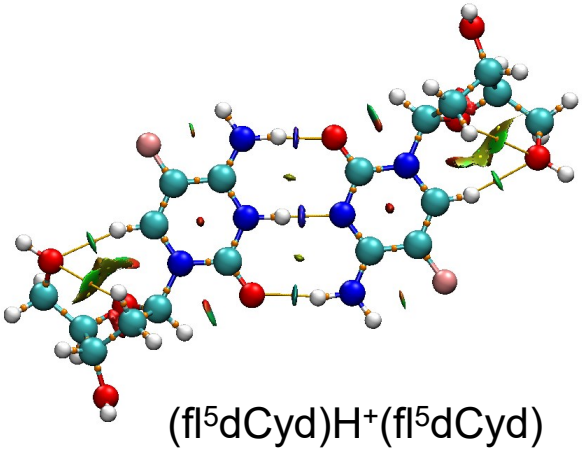

5′-Hydroxy Orientation
0
120
240
gauc

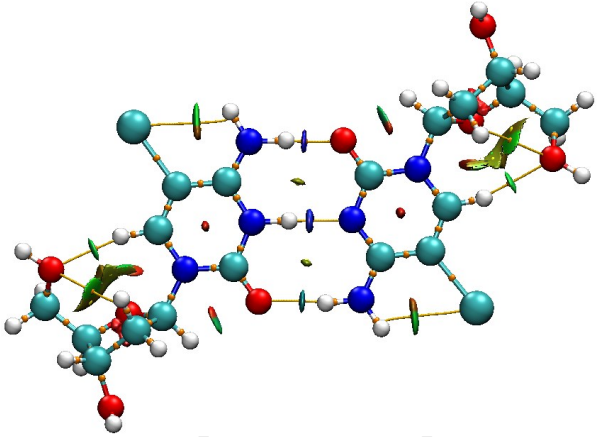
Figure 5.



(dCyd)H$^+$(dCyd)
***anti*, gauche$^+$**·*anti*, gauche$^+$
C3′-endo·C3′-endo (**^{3}T$_2$**· ^{3}T$_2$)



(fl^{5}dCyd)H$^+$(fl^{5}dCyd)
***anti*,gauche$^+$**·*anti*,gauche$^+$
C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)



(io^{5}dCyd)H$^+$(io^{5}dCyd)
***anti*,gauche$^+$**·*anti*,gauche$^+$
C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)

Figure 6.

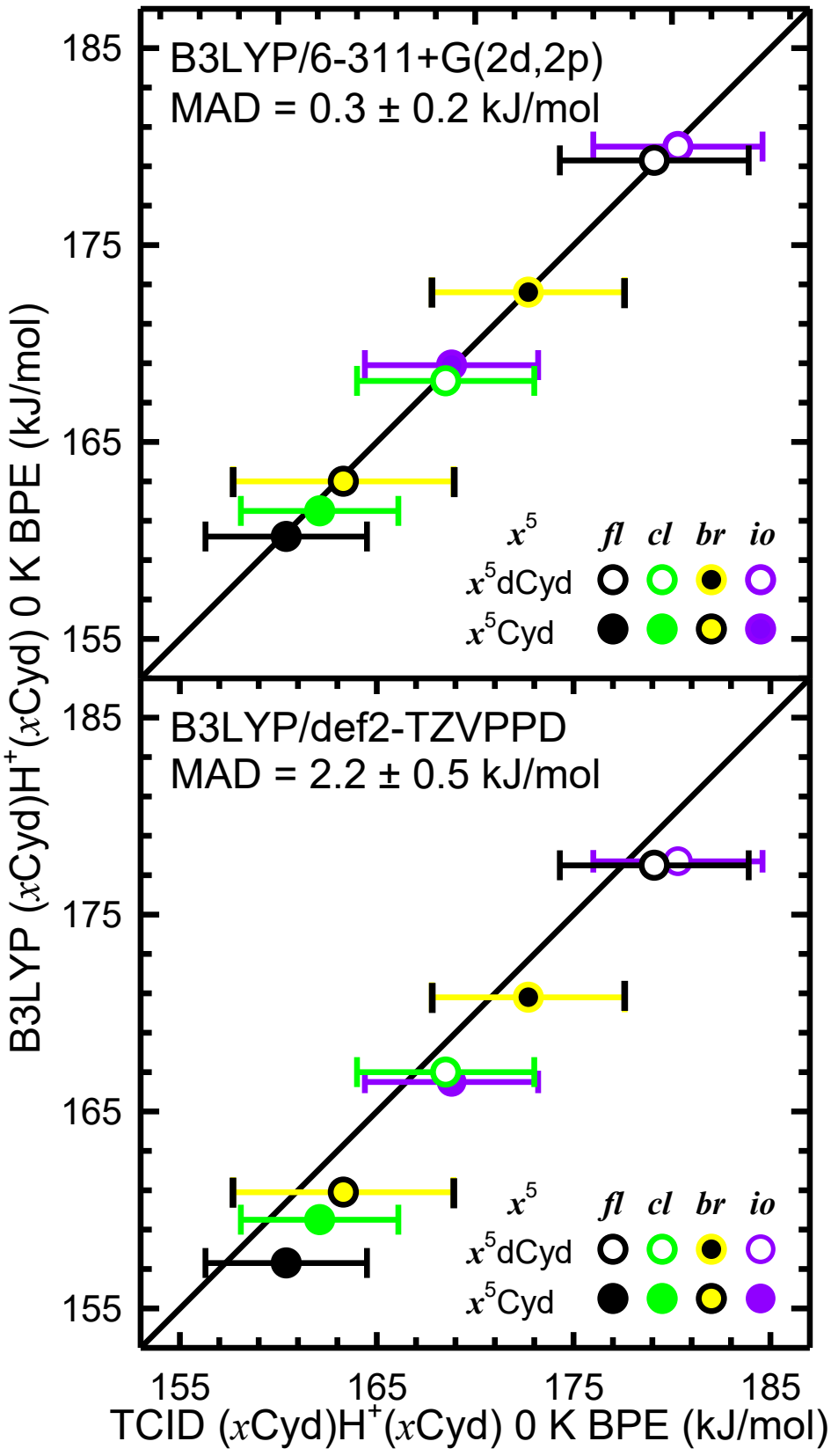

B3LYP/6-311+G(2d,2p)
MAD = 0.3 ± 0.2 kJ/mol
B3LYP/def2-TZVPPD
MAD = 2.2 ± 0.5 kJ/mol
x5 fl cl br io
x5dCyd
x5Cyd
185
175
165
155
B3LYP (xCyd)H+(xCyd) 0 K BPE (kJ/mol)
TCID (xCyd)H+(xCyd) 0 K BPE (kJ/mol)

Figure 7.

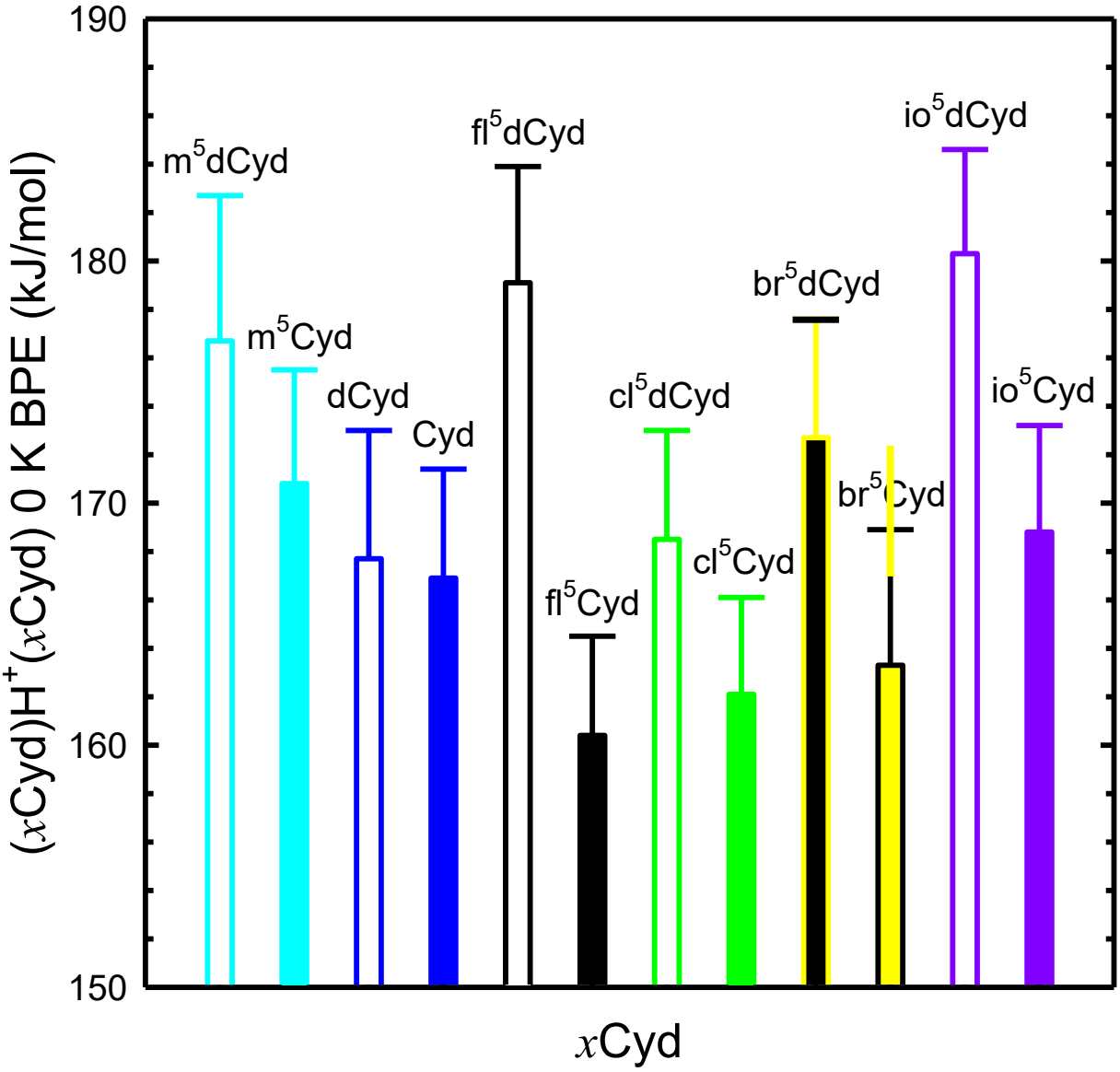

190
180
170
160
150
(xCyd)H+(xCyd) 0 K BPE (kJ/mol)
m5dCyd
m5Cyd
dCyd
Cyd
fl5dCyd
fl5Cyd
cl5dCyd
cl5Cyd
br5dCyd
br5Cyd
io5dCyd
io5Cyd
xCyd

Figure 8.

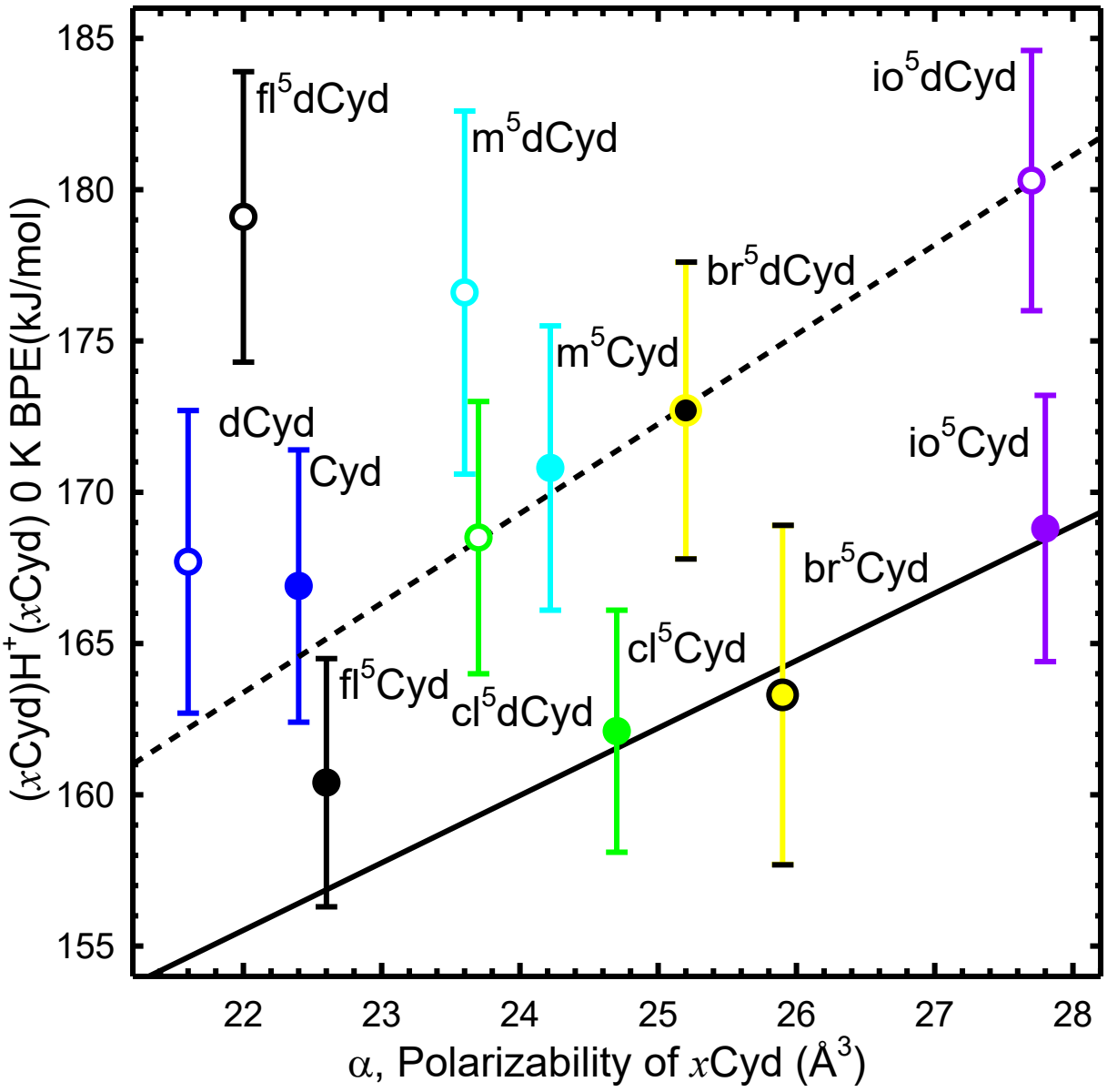
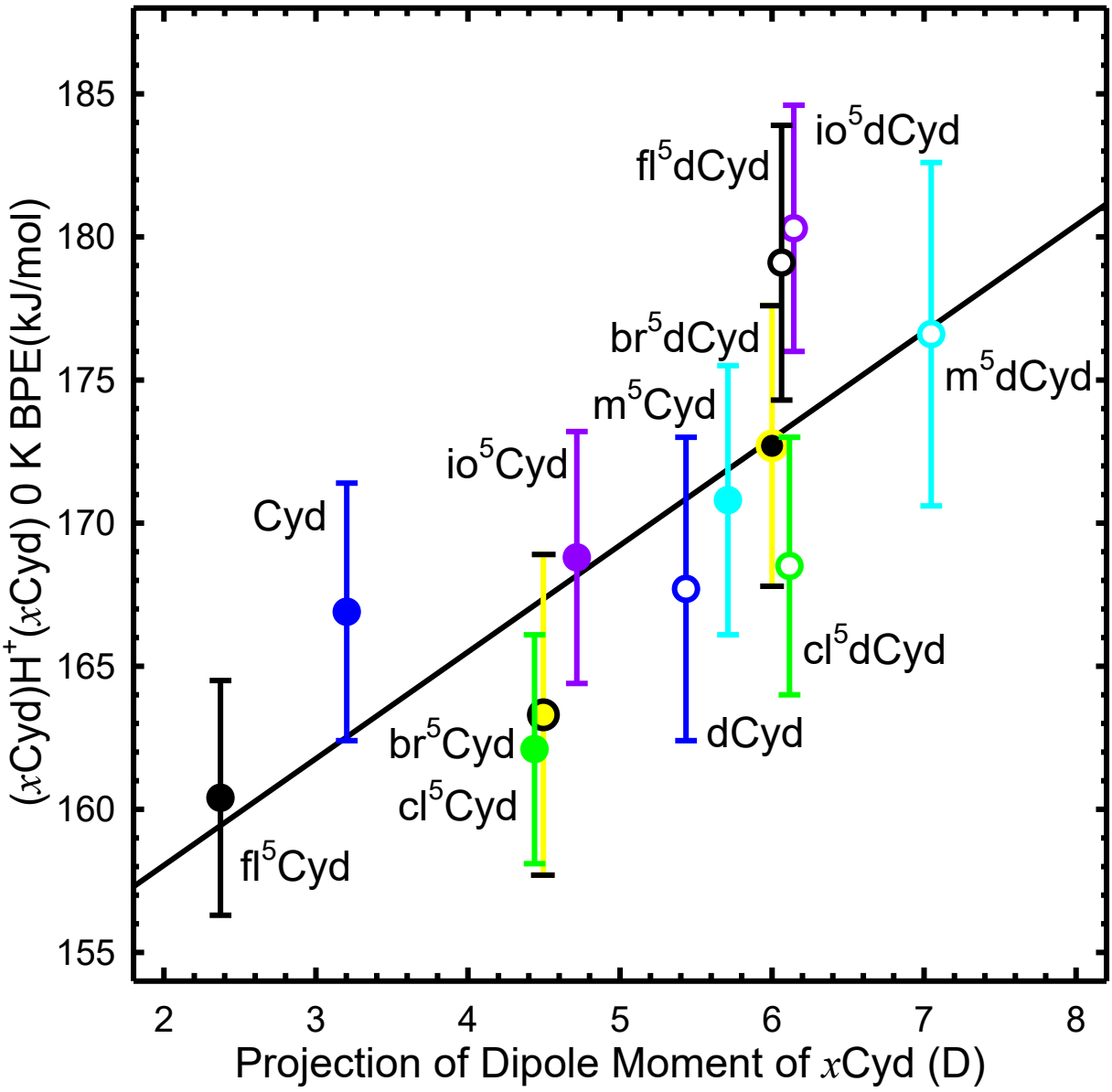
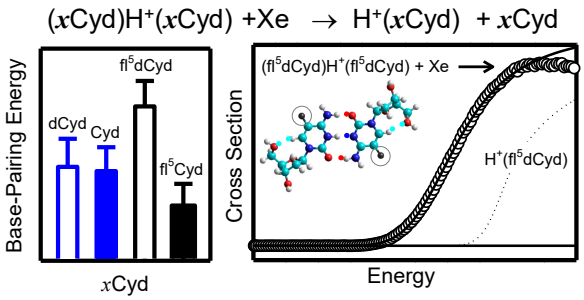

185
180
175
170
165
160
155
22
23
24
25
26
27
28
(xCyd)H+(xCyd) 0 K BPE(kJ/mol)
α, Polarizability of xCyd (Å3)
fl5dCyd
m5dCyd
io5dCyd
br5dCyd
m5Cyd
dCyd
Cyd
io5Cyd
br5Cyd
cl5Cyd
fl5Cyd
cl5dCyd

Figure 9.


185
180
175
170
165
160
155
2
3
4
5
6
7
8
(xCyd)H+(xCyd) 0 K BPE(kJ/mol)
Projection of Dipole Moment of xCyd (D)
Cyd
fl5Cyd
io5Cyd
br5Cyd
cl5Cyd
m5Cyd
dCyd
fl5dCyd
io5dCyd
br5dCyd
cl5dCyd
m5dCyd

For Table of Contents Use Only



Influence of 5-Halogenation on the Base Pairing Energies of Protonated Cytidine Nucleoside Analogue Base Pairs: Implications for the Stabilities of Synthetic *i*-Motif Structures for DNA Nanotechnology Applications

M. T. Rodgers*, Yakubu S. Seidu, and E. Israel

5-Halogenation of the cytosine nucleobases of the canonical DNA and RNA protonated cytidine nucleoside base pairs have been examined using complementary threshold collision-induced dissocation and computational methods. 5-Halogenation is found to enhance the base pairing of the DNA base pairs and weaken the base pairing for the RNA base pairs.

Figure 1.

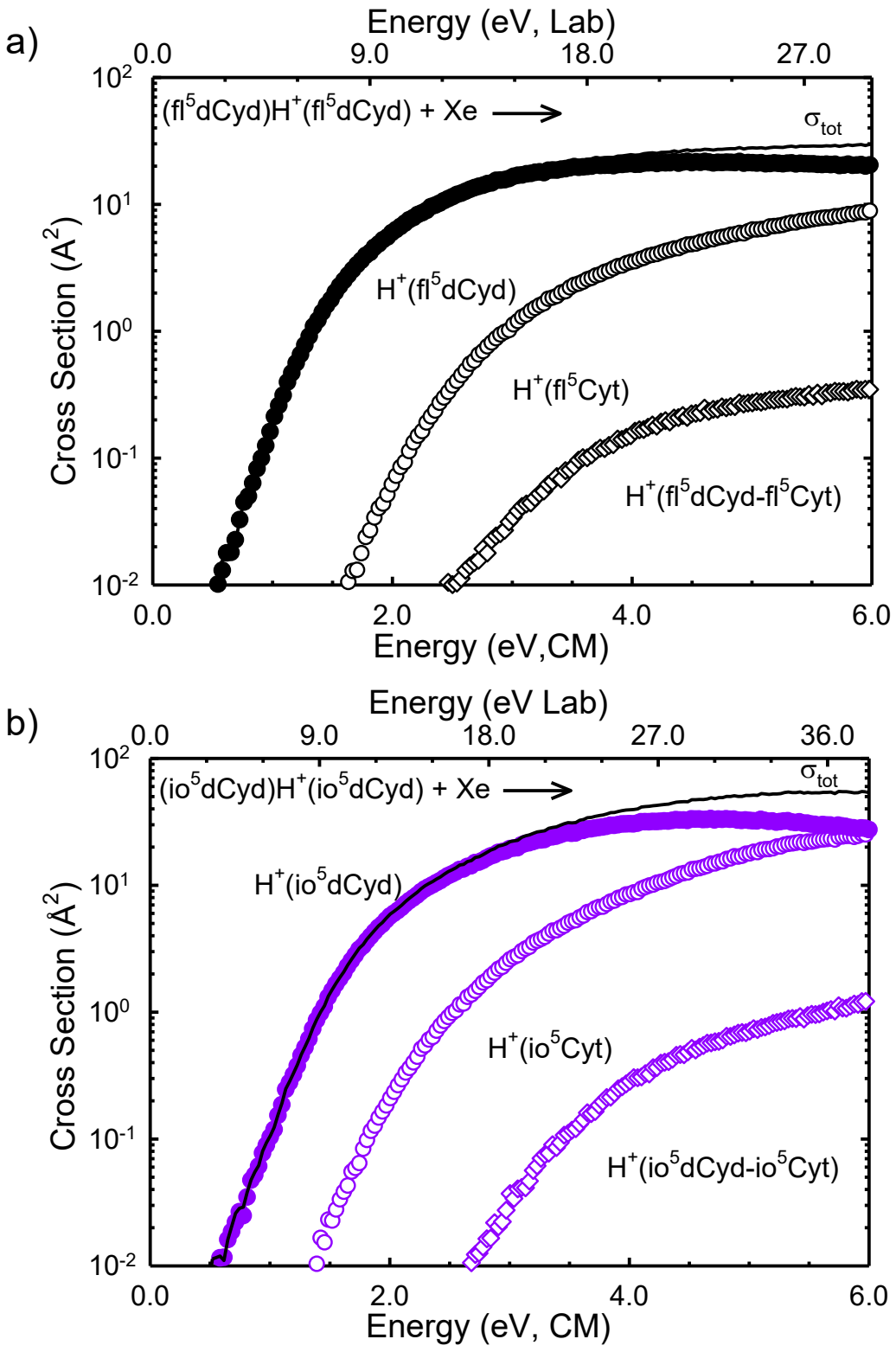

a)
Energy (eV, Lab)
0.0
9.0
18.0
27.0
(fl5dCyd)H+(fl5dCyd) + Xe
σtot
H+(fl5dCyd)
H+(fl5Cyt)
H+(fl5dCyd-fl5Cyt)
Cross Section (Å2)
10-2
10-1
100
101
102
0.0
2.0
4.0
6.0
Energy (eV,CM)
b)
Energy (eV Lab)
0.0
9.0
18.0
27.0
36.0
(io5dCyd)H+(io5dCyd) + Xe
σtot
H+(io5dCyd)
H+(io5Cyt)
H+(io5dCyd-io5Cyt)
Cross Section (Å2)
10-2
10-1
100
101
102
0.0
2.0
4.0
6.0
Energy (eV, CM)

Figure 2.

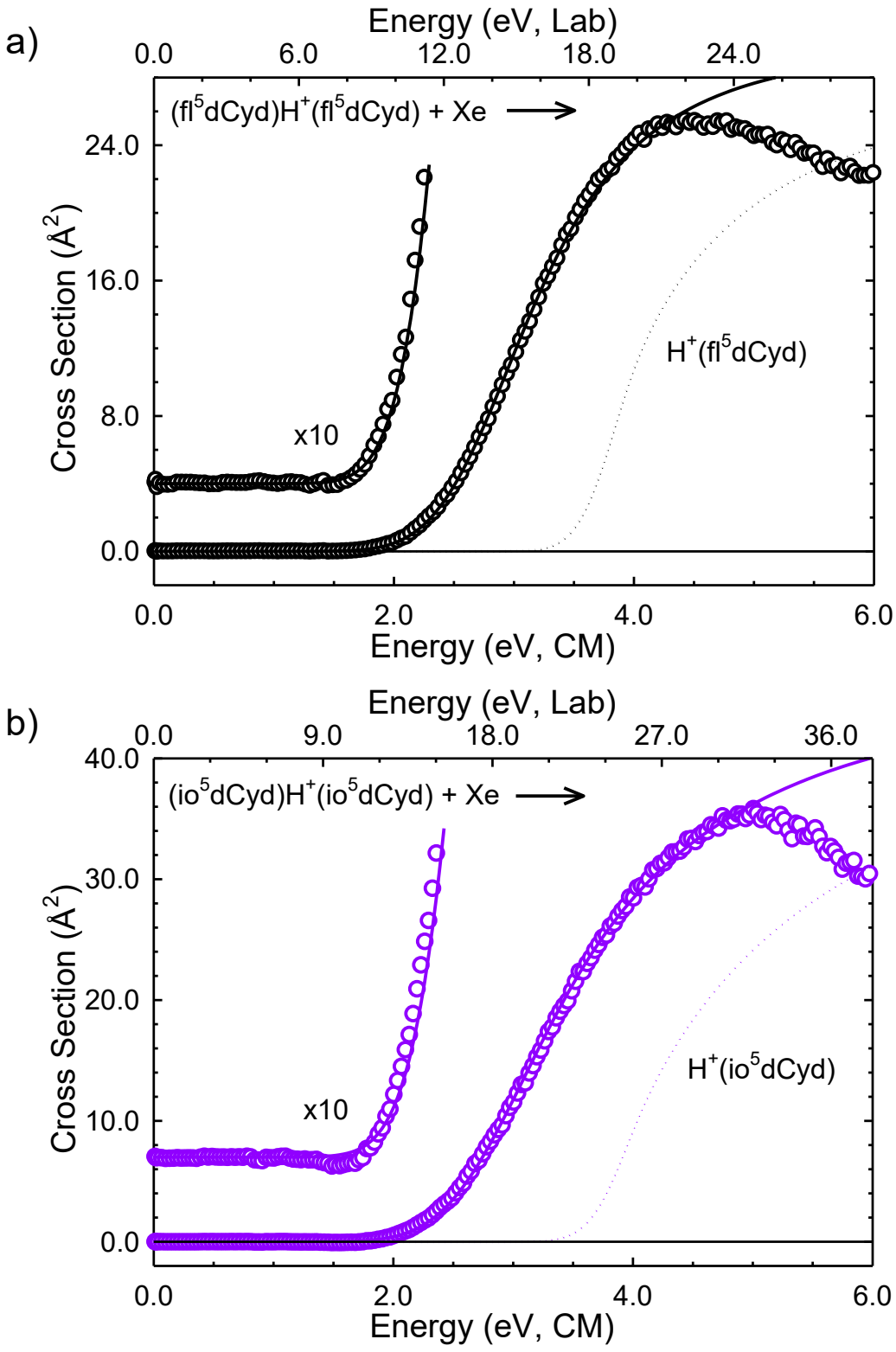
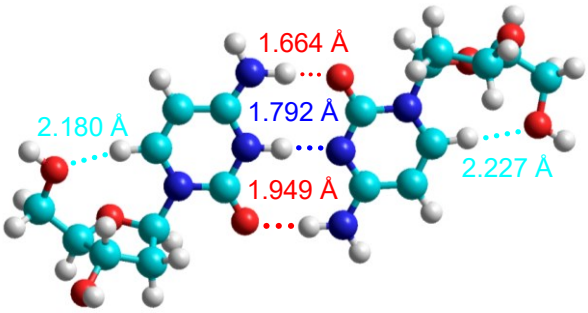

a)
Energy (eV, Lab)
0.0
6.0
12.0
18.0
24.0
(fl5dCyd)H+(fl5dCyd) + Xe
Cross Section (Å2)
24.0
16.0
8.0
0.0
x10
H+(fl5dCyd)
0.0
2.0
4.0
6.0
Energy (eV, CM)
b)
Energy (eV, Lab)
0.0
9.0
18.0
27.0
36.0
(io5dCyd)H+(io5dCyd) + Xe
Cross Section (Å2)
40.0
30.0
20.0
10.0
0.0
x10
H+(io5dCyd)
0.0
2.0
4.0
6.0
Energy (eV, CM)

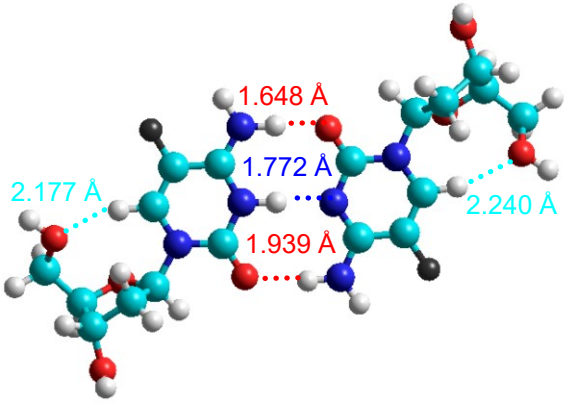
Figure 3.



(dCyd)H$^+$(dCyd)

***anti*, gauche$^+$**·*anti*, gauche$^+$

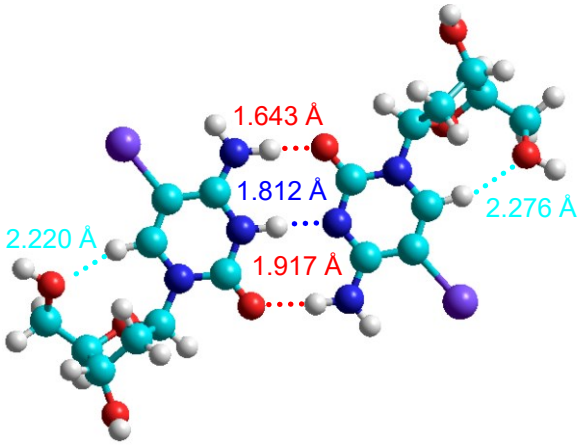
C3′-endo·C3′-endo (**^{3}T$_2$**· ^{3}T$_2$)



(fl^{5}dCyd)H$^+$(fl^{5}dCyd)

***anti*,gauche$^+$**·*anti*,gauche$^+$

C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)



(io^{5}dCyd)H$^+$(io^{5}dCyd)

***anti*,gauche$^+$**·*anti*,gauche$^+$

C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)

Figure 4.

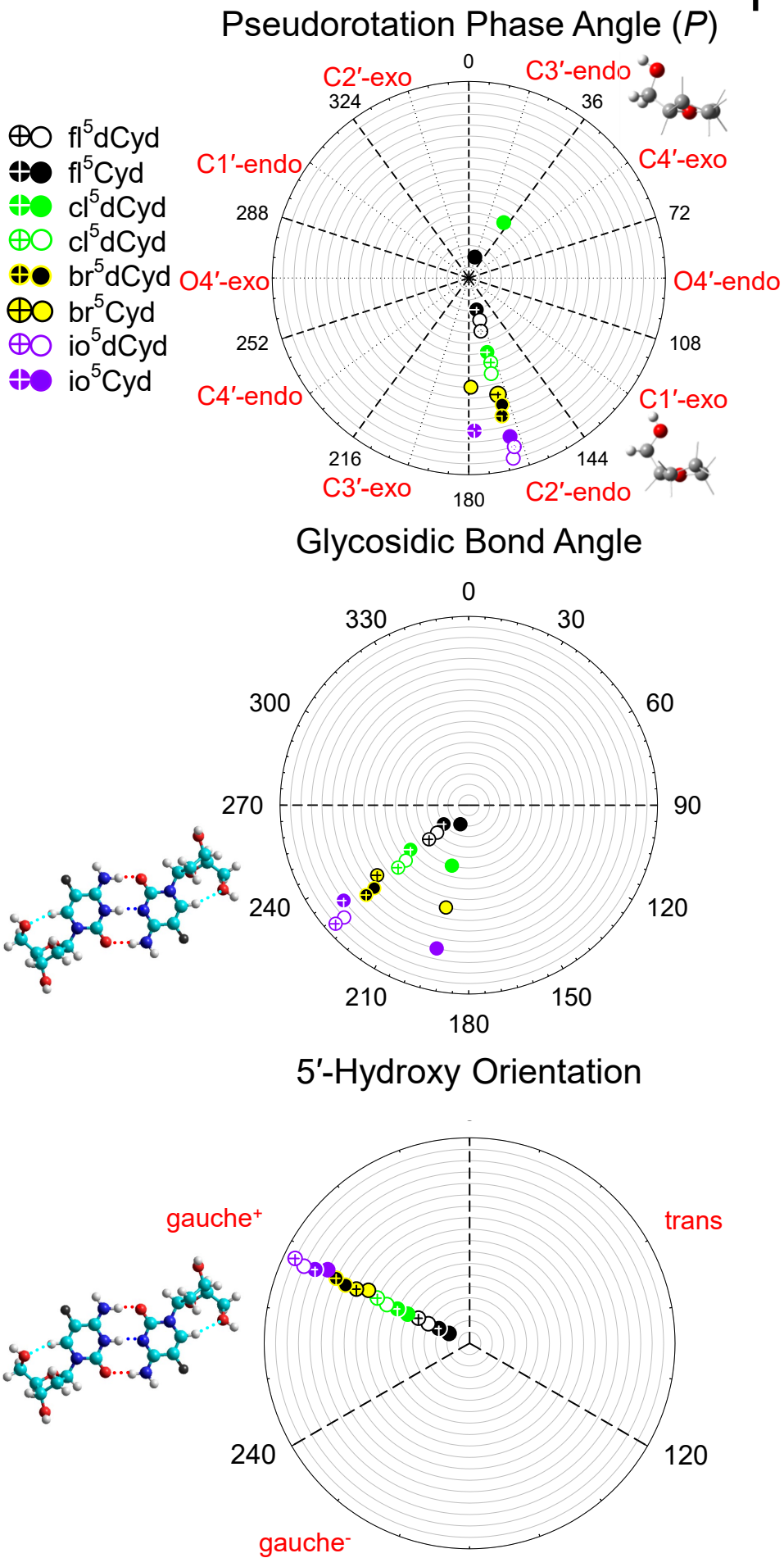
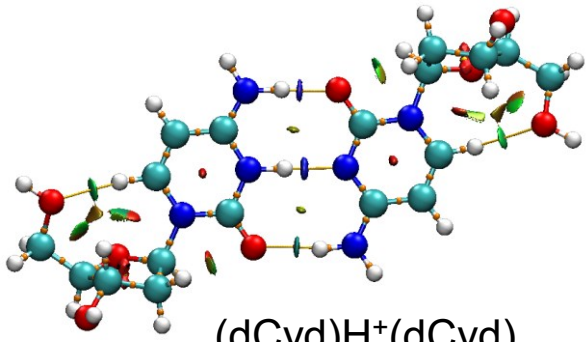
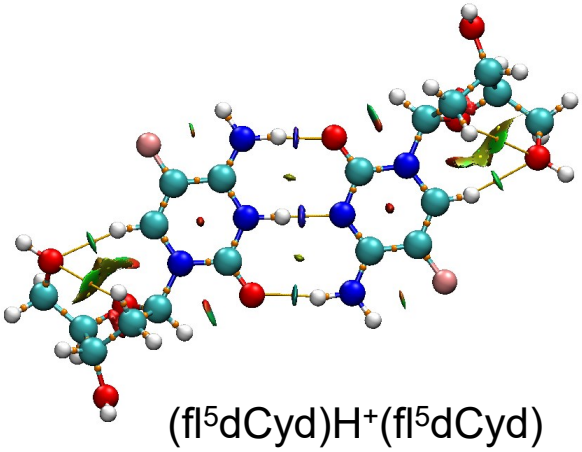

Pseudorotation Phase Angle (P)
fl5dCyd
fl5Cyd
cl5dCyd
cl5dCyd
br5dCyd
br5Cyd
io5dCyd
io5Cyd
0
36
72
108
144
180
216
252
288
324
C2′-exo
C3′-endo
C4′-exo
O4′-endo
C1′-exo
C2′-endo
C3′-exo
C4′-endo
O4′-exo
C1′-endo
Glycosidic Bond Angle
0
30
60
90
120
150
180
210
240
270
300
330
5′-Hydroxy Orientation
gauche+
trans
gauche-
120
240

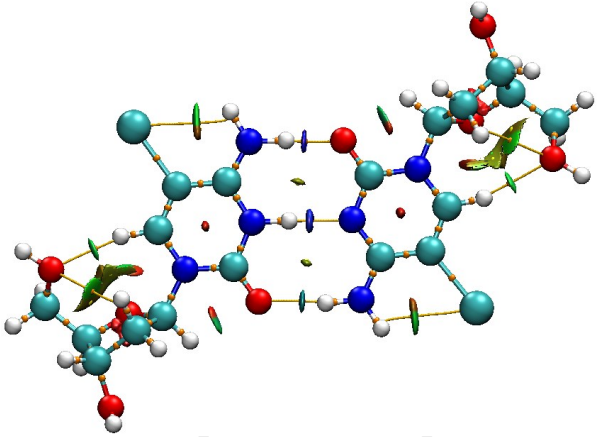
Figure 5.



(dCyd)H$^+$(dCyd)
***anti*, gauche$^+$**·*anti*, gauche$^+$
C3′-endo·C3′-endo (**^{3}T$_2$**· ^{3}T$_2$)



(fl^{5}dCyd)H$^+$(fl^{5}dCyd)
***anti*,gauche$^+$**·*anti*,gauche$^+$
C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)



(io^{5}dCyd)H$^+$(io^{5}dCyd)
***anti*,gauche$^+$**·*anti*,gauche$^+$
C2′-endo·C2′-endo (**^{2}T$_3$**· ^{2}T$_3$)

Figure 6.

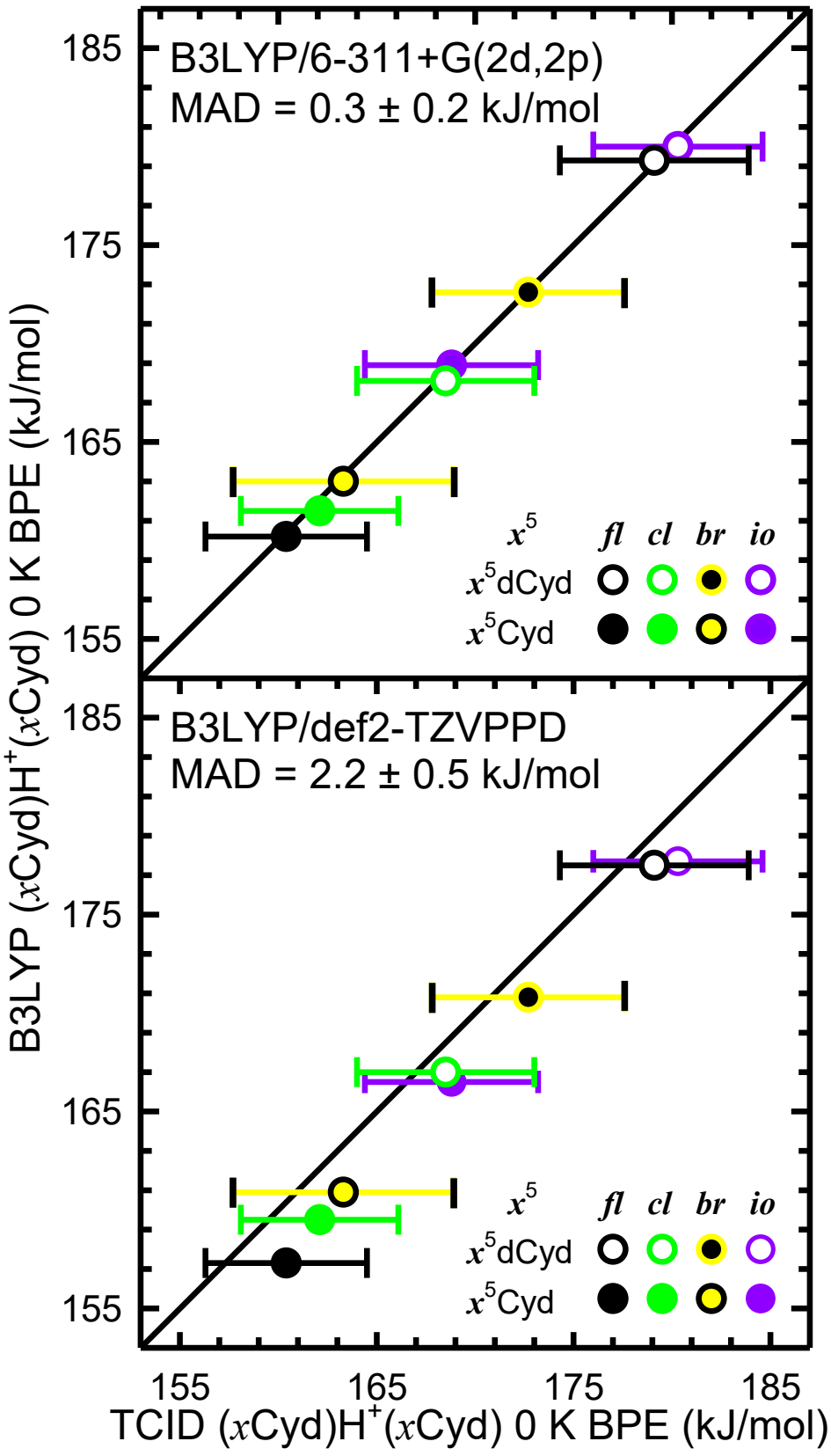

B3LYP/6-311+G(2d,2p)
MAD = 0.3 ± 0.2 kJ/mol
B3LYP/def2-TZVPPD
MAD = 2.2 ± 0.5 kJ/mol
x5 fl cl br io
x5dCyd
x5Cyd
185
175
165
155
B3LYP (xCyd)H+(xCyd) 0 K BPE (kJ/mol)
TCID (xCyd)H+(xCyd) 0 K BPE (kJ/mol)

Figure 7.

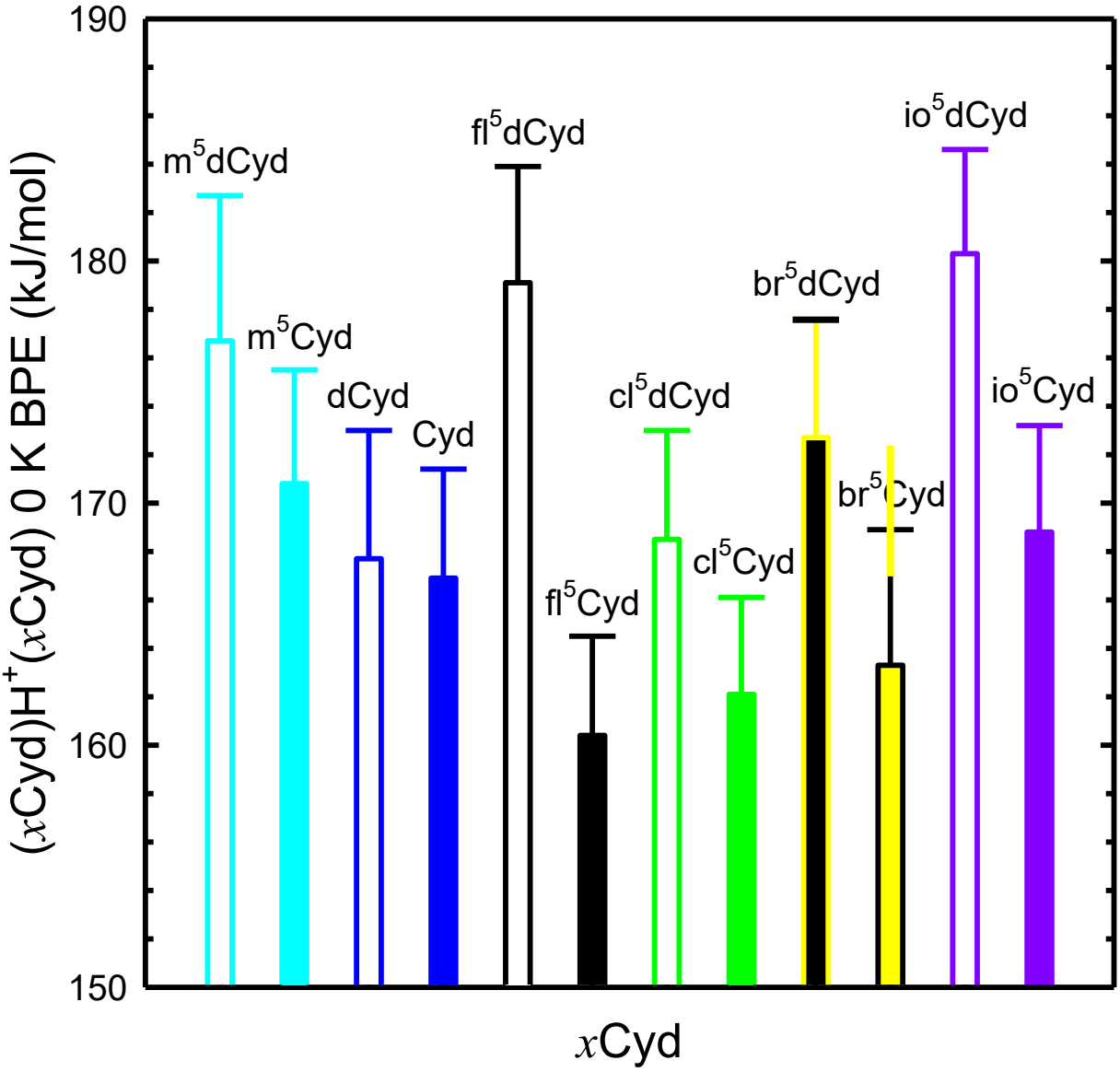

190
180
170
160
150
(xCyd)H+(xCyd) 0 K BPE (kJ/mol)
m5dCyd
m5Cyd
dCyd
Cyd
fl5dCyd
fl5Cyd
cl5dCyd
cl5Cyd
br5dCyd
br5Cyd
io5dCyd
io5Cyd
xCyd

Figure 8.

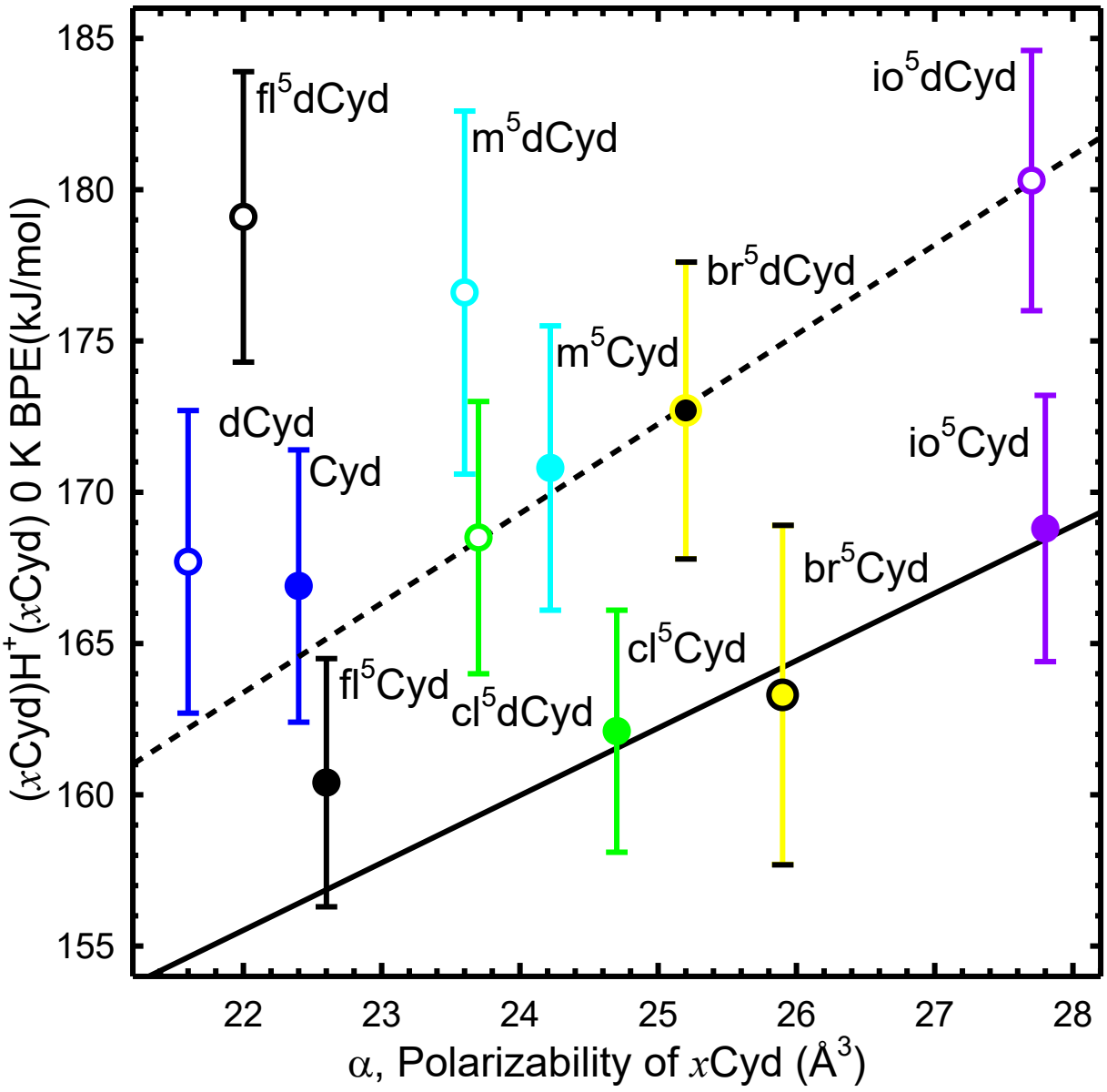
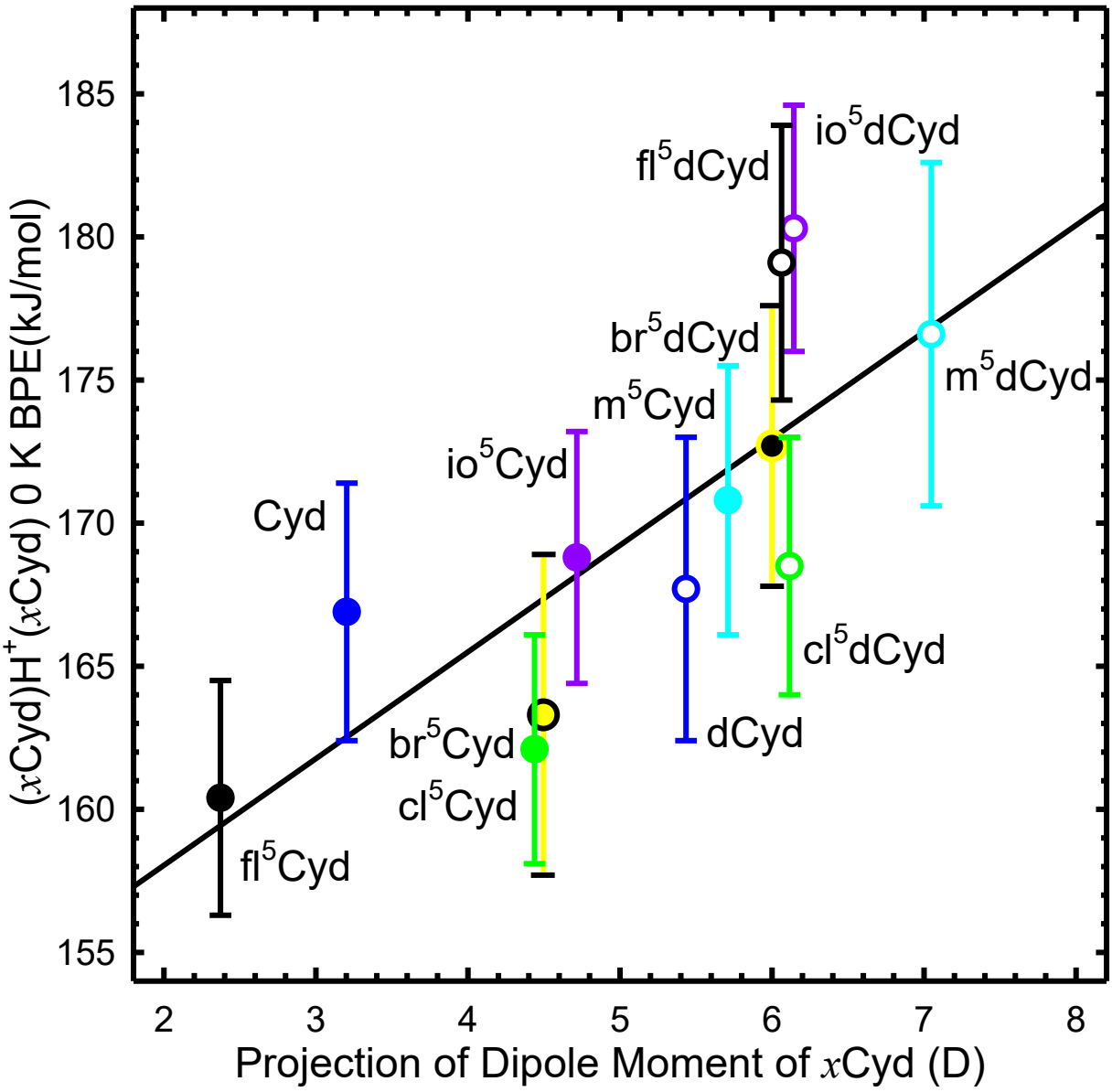
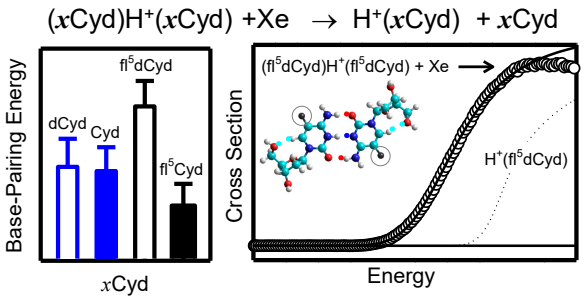

185
180
175
170
165
160
155
22
23
24
25
26
27
28
(xCyd)H+(xCyd) 0 K BPE(kJ/mol)
α, Polarizability of xCyd (Å3)
fl5dCyd
m5dCyd
io5dCyd
br5dCyd
m5Cyd
dCyd
Cyd
io5Cyd
br5Cyd
cl5Cyd
fl5Cyd
cl5dCyd

Figure 9.


185
180
175
170
165
160
155
2
3
4
5
6
7
8
(xCyd)H+(xCyd) 0 K BPE(kJ/mol)
Projection of Dipole Moment of xCyd (D)
Cyd
fl5Cyd
io5Cyd
br5Cyd
cl5Cyd
m5Cyd
dCyd
fl5dCyd
io5dCyd
br5dCyd
cl5dCyd
m5dCyd

For Table of Contents Use Only



Influence of 5-Halogenation on the Base Pairing Energies of Protonated Cytidine Nucleoside Analogue Base Pairs: Implications for the Stabilities of Synthetic *i*-Motif Structures for DNA Nanotechnology Applications

M. T. Rodgers*, Yakubu S. Seidu, and E. Israel

5-Halogenation of the cytosine nucleobases of the canonical DNA and RNA protonated cytidine nucleoside base pairs have been examined using complementary threshold collision-induced dissocation and computational methods. 5-Halogenation is found to enhance the base pairing of the DNA base pairs and weaken the base pairing for the RNA base pairs.