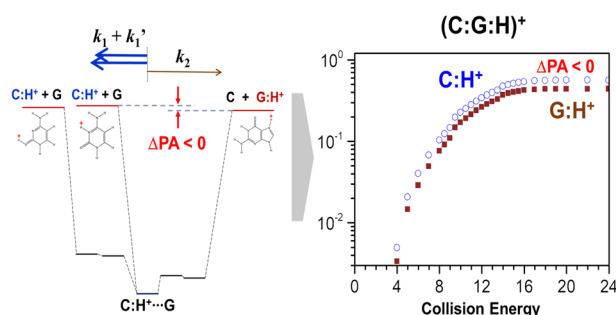


RESEARCH ARTICLE

Proton Transfer Accounting for Anomalous Collision-Induced Dissociation of Proton-Bound Hoogsteen Base Pair of Cytosine and Guanine

Jeong Ju Park,¹ Choong Sik Lee,² Sang Yun Han¹ ¹Department of Nanochemistry, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do 13120, Republic of Korea²Scientific Investigation Laboratory, Ministry of National Defense, 22 Itaewon-ro, Yongsan-gu, Seoul, 04383, Republic of Korea

Abstract. To understand the anomalous collision-induced dissociation (CID) behavior of the proton-bound Hoogsteen base pair of cytosine (C) and guanine (G), C:H⁺...G, we investigated CID of a homologue series of proton-bound heterodimers of C, 1-methylcytosine, and 5-methylcytosine with G as a common base partner. The CID experiments were performed in an energy-resolved way (ER-CID) under both multiple and near-single collision conditions. The relative

stabilities of the protonated complexes examined by ER-CID suggested that the proton-bound complexes produced by electrospray ionization in this study are proton-bound Hoogsteen base pairs. On the other hand, in contrast to the other base pairs, CID of C:H⁺...G exhibited more abundant productions of C:H⁺, the fragment protonated on the moiety with a smaller proton affinity, than that of G:H⁺. This appeared to contradict general prediction based on the kinetic method. However, further theoretical exploration of potential energy surfaces found that there can be facile proton transfers in the proton-bound Hoogsteen base pairs during the CID process, which makes the process accessible to an additional product state of O-protonated C for C:H⁺ fragments. The presence of an additional dissociation channel, which in other words corresponds to twofold degeneracy in the transition state leading to C:H⁺ fragments, effectively doubles the apparent reaction rate for production of C:H⁺. In this way, the process gives rise to the anomaly, the observed pronounced formation of C:H⁺ in the CID of the proton-bound Hoogsteen base pair, C:H⁺...G.

Keywords: Proton-bound Hoogsteen base pairs, Energy-resolved collision-induced dissociation (ER-CID), Guanine, Cytosine, Proton transfer

Received: 19 June 2018/Revised: 8 August 2018/Accepted: 8 August 2018/Published Online: 13 September 2018

Introduction

Protonation of nucleic acid bases has been known to cause certain alternations in DNA structures, which may affect related biological processes. A well-known example is the

Jeong Ju Park and Choong Sik Lee contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13361-018-2060-5>) contains supplementary material, which is available to authorized users.

Correspondence to: Sang Han; e-mail: sanghan@gachon.ac.kr

formation of DNA i-motif [1–4]. Base pairing involved in proton-bound cytosine (C) base pairs, C:H⁺...C, which occurs in the tetrameric structures in DNA i-motif, is known to include three hydrogen bonds, one of which is an ionic hydrogen bond [5–7]. The triply-bound structure of C:H⁺...C was clearly characterized by infrared multiple photon dissociation (IRMPD) spectroscopy of the proton-bound dimer of 1-methylcytosine (1-MeC) in the gas phase [8]. IRMPD spectroscopy was further used to interrogate the triply-bound base pairing involved in the proton-bound dimers of 1-MeC and its derivatives [9] and those of C and modified Cs [10, 11]. IRMPD spectroscopy combined with ion trap mass spectrometry has offered a

powerful means to probe the structures of gaseous proton- or metal ion-bound biomolecules and their complexes [12–15]; it has also successfully elucidated the structures and energetics of proton-bound base pairs of adenine (A), (A:A:H)⁺ [16], C and guanine (G), (C:G:H)⁺ [17], and G tetrads [18] in a combined approach with quantum chemical calculations.

As for the ionic hydrogen bond involved in the triply-bound C:H⁺⋯C, it was theoretically predicted that the proton transfer in the ionic hydrogen bond can occur at a high rate constant of $6 \times 10^{11} \text{ s}^{-1}$ due to the low energy barrier of 1.6 kcal/mol and quantum mechanical tunneling [7], which was later observed by an NMR study ($> 10^{-8} \text{ s}^{-1}$) [19]. While protonation of the N3-position of C has been found to participate in many base-pairing cases of protonated C [8–11, 19] including in DNA i-motif structures [1], proton-bound Hoogsteen base pairing of (C:G:H)⁺ [17, 20], (C:G:C:H)⁺ triplet [20] and triple helices [21], protonation of the O2-position of C is also known to occur and cause a mismatched base pairing of A-C in DNA [22, 23].

In a complementary relation to IRMPD spectroscopy, tandem mass spectrometry utilizing collision-induced dissociation (CID) has been an important tool in elucidating binding energetics and related thermochemistry. For example, Cooks' kinetic method has been useful in determining the thermochemical properties of molecules (B₁, B₂), when they are involved in a loosely bound ion pair by an ionic hydrogen bond, B₁⋯H⁺⋯B₂. The kinetic method suggests that the fragment abundances of [B₁:H⁺] and [B₂:H⁺] can be directly related to the enthalpy of protonation (ΔH_p) or the proton affinity (PA) of the moieties in a limiting case in which the entropy effect is essentially negligible [24–26].

In an approach known as threshold collision-induced dissociation (TCID) [27], the abundances of fragments, which approximate the reaction rate constants (*k*), were measured in near-single collision conditions using well-characterized guided-ion beam tandem mass spectrometers [27–32]. TCID experiments were applied to investigate various base pairs of protonated Cs, such as methylated and halogenated Cs and deoxycytidines [33–37]; these experiments yielded insight into base-pairing interaction by making accurate determination of binding energetics and PAs. In addition, CID experiments to measure depletion of parent ions in an energy-resolved way (ER-CID), in which parent ions experience multiple collisions in an ion trap, also provided a reliable measure of relative stability of parent ions. ER-CID experiments, often combined with IRMPD spectroscopy, have offered important information including relative stabilities and conformations of biomolecules and cluster ions such as nucleosides and G-tetrads in the gas phase [18, 38, 39].

On the other hand, an important example of DNA structural variations by protonation is proton-bound Hoogsteen base pairing [20, 40], in which protonated C binds to a Watson-Crick (WC) base pair of G-C, giving rise to the formation of a C:H⁺⋯G-C triplet and triple strands of DNA nucleotide oligomers [21, 41]. While WC base pairing is more stable at physiological conditions, protonated Hoogsteen base pairs can be stabilized and occur under certain circumstances [42–45]. This

formation of triplexes is related to human diseases [46] such as Friedreich's ataxia [47], and it also received research attention for its potential applications in therapeutics [48]. As for the motif of proton-bound Hoogsteen base pairing, the binding interaction of CH⁺⋯G dimer and CH⁺⋯G-C triplet was predicted to possess a strong interaction energy of 38.2 kcal/mol for the proton-bound Hoogsteen base pair. The energy barrier of proton transfer in the ionic hydrogen bond was also low at about 4 kcal/mol [20].

In contrast to the triply-bound protonated dimers of C, the experimental efforts to understand the base pairing involved in Hoogsteen base pairs at the molecular level have been rather limited to date. While the isolated protonated complex of C and G, (C:G:H)⁺, was previously produced by electrospray ionization (ESI) [49], its structures have not been identified until a recent IRMPD study. Very recently, an IRMPD study performed in the spectral range of 900–1900 cm⁻¹ along with differential ion mobility spectroscopy revealed that the isolated (C:G:H)⁺ complex produced by ESI under acidic conditions (pH = 3.2) dominantly possesses the proton-bound Hoogsteen conformation of C:H⁺⋯G with a population higher than 66% [17]. The IRMPD study also found that the preferred conformations of (C:G:H)⁺ varied depending on the pH of the ESI solutions. It was found that ESI preferentially produced the proton-bound WC base pair (91%) at pH = 5.8. The identification using IRMPD at the acidic conditions (pH = 3.2) is in a good accordance with a recent quantum chemical study for various conformations of (C:G:H)⁺ [50]. That theoretical study explored a variety of conformations of proton-bound dimeric structures complexed by multiple hydrogen bonds. As a result, the proton-bound Hoogsteen base pair was proposed as the most stable conformation in the gas phase and also in the aqueous environment; its calculated thermal free energy suggested that it would be predominant in the gas phase (> 90%).

In fact, the recent IRMPD spectroscopy [17] and early CID experiments [49], which have in common the excitation mechanism of slow heating, noted an intriguing feature in CID of a proton-bound Hoogsteen base pair, C:H⁺⋯G. In view of the kinetic method [24], CID is expected to occur with preferential formation of G:H⁺ rather than C:H⁺, as the PA of G is larger than that of C. However, both CID and IRMPD experiments found an anomaly in dissociation, in which the formation of C:H⁺ product was more pronounced than that of G:H⁺. Accordingly, while the recent IRMPD study has begun to shed light on the base pairing in isolated proton-bound Hoogsteen base pairs, the CID behavior still leaves an interesting question to answer. In this regard, in an effort to understand the intriguing behavior observed in the CID of (C:G:H)⁺, we performed a combined approach of ER-CID experiments in multiple and near-single collision conditions with theoretical calculations for a homologue of proton-bound base pairs of C, 1-MeC, and 5-methylcytosine (5-MeC) with G as a common pairing moiety.

Experimental and Theoretical Methods

Materials

Chemicals including cytosine (C), guanine (G), and 5-methylcytosine (5-MeC) were commercially obtained from Sigma-Aldrich Korea (Suwon, Korea). 1-Methylcytosine (1-MeC) was purchased from Synchem UG & Co. KG (Felsberg, Germany). Other chemicals including HPLC-grade acetonitrile and acetic acid were also obtained from Sigma-Aldrich Korea.

ER-CID Experiments in Multiple Collision Conditions

For ER-CID experiments in multiple collision conditions, an ion trap mass spectrometer (LTQ XL, Thermo Scientific, MA, USA) was utilized. The proton-bound heterodimers were produced by electrospray ionization of binary sample mixtures (200 μM each) dissolved in a solution of water/acetonitrile/acetic acid (49:49:2, v/v/v). The solution was directly infused into the mass spectrometer at a flow rate of 5 $\mu\text{l}/\text{min}$, and the temperature of the mass spectrometer inlet capillary was kept at 150 $^{\circ}\text{C}$. For the CID experiments, the excitation energy was set in terms of the normalized collision energy (N.C.E.), which is the instrument's own parameter in arbitrary unit (a.u.). The scan step was 0.1 or 0.05 depending on the required resolution. The collision gas was He (~ 1 mTorr) and the ions were excited for 30 ms, during which ions experienced multiple collisions with the collision gas in the linear ion trap. The resultant survival yield (S.Y.) for parent dimer ions as a function of N.C.E. was monitored as the ratio of the ion intensity of surviving parent ion, $I(p)$, to the sum of ion intensities of all fragments, $I(f)$, and $I(p)$ measured in the CID spectra; $\text{S.Y.} = I(p) / I_{\text{total}} = I(p) / (\sum I_i(f) + I(p))$. Similarly, the formation yield (F.Y.) of a fragment, f , was defined as $\text{F.Y.}(f) = I(f) / I_{\text{total}} = I(f) / (\sum I_i(f) + I(p))$. Using the least square method, the obtained S.Y. curves were fitted to the curve in this study, $\text{S.Y.} = [1 + (\text{N.C.E.} / \text{CID}_{50\%})^n]^{-1}$, where the proton-bound base pairs effectively underwent complete dissociation at high collision energies. The fitted value of $\text{CID}_{50\%}$ corresponds to the collision energy at which 50% depletion of parent ion occurs, which represents the relative stability of the proton-bound dimers. The value n represents the steepness of the declining region of the fitted curve [51].

ER-CID Experiments in Near-Single Collision Conditions

For near-single collision conditions, the ER-CID experiments were performed on a quadrupole tandem mass spectrometer (Xevo TQ, Waters Corporation, Manchester, UK), which is equipped with a stacked-ring ion-guide collision cell in place of a conventional multipole cell. The stacked-ring ion guide in the cell produces an effectively static pseudopotential that is much steeper at the ring edge when compared to that of common multipole ion guides. It is effectively immune to the loss of parent and fragment ions flying through the collision cell. The

advantages of this collision cell coming from static rather than radio-frequency potentials include low electrical noises as well as a large field-free region near the central axis of the guide. In this study, we turned off the DC voltages for the RF and DC barriers, and the traveling wave was only applied to the collision cell [52, 53]. Thus, the scanwave cell served just as a ring-electrode collision cell, i.e., a fragmentation and ion-guide device. For CID, Ar was used as the collision gas; this was the default setting for this instrument. The flow rate of collision gas into the collision cell was carefully controlled by adding an extra high-precision metering valve (Metering HR series, Parker Hannifin, Corp.) so that the pressure inside collision cell could be adjusted to as low as 0.06, 0.10, and 0.15 mTorr, which values correspond to near-single collision conditions in this instrument. The effects of collision gas pressure were also examined; a small effect of gas pressure, which slightly increased the fragmentation yield as the pressure increased, could be observed. The CID behaviors for all the proton-bound base pairs were carefully examined at the three pressures, which ensured no significant alteration of the ER-CID data. In addition, fragmentation that occurs without introduction of collision gas was measured and subtracted from the raw CID data for background correction. However, such background fragmentation was negligible in this threshold energy region.

The collision energy in the laboratory frame (E_{lab}) was controlled at an energy step of 0.2 or 0.5 eV. The center-of-mass energy (E_{com}) was obtained using the relation $E_{\text{com}} = (E_{\text{lab}} + V_c) \times m_{\text{Ar}} / (M_{\text{ion}} + m_{\text{Ar}})$, where M_{ion} is the mass of base pair ions and m_{Ar} is that of Ar. However, the setting value of collision energy (E_{lab}) can be influenced by certain factors such as residual fields in the inlet and outlet parts of the collision cell, inaccuracy in applied voltages, contact potentials, and so on, which are difficult to assess when using a commercial instrument. To account for such effects on the collision energy in the laboratory frame, we introduced a single-parameter calibration by employing a potential correction (V_c) in the above relation; this parameter was estimated in comparison with the reported CID results for (1-MeC:C:H)⁺, (5-MeC:C:H)⁺ [35], cytidine:H⁺ [54], and guanosine:H⁺ [55], and which led to an estimate of 2.54 V for V_c for this study. In addition, unlike guided ion-beam instruments specially designed or modified for measurement of threshold CID cross sections [27–30], instrumental parameters such as the effective length of the collision cell are not known. For this reason, we refrained from assessing collision cross sections from the observed ER-CID data obtained in the near-single collision conditions.

Theoretical Methods

Theoretical calculations were performed using the Gaussian 09 program package [56]. Density functional theory using the B3LYP hybrid functional was used in this study, which has been reliably employed in many theoretical investigations of nucleic acid bases and base pairs [20, 34–37]. The minimum energy structures were optimized without structural constraints.

Vibration frequency calculations for the optimized structures were carried out to examine if the structures represented true minima on the potential energy surfaces. The calculations also yielded thermochemical quantities such as enthalpy (H) and Gibbs free energy (G) at 298.15 K. For the energetics, the electronic energy difference values with and without zero-point energy corrections, ΔE_0 and ΔE_c , respectively, were calculated. As for the interaction energy between the two moieties in a base pair, $A^+\cdots B$, the dissociation energy (D_0 or D_e) defined as $D = [E(A^+) + E(B)] - E(A^+\cdots B)$ was evaluated. The complexation energy corrected for basis set superposition errors (D_{BSSE}) using the relaxed counterpoise method was also calculated [57]. Polarizable continuum model (PCM) calculations were performed when the energetics in the aqueous environment was examined [58].

To locate TS, potential energy scan was performed for large areas on the potential energy surfaces. TS was then located using QST3 calculations [59], which were further examined by frequency calculations. To accurately measure energies for monomers including PA and ΔG , the MP2/aug-cc-pVDT level of theory was also applied; this theory is known to perform in good agreement with the CCSD(T) theory [60, 61]. In comparison with the values obtained by B3LYP density functional, calculations using the MP2/aug-cc-pVDT theory at B3LYP-optimized geometries were also carried out.

Results and Discussion

Anomalous Fragmentation Yields in CID of $(C:G:H)^+$

The kinetic method has offered a simple and accurate way of determining thermochemical quantities by monitoring competitive cleavages of a series of weakly bound complexes that include a common ion [24–26]. This method suggests that in a competitive dissociation, the formation rate constants (k_i), which are approximated by the respective fragment ion abundances, depend on the enthalpy difference between the two competitive product channels, if the entropic effects for the two fragmentation pathways cancel. This gives the following relation for the dissociation of a proton-bound complex of $A^+\cdots B$, leading to the formation of $A:H^+$ and $B:H^+$ [24]: $\ln(k_1/k_2) = \ln([A:H^+]/[B:H^+]) \approx \Delta(\Delta G)/RT_{\text{eff}} \approx \Delta(\Delta H)/RT_{\text{eff}}$, where k_i accounts for the rate constant for fragmentation channel i , $[A:H^+]$ for the abundance of $A:H^+$ fragment ion, $\Delta(\Delta G)$ and $\Delta(\Delta H)$ for the differences in Gibbs free energy and enthalpy between the two product states, respectively, and T_{eff} is the effective temperature.

The schematic potential energy diagram for the proton-bound Hoogsteen base pair of $C:H^+\cdots G$ is given in Figure 1. In the case of $C:H^+\cdots G$, the TS is in general considered to be loose and product-like, so the property of TS can be deduced from the corresponding product state. The literature values for the entropies of protonation for C and G are small and essentially the same [62]. Thus, it is expected that the branching ratio between the two channels producing $G:H^+$ and $C:H^+$ is largely

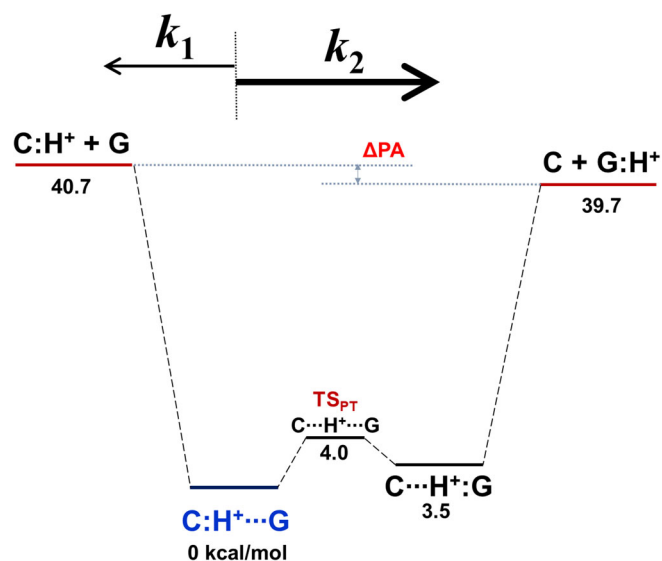


Figure 1. Schematic energy diagram for CID of the proton-bound base pair of C and G, $C:H^+\cdots G$. The values are ΔE_0 (kcal/mol) from the most stable, Hoogsteen structure (Table 2)

determined by the PAs of the two moieties; G and C. G is known to possess a higher PA (229.3 kcal/mol) than that of C (227.0 kcal/mol); these values were experimentally measured by the kinetic method using pyrrolidine [62, 63].

This study calculated the thermochemical quantities for protonation of G and C; a larger ΔH_p (298.15 K) value of 228.1 kcal/mol for G than that of 226.8 kcal/mol for C when C is protonated at the N3-position at the theory level of B3LYP/6-311+G(2d,p) (Table 1). The molecular structures of N3-protonated C, $C(N3):H^+$, O2-protonated C, $C(O2):H^+$, and N7-protonated G, $G(N7):H^+$, along with that of proton-bound Hoogsteen base pair, $C:H^+\cdots G$, are given in Figure 1S (Supporting Information). As for ΔG_p (298.15 K), values of 228.5 and 226.7 kcal/mol were predicted for G and C, respectively. In addition, the difference between $\Delta(\Delta G_p)$ and $\Delta(\Delta H_p)$, i.e., $T\Delta(\Delta S)$, is calculated and found to be small at only 0.5 kcal/mol. This suggests that $\Delta(\Delta H_p)$ largely accounts for $\Delta(\Delta G_p)$ suggesting that the entropic effect in product states and probably that in the product-like TSs are negligible. In fact, the kinetic method is known to be suitable for accurate determination of PAs of nucleic acid bases and nucleosides [64]. In this regard, the application of the kinetic method is reasonable for predicting the dissociation behavior of the $C:H^+\cdots G$ base pair.

Form the kinetic method, it can be expected that $G:H^+$ would be more abundant than $C:H^+$ in fragmentation of $C:H^+\cdots G$ near the dissociation thresholds due to the larger PA of G; this was, however, not the case according to the previous findings obtained by CID and IRMPD experiments [17, 49]. There may be certain dynamic reasons that cannot be taken into account by simple consideration using the kinetic method. To address this unexpected observation more carefully, we investigated the ER-CID of $(C:G:H)^+$, $(1\text{-MeC}:G:H)^+$, and $(5\text{-MeC}:G:H)^+$ both in multiple and near-single collision conditions. Methylation of C was exploited to vary the PA of C

Table 1. Predicted Thermochemical Properties for Protonation of Monomers (298.15 K); Protonation Was Considered at the N3-Position of C and the N7-Position of G. The Values Obtained for Protonation at the O2-Position of C Are Also Given in Parenthesis (B3LYP/6-311+G(2d,p)) (kcal/mol) (see also Table 1S)

	PA	ΔH_p	ΔG_p
C	226.75 (227.00)	226.78 (227.12)	226.68 (226.79)
1-MeC	230.32 (229.90)	230.26 (229.96)	230.37 (229.76)
5-MeC	229.28 (230.40)	229.27 (230.47)	229.24 (230.21)
G	228.28	228.06	228.51

moiety systematically, while keeping G as a common counterpart for protonated Cs. The predicted energy order of PA was 1-MeC > 5-MeC > G > C (Table 1), which is in a good agreement with the TCID results [33, 35]. (1-MeC:C:H)⁺ was also examined for comparison, as its structure and threshold dissociation have been well understood from previous IRMPD and TCID studies [10, 35].

Lowest Energy Structures: Proton-Bound Hoogsteen Base Pairs

To gain insight into the structures and energies of protonated complexes of (C:G:H)⁺, (1-MeC:G:H)⁺, and (5-MeC:G:H)⁺, which were produced by ESI in this study, we carried out a theoretical study at the B3LYP/6-311+G(2d,p) theory level. Figure 2S displays the six lowest energy conformers for (G:C:H)⁺, which were adopted from the previous study [50], along with the predicted energy differences of ΔE_0 and ΔG obtained at the present theory level from the most stable conformer, the proton-bound Hoogsteen base pair (I). Predicted energetics for the lowest energy structures can be found in Table 2S. Briefly, an ionic hydrogen bond is involved in the base pairing of all the lowest energy structures. The Hoogsteen base pair (I) is the most stable structure; other conformers (II–VI) lie far above, at least 2.5 kcal/mol higher, in energy (ΔE_0). The other conformers of II–VI are located very close to one another within the energy range of 0.5 kcal/mol. Their energy order varies slightly depending on the employed theory levels [50]. Some conformers are related by a proton transfer reaction in the ionic hydrogen bond such as I and IV as well as II and V. In the Hoogsteen base pair (I), the proton is attached to the N3-position of C rather than to the N7-position of G moiety with a PA that is higher by 1.5 kcal/mol [20]. In II, the structure that is antiparallel to I, the proton is attached to the N3-position of the G moiety rather than the N3-position of C in V. While the N3-position is the frequent protonation site in the C moiety, O2-protonated C is also found to participate in the base pairing as in III.

In this study, we considered the exhibited structures of the six lowest conformers as favorable ways of base pairing between C- and G-based molecules; we extended this idea to examine the proton-bound base pairs of (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺, as they differ only in methylation at C. The calculated results for the conformers of (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺ are presented in Figures 3S and 4S, with predicted energetics summarized in Tables 3S and 4S, respectively. As shown in the results for (1-MeC:G:H)⁺, base pairing of

Hoogsteen type represented the most stable conformer (I). The other conformers lie more than 2.8 kcal higher in energy and are very close to one another, within 0.5 kcal/mol. Similarly, the proton-bound Hoogsteen base pair is also the most stable conformer for (5-MeC:G:H)⁺ and is at least 2.7 kcal/mol more stable than the next most stable conformer (II). Solvent effects considered by PCM calculations suggested that the Hoogsteen base pairs are one of the most stable conformers in the aqueous environment. In addition, it was predicted that the protonated WC G-C base pairs were far less stable by 4.9, 5.9, and 5.5 kcal/mol (ΔE_0) than their respective Hoogsteen base pairs of C:H⁺⋯G (I), 1-MeC:H⁺⋯G (I), and 5-MeC:H⁺⋯G (I). Therefore, this theoretical study suggests that the protonated Hoogsteen base pairs may be the most stable and thus the most probable conformers in the gas phase.

The optimized structures of proton-bound Hoogsteen base pairs are given in Figure 2. The dissociation energies were calculated and are given in Table 2; these values are predicted to be close in energy, in the order of (1-MeC:C:H)⁺ (39.5 kcal/mol in D_0) > (5-MeC:G:H)⁺ (38.9 kcal/mol) > (C:G:H)⁺ (38.2 kcal/mol) > (1-MeC:G:H)⁺ (38.0 kcal/mol), if the dissociation to formation of lowest energy fragments between the two competitive channels was considered. While dissociation of (1-MeC:C:H)⁺, as found in TCID studies, does not give rise to the formation of O2-protonated C, C(O2):H⁺ [33, 35], the dissociation energies for the fragmentation channels involving formation of C(O2):H⁺ were also evaluated (Table 2). In fact, O2-protonated molecules of C and nucleosides have been known to exist in the gas phase [65, 66].

ER-CID Experiments: Anomalous CID Behaviors of Proton-Bound Hoogsteen Base Pairs

ER-CID in multiple collision conditions offers a reliable probe of the relative stability of protonated biomolecules and complexes [18, 38, 39]. In this study, we employed the ER-CID method to examine the proton-bound base pairs produced by ESI. Briefly, proton-bound base pairs were exposed to multiple collisions with He (~1 mTorr, 30 ms) in a linear ion trap, and the resulting CID spectra were examined in terms of S.Y. and F.Y. as a function of N.C.E. (a.u.). The CID_{50%} value obtained for S.Y. represents the energy required to deplete 50% of the parent base pairs; S.Y. thus serves as a parameter indicating the relative stability of the complexes. The plots of S.Y. of (C:G:H)⁺, (1-MeC:G:H)⁺, and (5-MeC:G:H)⁺ as a function of N.C.E. are given in Figure 3; least square fittings gave the CID_{50%} values of 10.9, 10.6, and 10.3, respectively. These are lower than the

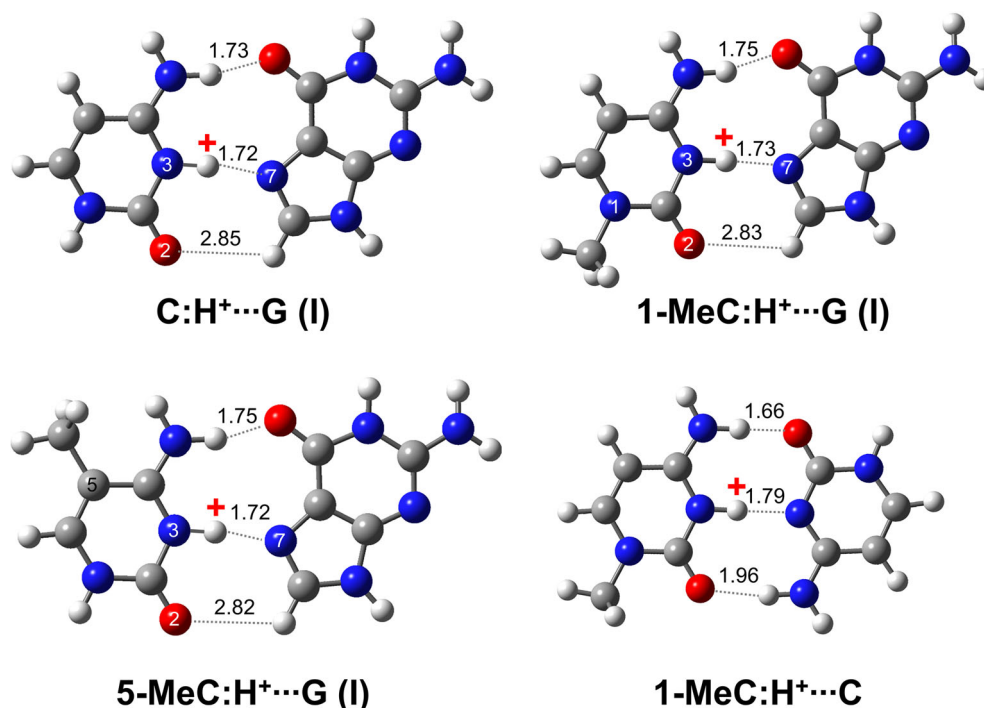


Figure 2. The predicted lowest energy structures, proton-bound Hoogsteen base pairs, for $(\text{C}:\text{G}:\text{H})^+$, $(1\text{-MeC}:\text{G}:\text{H})^+$, and $(5\text{-MeC}:\text{G}:\text{H})^+$ as well as $(1\text{-MeC}:\text{C}:\text{H})^+$. Plus sign denotes the attached proton

$\text{CID}_{50\%}$ value of 11.8 obtained for the triply-bound $(1\text{-MeC}:\text{C}:\text{H})^+$ as well. In fact, the measured $\text{CID}_{50\%}$ values for the proton-bound base pairs are very close, which suggests that their base-pairing strengths are comparable, perhaps via the same type of base-pairing interaction, i.e., Hoogsteen type, as predicted in the theoretical study (Figure 2). However, the order deviates slightly from the predicted energy order (D_0) of $(1\text{-MeC}:\text{C}:\text{H})^+ > (5\text{-MeC}:\text{G}:\text{H})^+ > (\text{C}:\text{G}:\text{H})^+ > (1\text{-MeC}:\text{G}:\text{H})^+$.

The ER-CID data were further examined in terms of F.Y. In this CID study of weakly bound non-covalent base pairs, the excitation energy range was quite low, so only the main channel products of $\text{C}:\text{H}^+$ (or methylated $\text{C}:\text{H}^+$) and $\text{G}:\text{H}^+$ appeared in the CID spectra. Figure 4 displays F.Y. as a function of N.C.E. The dissociation of $(1\text{-MeC}:\text{G}:\text{H})^+$ and $(5\text{-MeC}:\text{G}:\text{H})^+$ in multiple collision conditions showed predominant generation of $1\text{-MeC}:\text{H}^+$ ($\Delta\text{PA}_0 = \text{PA}_0(1\text{-MeC}) - \text{PA}_0(\text{G}) = 2.0 \text{ kcal/mol} > 0$)

and $5\text{-MeC}:\text{H}^+$ ($\Delta\text{PA}_0 = 2.1 \text{ kcal/mol} > 0$), which are the moieties with higher PA values than that of the counterpart, G. The observed formation of $\text{G}:\text{H}^+$ was very weak at only a few percent of the population. The dissociation of $(1\text{-MeC}:\text{C}:\text{H})^+$ was also similar, and $1\text{-MeC}:\text{H}^+$ ($\Delta\text{PA}_0 = 3.6 \text{ kcal/mol} > 0$) was pronounced in the CID spectra. Accordingly, the results qualitatively agree well with those expected from the kinetic method. However, the CID of $(\text{C}:\text{G}:\text{H})^+$ is drastically different and indeed anomalous as noticed in the previous experiments [17, 49]. In the ER-CID of $(\text{C}:\text{G}:\text{H})^+$, formation of $\text{C}:\text{H}^+$ ($\Delta\text{PA}_0 = -1.3 \text{ kcal/mol} < 0$) was more pronounced than that of $\text{G}:\text{H}^+$, despite the smaller PA of C moiety. The observed ratio of $\text{C}:\text{H}^+$ to $\text{G}:\text{H}^+$ was about 5.5:4.5. The experiments were performed using the acidic sample solution (pH = 3.1). However, the anomaly, i.e., formation of the most abundant $\text{C}:\text{H}^+$ fragments, was still observed even when using the sample solution without the addition of acid (pH = 6.4), which suggests that proton-bound WC base pairs were not significantly produced under our experimental conditions (Figure 5S).

Table 2. Predicted Energetics for Dissociation of the Proton-Bound Hoogsteen Base Pairs of $(\text{C}:\text{G}:\text{H})^+$, $(1\text{-MeC}:\text{G}:\text{H})^+$, and $(5\text{-MeC}:\text{G}:\text{H})^+$ (B3LYP/6-311+G(2d,p)) (kcal/mol)

	D_c	D_0	ΔG
$\text{C}:\text{H}^+\cdots\text{G} \rightarrow \text{C}(\text{N}3):\text{H}^+ + \text{G}$	40.70	39.74	27.77
($\rightarrow \text{C}(\text{O}2):\text{H}^+ + \text{G}$)	(40.33)	(39.49)	(27.66)
$\rightarrow \text{C} + \text{G}(\text{N}7):\text{H}^+$	39.67	38.21	25.94
$1\text{-MeC}:\text{H}^+\cdots\text{G} \rightarrow 1\text{-MeC}(\text{N}3):\text{H}^+ + \text{G}$	38.98	37.97	26.03
($\rightarrow 1\text{-MeC}(\text{O}2):\text{H}^+ + \text{G}$)	(39.21)	(38.40)	(26.03)
$\rightarrow 1\text{-MeC} + \text{G}(\text{N}7):\text{H}^+$	41.42	40.02	27.89
$5\text{-MeC}:\text{H}^+\cdots\text{G} \rightarrow 5\text{-MeC}(\text{N}3):\text{H}^+ + \text{G}$	39.80	38.93	27.07
($\rightarrow 5\text{-MeC}(\text{O}2):\text{H}^+ + \text{G}$)	(38.62)	(37.81)	(26.09)
$\rightarrow 5\text{-MeC} + \text{G}(\text{N}7):\text{H}^+$	41.23	39.93	27.80
$1\text{-MeC}:\text{H}^+\cdots\text{C} \rightarrow 1\text{-MeC}(\text{N}3):\text{H}^+ + \text{C}$	41.21	39.47	27.01
$\rightarrow 1\text{-MeC} + \text{C}(\text{N}3):\text{H}^+$	44.68	43.04	30.70

To examine the intriguing CID behavior of $(\text{C}:\text{G}:\text{H})^+$, CID experiments were further performed in near-single collision conditions using a quadrupole tandem mass spectrometer. The CID data taken at the collision gas (Ar) pressure of 0.10 mTorr are given in Figure 5, in which it can be seen that these are essentially the same results observed in multiple collision experiments. The results of $(1\text{-MeC}:\text{C}:\text{H})^+$ accord well with the previous TCID experiments [35], where formation of $1\text{-MeC}:\text{H}^+$ ($\Delta\text{PA}_0 > 0$) was dominant over that of $\text{C}:\text{H}^+$. Again, dissociation in the near-single collision conditions also led to pronounced formations of $1\text{-MeC}:\text{H}^+$ and $5\text{-MeC}:\text{H}^+$ for $(1\text{-MeC}:\text{G}:\text{H})^+$ and $(5\text{-MeC}:\text{G}:\text{H})^+$, respectively, which are

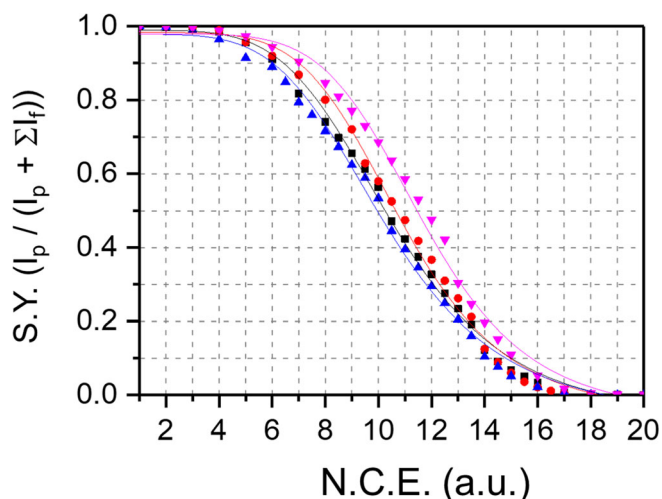


Figure 3. S.Y. data as a function of N.C.E. (a.u.) obtained by ER-CID in multiple collision conditions for (C:G:H)⁺ (red circle), (1-MeC:G:H)⁺ (black square), (5-MeC:G:H)⁺ (blue up-pointing triangle), and (1-MeC:C:H)⁺ (pink down-pointing triangle)

moieties with level of PA higher than that of G. As for the CID of (C:G:H)⁺; the intriguing aspect of CID of producing more abundant C:H⁺ is still evident. Therefore, it can be suggested that there might be certain causes involved in the dissociation process of (C:G:H)⁺; this will require further understanding of the potential energy surface of (C:G:H)⁺ along the dissociation

coordinates. The stronger formation of C:H⁺, which was a characteristic of the proton-bound Hoogsteen base pair identified by IRMPD [17], indicates that the (C:G:H)⁺ produced under acidic conditions in this study is indeed the Hoogsteen complex.

Potential Energy Surfaces: Facile Proton Transfer Leading to O-Protonated C Formation

In an effort to understand the dissociation process of the proton-bound Hoogsteen base pair, we theoretically explored the potential energy surface of (C:G:H)⁺. The resulting potential energy diagrams are given in Figure 6, presented in terms of ΔE_c and ΔG (298.15 K) from the lowest energy structure, C:H⁺⋯G (I). The proton-bound Hoogsteen base pair is related to the IV structure by an intermolecular proton transfer from the N3-position of C to the N7-position of G with a low PT energy barrier of 4.0 kcal/mol via TS_{PT(N7)} [17, 20, 50]. In fact, this PT offers a pathway relevant to the dissociation into the product state of C + G:H⁺, in which the proton is transferred to the G fragment. When a potential energy scan was performed by adiabatically elongating the intermolecular distance between the two moieties of the Hoogsteen base pair, the energy gradually increased until the product state of C(N3):H⁺ + G was reached, where no energy barrier for

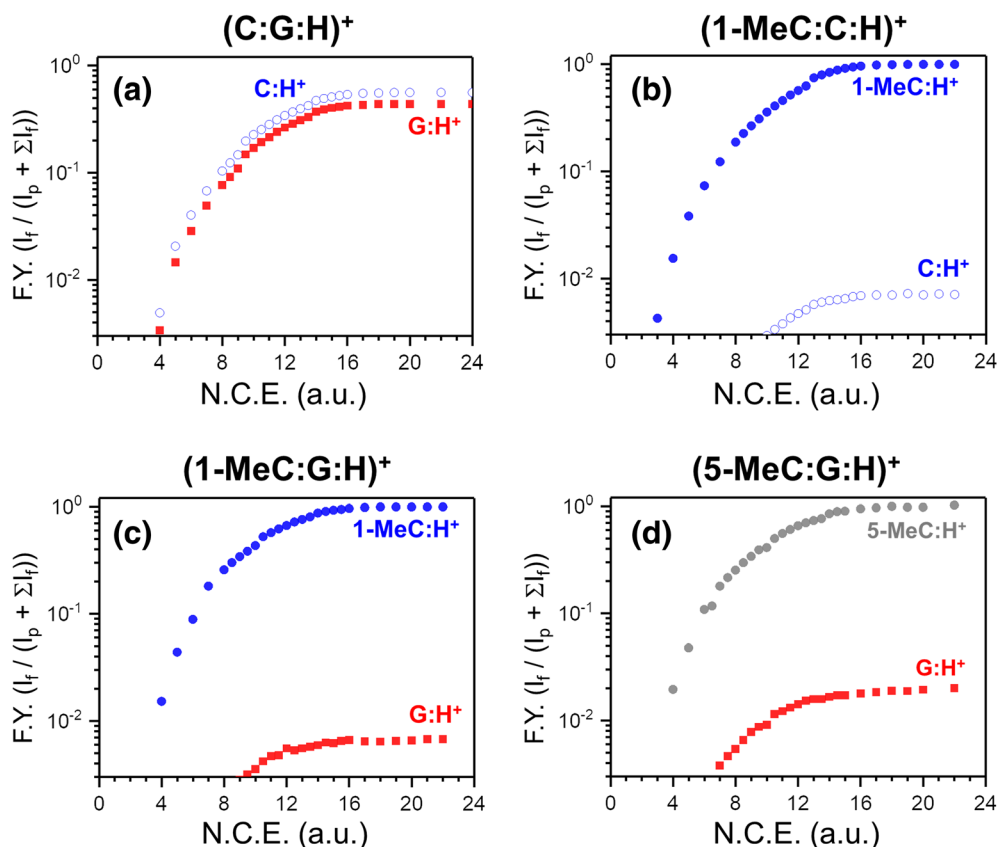


Figure 4. F.Y. data as a function of N.C.E. (a.u.) obtained by ER-CID in multiple collision conditions for (C:G:H)⁺, (1-MeC:G:H)⁺, (5-MeC:G:H)⁺, and (1-MeC:C:H)⁺

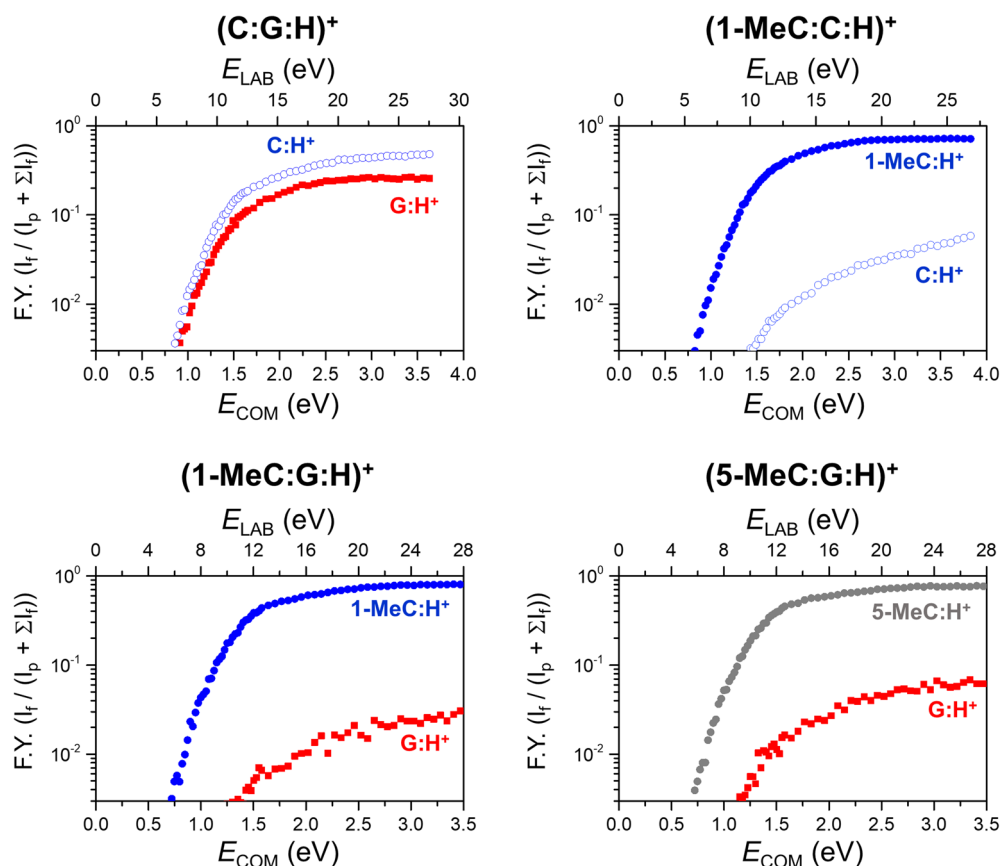


Figure 5. F.Y. data as a function of E_{com} (eV) obtained by ER-CID in near-single collision conditions for $(\text{C}:\text{G}:\text{H})^+$, $(1\text{-MeC}:\text{G}:\text{H})^+$, $(5\text{-MeC}:\text{G}:\text{H})^+$, and $(1\text{-MeC}:\text{C}:\text{H})^+$. The collision gas (Ar) pressure was 0.10 mTorr

dissociation was noticeable. Both represent the two dissociation channels that were considered as the competitive pathways of dissociation for the proton-bound Hoogsteen

base pair; $\text{C}:\text{H}^+\cdots\text{G} \rightarrow \text{C} + \text{GH}^+$ and $\text{C}:\text{H}^+\cdots\text{G} \rightarrow \text{C}(\text{N}3):\text{H}^+ + \text{G}$ (Figure 1), for which the observed branching ratio was not fully understood yet [17, 49].

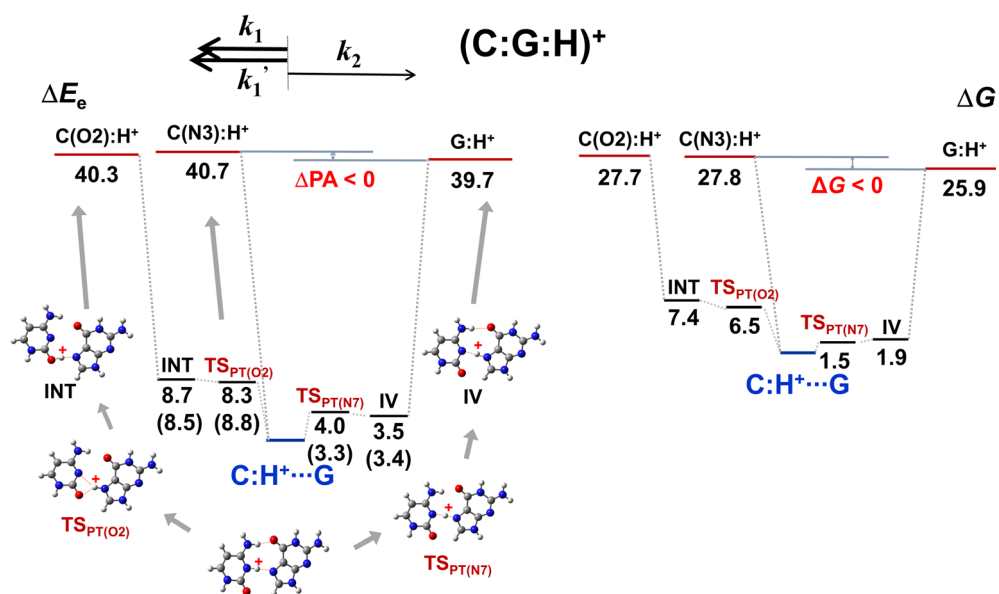


Figure 6. Potential energy diagrams for $(\text{C}:\text{G}:\text{H})^+$ constructed in terms of ΔE_e and ΔG (298.15 K). The energy values (kcal/mol) are calculated at the B3LYP/6-311+G(2d,p) theory level. The MP2/aug-cc-pVDT energies obtained at the B3LYP optimized geometries are also presented in parenthesis

On the other hand, protonation at the N3-position of C is the most important aspect in base pairing, as found in DNA i-motif and Hoogsteen base pairing. However, protonation of the O2-position in isolated C molecules is also known in the gas phase [65, 66], which is energetically even more favorable than that of N3 in C molecule (Table 1). Although the fragmentation to the product state forming C(O2):H⁺ may thus be possible, it was found that the dissociation to the product state of 1-MeC(O2):H⁺ + C was not involved in the TCID of (1-MeC:C:H)⁺. To produce the 1-MeC(O2):H⁺ fragment, the proton has to migrate from the N3-position to the O2-position of C via the collisional excitation. However, to transfer the proton from the N3 to O2-position of C in the dissociation process, the reaction needs to surmount a PT barrier of 43.4 kcal/mol, which is approximately 3 kcal/mol larger than the threshold dissociation energy to the corresponding product state. Thus, the fragmentation pathway yielding the 1-MeC(O2):H⁺ fragment was not available in the TCID of (1-MeC:C:H)⁺ [35].

In the course of exploring the potential energy surface of (C:G:H)⁺, we also examined the possible PT from the N3 to the O2-position of C in the dissociation of base pairs. Interestingly, in contrast to the (1-MeC:C:H)⁺ case, it was found there is a PT pathway with a low energy barrier of about 8 kcal/mol (TS_{PT(O2)}), which migrates the proton from the N3-position of C in the Hoogsteen geometry to the O2-position of C in a loosely bound complex formed by a single ionic hydrogen bond (INT), through which further potential energy scan smoothly dissociated the complex to the product state of C(O2):H⁺ + G. When compared with the dissociation energy threshold of 40.3 kcal/mol, the TS_{PT(O2)} is very low (8 kcal/mol). This is easily surmountable in the energy regime in which actual collision-induced dissociation can take place. In Figure 7, the potential energy diagrams for (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺ are also displayed. The theoretical study also indicated facile PTs upon activation of the base pairs of (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺, of which the energetics are similar to that of (C:G:H)⁺.

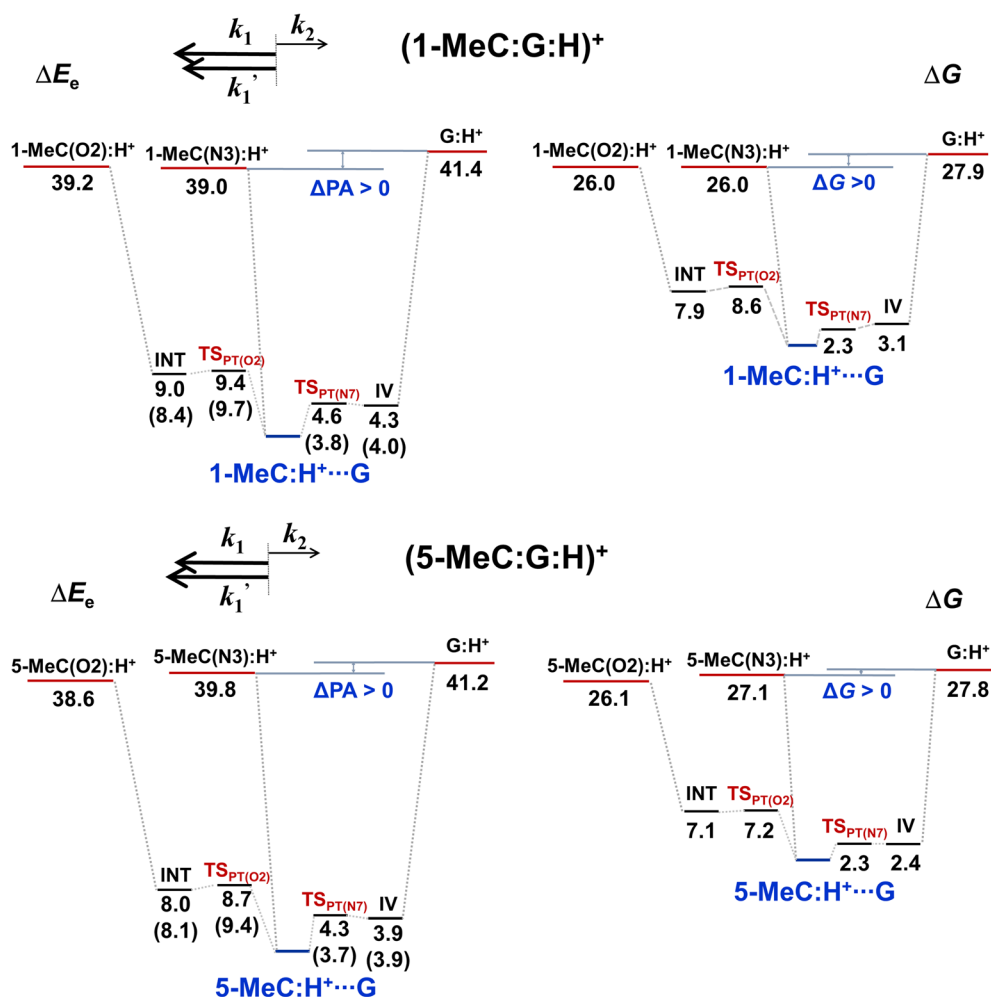


Figure 7. Potential energy diagrams for (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺ constructed in terms of ΔE_e and ΔG (298.15 K). The energy values (kcal/mol) are calculated at the B3LYP/6-311+G(2d,p) theory level. The MP2/aug-cc-pVDZ energies obtained at the B3LYP optimized geometries are also presented in parenthesis

This finding of the presence of an additional fragmentation pathway producing O2-protonated C, C(O2):H⁺, accounts for many important questions in this study. First, although mass spectrometry, which measures only ion masses, cannot characterize observed protonated C fragments, C:H⁺, it is most likely that both N3- and O2-protonated Cs are produced and coexist in the observed C:H⁺ fragments. Second, the availability of a new product channel of O2-protonated C accounts for the deviation in the order of relative stability (CID_{50%}), (1-MeC:C:H)⁺ > (C:G:H)⁺ > (1-MeC:G:H)⁺ > (5-MeC:G:H)⁺, from that of predicted *D*₀. With the additional accessible channel to O2-protonated C, the energy order for the lowest product channels is altered (1-MeC:C:H)⁺ (39.5 kcal/mol) > (C:G:H)⁺ (38.2 kcal/mol) > (1-MeC:G:H)⁺ (38.0 kcal/mol) > (5-MeC:G:H)⁺ (37.8 kcal/mol), which now agrees well with the CID_{50%} results. In addition to the theoretical prediction of the lowest energy structures (Figure 2), the agreement of relative stability (CID_{50%}) with the predicted energetics (ΔD_0) supports the idea that the protonated base pairs produced by ESI in this study are indeed proton-bound Hoogsteen base pairs of C:H⁺⋯G, 1-MeC:H⁺⋯G, and 5-MeC:H⁺⋯G.

Finally, the anomaly in the CID of (C:G:H)⁺ was the contradiction to the prediction made using two competitive channels. However, this finding of an additional reaction channel effectively doubles the number of available states in the TS to C:H⁺. This greatly promotes the apparent reaction rate for production of C:H⁺ fragments. This explains the observed anomaly in dissociation of (C:G:H)⁺ and pronounced production of protonated C fragments in the CID of (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺ as well.

As for ER-CID of (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺, formation of the most abundant (1-MeC:H)⁺ and (5-MeC:H)⁺ does not appear to be anomalous as the PAs of 1-MeC and 5-MeC are larger than that of the counterpart, G ($\Delta PA > 0$). However, the branching ratios of [1-MeC:H⁺]/[G:H⁺]⁺ and [5-MeC:H⁺]/[G:H⁺]⁺ are measured to be very large, 149 and 51, respectively (Figure 4). If it is assumed that the three competitive dissociation pathways leading to formation of N-protonated C (*k*₁), O-protonated C (*k*₁'), and protonated G (*k*₂) are completely independent, the basic relation of kinetic method can be extended as $\ln((k_1 + k_1')/k_2) = \ln([x\text{-MeC:H}^+]/[G:H^+]) \approx (\Delta G_{k_1} + \Delta G_{k_1'} - \Delta G_{k_2})/RT_{\text{eff}}$. Assuming that *T*_{eff} is 200 K, the branching ratios are estimated to be very large, 131 and 78 for [1-MeC:H⁺]/[G:H⁺]⁺ and [5-MeC:H⁺]/[G:H⁺]⁺, respectively, using the calculated ΔG_p values (Table 1S). This explains the large observed branching ratios, which further reveals the same *anomalous* CID behaviors for (1-MeC:G:H)⁺ and (5-MeC:G:H)⁺ as found for the proton-bound Hoogsteen base pair of (C:G:H)⁺.

Conclusion

In this study, ER-CID of (C:G:H)⁺ exhibited an anomaly in fragment abundances between two competitive fragmentation pathways of generating C:H⁺ and G:H⁺ fragments. In contrast to the other ER-CID results for (1-MeC:G:H)⁺ and (5-

MeC:G:H)⁺ ($\Delta PA > 0$), in which the protonated fragments of methylated Cs were more abundant than the common ion, G:H⁺, protonated C was still more pronounced in the ER-CID of (C:G:H)⁺, despite the lower PA of C ($\Delta PA < 0$). However, in further theoretical study on the potential energy surfaces of the proton-bound Hoogsteen base pairs, it was found that in contrast to the triply-bound complex of (1-MeC:C:H)⁺, upon collisional excitation, proton transfers are feasible not only from the N3-position of C to the N7-position of G (intermolecular PT barrier of 4 kcal/mol) but also from the N3-position to the O2-position of C (intramolecular PT barrier of 8 kcal/mol). This suggests that the proton in the ionic hydrogen bond is rather mobile in the excited complexes, leading to formation of both N3- and O2- protonated Cs. This availability of an extra channel for C:H⁺, which suggests two-fold degeneracy or increased number of available states in TS, may give rise to pronounced formation of C:H⁺ product. It can therefore be considered that it is the proton transfer that accounts for the apparent anomaly observed in the CID of the proton-bound Hoogsteen base pair of C:H⁺⋯G.

Acknowledgements

This work was supported by MSIT via NRF (Grant No. NRF-2016R1D1A1B03931987). This work was also supported by MOTIE via KEIT (Grant No. 10063335). S.Y.H. is thankful to Prof. Han Bin Oh (Sogang University) for his valuable comments on this research subject.

References

- Day, H.A., Pavlou, P., Waller, Z.A.E.: i-Motif DNA: structures, stability, and targeting ligands. *Bioinorg. Med. Chem.* **22**, 4407–4418 (2014)
- Gehring, K., Leroy, J.-L., Guéron, M.: A tetrameric DNA structure with protonated cytosine–cytosine base pairs. *Nature*. **363**, 561–565 (1993)
- Darlow, J.M., Leach, D.R.F.: Secondary structure in d(CGG) and d(CCG) repeat tracts. *J. Mol. Biol.* **275**, 3–16 (1998)
- Fotjik, P., Vorlickova, M.: The fragile X chromosome (GCC) repeat folds into a DNA tetraplex at neutral pH. *Nucleic Acids Res.* **29**, 4684–4690 (2001)
- Armentano, D., De Munno, G., Di Donna, L., Sindona, G., Giorgi, G., Salvini, L., Napoli, A.: Self-assembling of cytosine nucleoside into triply-bound dimers in acid media. A comprehensive evaluation of proton-bound pyrimidine nucleosides by electrospray tandem mass spectrometry, X-rays diffractometry, and theoretical calculations. *J. Am. Soc. Mass Spectrom.* **15**, 268–279 (2004)
- Šponer, J., Leszczynski, J., Vetterl, V., Hobza, P.: Base stacking and hydrogen bonding in protonated cytosine dimer: the role of molecular ion-dipole and induction interactions. *J. Biomol. Struct. Dyn.* **13**, 695–706 (1996)
- Han, S.Y., Oh, H.B.: Theoretical study of the ionic hydrogen bond in the isolated proton-bound dimer of cytosine. *Chem. Phys. Lett.* **432**, 269–274 (2006)
- Oomens, J., Moehlig, A.R., Morton, T.H.: Infrared multiple photon dissociation (IRMPD) spectroscopy of the proton-bound dimer of 1-methylcytosine in the gas phase. *J. Phys. Chem. Lett.* **1**, 2891–2897 (2010)
- Ung, H.U., Moehlig, A.R., Kudla, R.A., Mueller, L.J., Oomens, J., Berden, G., Morton, T.H.: Proton-bound dimers of 1-methylcytosine and its derivatives: vibrational and NMR spectroscopy. *Phys. Chem. Chem. Phys.* **15**, 19001–19012 (2013)

10. Yang, B., Wu, R.R., Berden, G., Oomens, J., Rodgers, M.T.: Infrared multiple photon dissociation action spectroscopy of proton-bound dimers of cytosine and modified cytosines: effects of modifications on gas-phase conformations. *J. Phys. Chem. B* **117**, 14191–14201 (2013)
11. Kwon, S., Oh, H.B., Han, S.Y.: Infrared multiple photon depletion of the gas-phase proton-bound cytosine dimer. *Chem. Lett.* **44**, 1756–1758 (2015)
12. Oh, H.B., Lin, C., Hwang, H.Y., Zhai, H.Y., Breuker, K., Zbrouskov, V., Carpenter, B.K., McLafferty, F.W.: Infrared photodissociation spectroscopy of electrosprayed ions in a Fourier transform mass spectrometer. *J. Am. Chem. Soc.* **127**, 4076–4083 (2005)
13. Eyler, J.R.: Infrared multiple photon dissociation spectroscopy of ions in penning traps. *Mass Spectrom. Rev.* **28**, 448–467 (2009)
14. Fridgen, T.D.: Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes. *Mass Spectrom. Rev.* **28**, 586–607 (2009)
15. Chiavarino, B., Crestoni, M.E., Fornarini, S., Scuderi, D., Salpin, J.-Y.: Interaction of cisplatin with adenine and guanine: a combined IRMPD, MS/MS, and theoretical study. *J. Am. Chem. Soc.* **135**, 1445–1455 (2013)
16. Rajabi, K., Theel, K., Gillis, E.A.L., Beran, G., Fridgen, T.D.: The structure of the protonated adenine dimer by infrared multiple photon dissociation spectroscopy and electronic structure calculations. *J. Phys. Chem. A* **113**, 8099–8107 (2009)
17. Cruz-Ortiz, A.F., Rossa, M., Berthias, F., Berdakin, M., Maitre, P., Pino, G.A.: Fingerprints of both Watson-Crick and Hoogsteen isomers of the isolated (cytosine-guanine)H⁺ pair. *J. Phys. Chem. Lett.* **8**, 5501–5506 (2017)
18. Azargun, M., Fridgen, T.D.: Guanine tetrads: an IRMPD spectroscopy, energy resolved SORI-CID, and computational study of M(9-ethylguanine)₄⁺ (M = Li, Na, K, Rb, Cs) in the gas phase. *Phys. Chem. Chem. Phys.* **17**, 25778–257853 (2015)
19. 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.* **51**, 4067–4070 (2012)
20. Han, S.Y., Lee, S.H., Chung, J., Oh, H.B.: Base-pair interactions in the gas-phase proton-bonded complexes of C⁺G and C⁺GC. *J. Chem. Phys.* **127**, 245102–245110 (2007)
21. Povsic, T.J., Dervan, P.B.: Triple helix formation by oligonucleotides on DNA extended to the physiological pH range. *J. Am. Chem. Soc.* **111**, 3059–3061 (1989)
22. Hunter, W.N., Brown, T., Anand, N.N., Kennard, O.: Structure of an adenine cytosine base pair in DNA and its implications for mismatch repair. *Nature* **320**, 552–555 (1986)
23. Florián, J., Baumruk, V., Leszczynski, J.: IR and Raman spectra, tautomeric stabilities, and scaled quantum mechanical force fields of protonated cytosine. *J. Phys. Chem.* **100**, 5578–5589 (1996)
24. Cooks, R.G., Koskinen, J.T., Thomas, P.D.: The kinetic method of making thermochemical determinations. *J. Mass Spectrom.* **34**, 85–92 (1999)
25. Cooks, R.G., Patrick, J.S., Kotiaho, T., McLuckey, S.A.: Thermochemical determinations by the kinetic method. *Mass Spectrom. Rev.* **13**, 287–339 (1994)
26. Cheng, X., Wu, Z., Fenselau, C.: Collision energy dependence of proton-bound dimer dissociation: entropy effects, proton affinities, and intramolecular hydrogen-bonding in protonated peptides. *J. Am. Chem. Soc.* **115**, 4844–4848 (1993)
27. Armentrout, P.B.: Mass spectrometry – not just a structural tool: the use of guided ion beam tandem mass spectrometry to determine thermochemistry. *J. Am. Soc. Mass Spectrom.* **13**, 419–434 (2002)
28. 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* **105**, 2374–2383 (2001)
29. Romanov, V., Verkerk, U.H., Siu, C.-K., Hopkinson, A.-C., Siu, K.W.M.: Threshold collision-induced dissociation measurements using a ring ion guide as the collision cell in a triple-quadrupole mass spectrometer. *Anal. Chem.* **81**, 6805–6812 (2009)
30. Fang, Y., Liu, J.: Reaction of protonated tyrosine with electronically excited singlet molecular oxygen ($a^1\Delta_g$): an experimental and trajectory study. *J. Phys. Chem. A* **113**, 11250–11261 (2009)
31. Muntean, F., Armentrout, P.B.: Guided ion beam study of collision-induced dissociation dynamics: integral and differential cross sections. *J. Chem. Phys.* **115**, 1213–1228 (2001)
32. Rodgers, M.T., Ervin, K.M., Armentrout, P.B.: Statistical modeling of collision-induced dissociation thresholds. *J. Chem. Phys.* **106**, 4499–4508 (1997)
33. 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.* **85**, 11000–11006 (2013)
34. 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.* **136**, 282–290 (2014)
35. Yang, B., Moehlig, A.R., Frieler, C.E., Rodgers, M.T.: Base-pairing energies of protonated nucleobase pairs and proton affinities of 1-methylated cytosines: model systems for the effects of the sugar moiety on the stability of DNA i-motif conformations. *J. Phys. Chem. B* **119**, 1857–1868 (2015)
36. 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.* **26**, 1469–1482 (2015)
37. Yang, B., Rodgers, M.T.: Base-pairing energies of protonated nucleoside base pairs of dCyd and m⁵dCyd: implications for the stability of DNA i-motif conformations. *J. Am. Soc. Mass Spectrom.* **26**, 1394–1403 (2015)
38. Zhu, Y., Roy, H.A., Cunningham, N.A., Strobehn, S.F., Gao, J., Munshi, M.U., Berden, G., Oomens, J., Rodgers, M.T.: IRMPD action spectroscopy, ER-CID experiments, and theoretical studies of sodium cationized thymidine and 5-methyluridine: kinetic trapping during the ESI desolvation process preserves the solution structure of [Thd+Na]⁺. *J. Am. Soc. Mass Spectrom.* **28**, 2423–2437 (2017)
39. Zhu, Y., Hamlow, L.A., He, C.C., Lee, J.K., Gao, J., Berden, G., Oomens, J., Rodgers, M.T.: Gas-phase conformations and N-glycosidic bond stabilities of sodium cationized 2'-deoxyguanosine and guanosine: sodium cations preferentially bind to the guanine residue. *J. Phys. Chem. B* **127**, 4048–4060 (2017)
40. Hoogsteen, K.: The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **16**, 907–916 (1963)
41. Rajagopal, P., Feigon, J.: Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)₄ and d(T-C)₄. *Nature* **339**, 637–640 (1989)
42. Pramanik, S., Nakamura, K., Usui, K., Nakano, S.-I., Saxena, S., Matsui, J., Miyoshi, D., Sugimoto, N.: Thermodynamic stability of Hoogsteen and Watson-Crick base pairs in the presence of histone H3-mimicking peptide. *Chem. Commun.* **47**, 2790–2792 (2011)
43. Halder, A., Halder, S., Bhattacharyya, D., Mitra, A.: Feasibility of occurrence of different types of protonated base pairs in RNA: a quantum chemical study. *Phys. Chem. Chem. Phys.* **16**, 18383–18396 (2014)
44. Nikolova, E.N., Kim, E., Wise, A.A., O'Brien, P.J., Andricioaei, I., Al-Hashimi, H.M.: Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498–504 (2011)
45. Alvey, H.S., Gottardo, F.L., Nikolova, E.N., Al-Hashimi, H.M.: Widespread transient Hoogsteen base pairs in canonical duplex DNA with variable energetics. *Nat. Commun.* **5**, 4786/1–4786/8 (2014)
46. Mirkin, S.M.: Expandable DNA repeats and human disease. *Nature* **447**, 932–940 (2007)
47. Gacy, A.M., Goellner, G.M., Spiro, C., Chen, X., Gupta, G., Bradbury, E.M., Dyer, R.B., Mikesell, M.J., Yao, J.Z., Johnson, A.J., Richter, A., Melancon, S.B., McMurray, C.T.: GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol. Cell.* **1**, 583–593 (1998)
48. Chubb, J.M., Hogan, M.E.: Human therapeutics based on triple helix technology. *Trends Biotechnol.* **10**, 132–136 (1992)
49. Seong, Y., Han, S.Y., Jo, S.-C., Oh, H.B.: An anomalous dissociation of protonated cluster ions of DNA guanine-cytosine base-pair. *Mass Spectrom. Lett.* **2**, 72–75 (2011)
50. Jun, J., Han, S.Y.: Theoretical exploration of gas-phase conformers of proton-bound non-covalent heterodimers of guanine and cytosine rare tautomers: structures and energies. *Theor. Chem. Accounts* **136**, 136–146 (2017)
51. Kertesz, T.M., Hall, L.H., Hill, D.W., Granta, D.F.: CE₅₀: quantifying collision induced dissociation energy for small molecule characterization and identification. *J. Am. Soc. Mass Spectrom.* **20**, 1759–1767 (2009)
52. Kenny, D.J., Worthington, K.R., Hoyes, J.B.: Scanwave: a new approach to enhancing spectral data on a tandem quadrupole mass spectrometer. *J. Am. Soc. Mass Spectrom.* **21**, 1061–1069 (2010)

53. Guan, S., Marshall, A.G.: Stacked-ring electrostatic ion guide. *J. Am. Soc. Mass Spectrom.* **7**, 101–106 (1996)
54. Wu, R.R., Rodgers, M.T.: O2 protonation controls threshold behavior for N-glycosidic bond cleavage of protonated cytosine nucleosides. *J. Phys. Chem. B.* **120**, 4803–4811 (2016)
55. Wu, R.R., Chen, Y., Rodgers, M.T.: Mechanisms and energetics for N-glycosidic bond cleavage of protonated 2'-deoxyguanosine and guanosine. *Phys. Chem. Chem. Phys.* **18**, 2968–2980 (2016)
56. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., 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., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J.: Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT (2013)
57. 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.* **100**, 65–73 (2002)
58. Tomasi, J., Mennucci, B., Cammi, R.: Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999–3094 (2005)
59. Peng, C., Schlegel, H.B.: Combining synchronous transit and quasi-Newton methods for finding transition states. *Israel J. Chem.* **33**, 449–454 (1993)
60. Jurečka, P., Šponer, J., Černý, J., Hobza, P.: Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **8**, 1985–1993 (2006)
61. Mejía-Mazariegos, L., Robles, J., García-Revilla, M.A.: Tautomerism in some pyrimidine nucleoside analogues used in the treatment of cancer: an ab initio study. *Theor. Chem. Accounts.* **135**, 233–244 (2016)
62. Hunter, E.P.L., Lias, S.G.: Evaluated gas phase basicities and proton affinities of molecules: an update. *J. Phys. Chem. Ref. Data.* **27**, 413–656 (1996)
63. Greco, F., Liguori, A., Sindona, G., Uccella, N.: Gas-phase proton affinity of deoxyribonucleosides and related nucleobases by fast atom bombardment tandem mass spectrometry. *J. Am. Chem. Soc.* **112**, 9092–9096 (1990)
64. Di Donna, L., Napoli, A., Sindona, G., Athanassopoulos, C.: A comprehensive evaluation of the kinetic method applied in the determination of the proton affinity of the nucleic acid molecules. *J. Am. Soc. Mass Spectrom.* **15**, 1080–1086 (2004)
65. Salpin, J.-Y., Guillaumont, S., Tortajada, J., MacAleese, L., Lemaire, J., Maitre, P.: Infrared spectra of protonated uracil, thymine and cytosine. *ChemPhysChem.* **8**, 2235–2244 (2007)
66. Filippi, A., Fraschetti, C., Rondino, F., Piccirillo, S., Steinmetz, V., Guidoni, L., Speranza, M.: Protonated pyrimidine nucleosides probed by IRMPD spectroscopy. *Int. J. Mass Spectrom.* **354–355**, 54–61 (2013)