PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

PAPER

Absolute electron total ionization cross-sections: Molecular analogues of DNA and RNA nucleobase and sugar constituents

James N Bull,[a,b] Jason W L Lee[a] and Claire Vallance[a,c]

Received (in XXX, XXX) Xth XXXXXXXXX 2014, Accepted Xth XXXXXXXXX 2014

DOI: 10.1039/b000000xt

Accurate ionization cross-sections for DNA and RNA constituents in the condensed or aqueous phase are important parameters for models simulating radiation damage to genetic material in living cells. In this work, absolute gas-phase electron total ionization cross-sections (TICSs) have been measured for a series of six aromatic and eight non-aromatic cyclic species that can be considered as prototype functional group analogues for the nucleobases and sugar backbone constituents of DNA and RNA. TICSs for water, hexane, and ethylacetamide (a peptide bond analogue) are also reported. The experimental apparatus utilizes a cylindrical ion collector that surrounds the ionization region, providing essentially unit detection efficiency. Two theoretical models, the polarizability-correlation method and binary-encounter Bethe theory, are able to reproduce the measured maximum TICS well for all species studied. An empirical energy-dependent correction is found to yield improvement in the agreement between experimental energy-dependent cross sections and the predictions of the BEB model. Having characterised and optimised the performance of both models, they are then used to predict TICSs for gas-phase DNA and RNA nucleobases and sugars. Direct experimental determinations of TICSs for these species are difficult because of their low volatility, which makes it difficult to prepare suitable gas-phase samples for measurement.

Introduction

The electron total ionization cross-section (TICS), $\sigma(E)$, for a given molecule is an energy-dependent measure of the total ionization efficiency.[1] Accurate *absolute* TICSs are important parameters for mass spectrometer calibration, modelling of plasma processes, and normalization of relative electron ionization measurements.[2] Over the last one and a half decades, Harland and co-workers have measured absolute TICSs for over one hundred C_1 to C_6 organic and halocarbon species, yielding one of the largest existing collections of such cross-sections – see, for example, references 3 and 4. The measurements employed an ion collector designed to surround the ionization region in order to achieve essentially unit detection efficiency. A recent review of theoretical models for electron ionization by Bull *et al*.[5] compared experimental and theoretical data for many of these species, and concluded that two relatively simple models are able to predict maximum TICSs to a precision of ~7%, provided that the required *ab initio* input parameters are of sufficiently high quality. The species included in the work of Harland and coworkers do not include any cyclic or aromatic molecules. However, absolute cross-sections measurements on these classes of molecule are of considerable interest, since they can be considered as prototypes for the nucleobase and sugar backbone constituents of DNA and RNA. Direct measurements of electron ionization cross-sections for DNA or RNA constituents are difficult because they are solids at room temperature, with melting points higher than their thermal decomposition or pyrolysis temperatures.

Knowledge of the collision dynamics between slow electrons and biomolecules or model analogues has been the subject of considerable interest over the last two decades, in the context of understanding radiation damage to biological cells. Starting just over a decade ago, Sanche and co-workers[6-9] have investigated experimentally the interaction of high-energy radiation (e.g. X-rays and γ-rays) with short strands of DNA, demonstrating that the majority of radiation damage occurs not due to a direct interaction with the incoming radiation, but through interactions with secondary (ballistic) low energy electrons, which are produced with kinetic energies up to a few tens of electronvolts. Most of these electrons are formed through interaction of the incoming radiation with water within the cell cytoplasm; inner-shell ionization, valence ionization, and intermolecular Coulombic decay processes produce around 40,000 low-energy electrons per MeV of incident radiation, with a mean electron energy of around 9 eV. [8,10-12] Because the genetic material in mammalian cells constitutes less than 1% of the cellular contents, low energy electron production events of the type described above vastly outnumber direct primary interactions between the incoming radiation and DNA within the cell nucleus. On the femtosecond timescale, the low energy electrons either dissociate or ionize water to produce ·OH radicals, are captured by DNA/RNA constituents to form a radical anion, or ionize and/or electronically excite DNA/RNA constituents[8,9,13], with the first two processes dominating for ~9 eV electrons. Ballistic low energy electrons become thermalized and solvated on the picosecond timescale, and the first stages of chemically-induced DNA damage, including strand breaks and/or base loss, occur on the microsecond to millisecond timescale. The initial site of ionization on a DNA or RNA substrate is not necessarily the site of subsequent radiation damage; the nascent radical cation can migrate reversibly along the DNA or RNA substrate over distances exceeding 200 Å, until eventually the cation becomes

trapped through an irreversible chemical reaction, for example, a 'GGG trap' inducing a strand break or base ejection.[14-16] So-called secondary chemical reaction damage through species other than water may also occur, for example, the ionization of chromosomal proteins or lipids, which then generate secondary species which in turn go on to chemically react with DNA.[8,9] The immense complexity of condensed-phase biological systems, and the ultrafast time scales involved, means that confidently identifying and quantifying the important mechanisms is an exceptionally challenging task. Developing a detailed understanding of such processes relies heavily on combined data from both experiments and theoretical models. Baccarelli *et al*.[13] remark "From the theoretical point of view, a greater effort is needed to better evaluate the ionization cross-sections in different media as well as methods for evaluating the yield of secondary species, the products of excitation, which may evolve along different paths and lead to different channels of biodamage and so on." Knowledge of accurate absolute cross-sections for the many possible low energy electron interactions, and the extent to which these cross-sections may be treated as additive, is therefore vital when attempting to develop the sophisticated radiation models needed for applications such as radiotherapy dosimetry.[9,14,17-19]

This paper reports what are believed to be the most reliable absolute gas-phase electron TICSs to date for a number of molecules that can be considered as models for various substituents of DNA and RNA chains. Cross-sections have been measured over the electron energy range from 10 eV to 285 eV in 5 eV increments. Theoretical calculations employing the polarizability correlation method and the binary-encounter Bethe (BEB) theory are shown to be in very good agreement with experiment when various corrections are applied. This allows the models to be used with confidence to predict TICSs for the actual DNA and RNA constituents. To the authors' knowledge, there have been no reliable reports on direct *absolute* TICS for DNA and RNA constituents in the gas phase. A number of studies have considered different theories for calculation of their TICSs;[20-25] however it is difficult to assess the accuracy or reliability of the theoretical approach without experimental comparisons for similar species, e.g. species containing multiple functional groups combined with saturated and aromatic rings.

Experimental

The apparatus used to measure all absolute TICS is a rebuilt version of the original instrument used by Harland and co-workers[4,26] (see Acknowledgements), and records a *gross* or current-counting TICS. The ionization cell, shown schematically in Fig. 1, is housed inside a vacuum chamber maintained at a background pressure of ~4×10^{-8} Torr. Electrons are emitted from a filament, and focused into a beam by an Einzel lens before passing through the collision region and terminating at a positively biased electron trap. The electron current is regulated to between 30 and 50 nA to avoid space-charge effects, and the beam is constrained to the axis of the ionization cell by a magnetic field maintained between two rare-earth permanent magnets (several thousand Gauss). The electron energy is controlled by the potential applied to the filament, and the beam has an energy FWHM of around 1 eV, characterised at several different electron energies through measurement of the beam current as a function of a variable retarding potential. The inner wall of the cylindrical ionization cell is coated with graphite aerosol in order to prevent surface scattering of incident ions, and serves as the ion collector, being held at a slightly negative potential relative to the electron beam axis. Due to the non-volatile and "sticky" nature of the samples considered in this study, it was necessary to dismantle and clean the cell after every few samples to prevent residual sample within the graphite layer leading to the build-up of contact potentials.

To perform a measurement, the sample gas of interest is admitted into the collision cell to a pressure of around 10^{-4} Torr (constantly measured using an MKS baratron model 690A capacitance manometer), and the electron and ion currents are recorded as a function of electron energy. The ion current is recorded using a Keithley model 486 picoammeter, and is typically between 2 - 10 nA at the peak of the ionization efficiency (TICS) curve. Apart from the 2 mm electron beam entrance aperture, the collision cell is gas tight. All electrical potentials are supplied by a computer-controlled multichannel power supply, and data acquisition is fully automated through a LabVIEW interface and LabJack model U3 digital-to-analogue converter. Temperature is monitored *via* a calibrated thermistor drilled and embedded into the exterior of the brass ion collector. The sample inlet line, variable leak valve (Granville-Phillips Co. series 203) and short tube connecting the ionization cell to the baratron are heated to a temperature of ~42°C, which corresponds to the average working temperature of the ionization cell and natural operating temperature of the baratron. The sample gas is assumed to be in thermal equilibrium with the walls of the sample line and ionization cell during measurements. The ionization cell was designed so that the positive ion collector surface essentially surrounds the ionization region, which increases the probability of ion detection to essentially unity – this condition was further confirmed through ion trajectory simulations carried out using the SIMION 8.0.4 software package.[27] The collection of all ions is an important consideration, particularly when fragment ions may exhibit high kinetic energy release (KER). Such ions are often discriminated against in mass spectrometers.[28-30]

The absolute TICS, $\sigma(E)$, is calculated from a Beer-Lambert type law. Under the assumption that the fraction of electrons lost from the beam due to ionizing collisions is small, this can be written

$$\frac{I_+}{I_-} = \frac{P\sigma(E)l}{k_B T}$$

where I_+ and I_- are the measured ion and electron currents, P and T are temperature and pressure of the sample gas, l is the electron path length allowing for cyclotron motion in the magnetic field within the collision cell, and k_B is the Boltzmann constant.

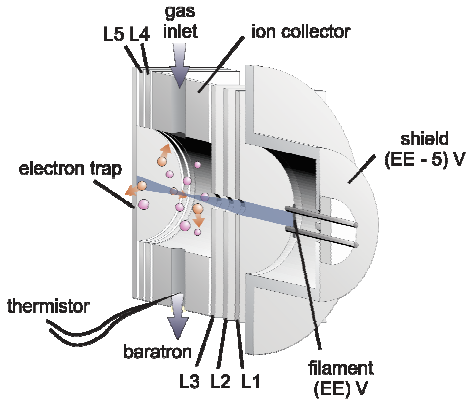

Fig. 1 Absolute total ionization cross-section cell. Cell length from filament block to trap is about 35 mm. Rare earth magnets are mounted at each end of the cell to magnetically constrain the electron beam. Lenses L1, L2 and L3 form an Einzel arrangement to focus the electron beam, and L4 and L5 have a small positive electrical bias.

Ionization efficiency curves are highly reproducible from run to run, even for experiments carried out several weeks apart. The results reported herein are the averages of between six to eight repeated determinations for each target gas, recorded over a period of about eight weeks. Before and after each data run, several measurements of the ionization efficiency curve for N2 were performed as both an independent assessment of the accuracy of the data and to ensure no organic sample was still present in the sample inlet line before the next sample was introduced. The "recommended" maximum cross-section for N_2 is given as 2.51 Å^2 (±5%) and the absolute values measured in this work always fall within the range of 2.50–2.53 Å^2.[31] The maximum instrumental uncertainty is estimated by summing the uncertainties associated with each measurement device to be ±4%. While the total uncertainty introduced by the electron energy bandwidth could in principle be larger than this, we have found that the estimated 4% uncertainty was always larger than the 95% confidence interval error for any of the sample data sets.

The fifteen cyclic compounds measured in this work can be divided into six aromatic and eight non-aromatic species, and are listed in Table 1. Geometrical structures for each species considered in this study are given in the ESI.†

Theoretical

TICSs were modelled using two different theories: the polarizability correlation model, σ_{POL}; and the binary-encounter Bethe model, σ_{BEB}. All required electronic structure parameters were calculated using the Gaussian 09 computational package.[32] The first step in all calculations involved optimisation of molecular geometry, followed by confirmation of these geometries as energetic minima through computation of harmonic vibrational frequencies. For calculations on isolated nucleobases, the carbon that is usually bonded with the sugar group was capped with a hydrogen atom. Further details of the calculations are given in the following sections.

Polarizability correlation

The polarizability correlation model considers the relationship between the maximum in the TICS and the molecular static electronic polarizability volume.[5,33] This empirical relation was first established by Lampe, Franklin and Field[34], and can be rationalized classically by the fact that both molecular properties have an analogous mathematical dependence on the electronic dipole matrix. Although the correlation only provides an estimate of the maximum TICS, the model has proven to be applicable universally across many different series of functional groups. Following a recent comprehensive review of this model based on data from some 65 different organic and halocarbon species that identified the optimum computational methodology,[5] static isotropic polarizability volumes were calculated using PBE0 density functional theory[35] with either a slightly modified version of the Sadlej[36] triple-ζ basis set (an uncontracted set of *d* orbitals with exponent ζ=0.1 has been added to each hydrogen atom) or the Z3Pol basis set of Benkova *et al*.[37]. The Z3Pol basis set is a reduced size and coefficient re-optimized variant of the parent Sadlej triple-ζ basis, and is targeted for improved computational efficiency when dealing with larger molecules with minimal loss of performance. A selected test set of species gives polarizabilities within 1% of those obtained using the parent basis set.

BEB model

The semi-classical binary-encounter Bethe (BEB) model of Kim and Rudd[38,39] is probably the simplest and most successful model for prediction of the total electron ionization efficiency as a function of incident electron kinetic energy. The model is an integrated approximation to the parent binary-encounter dipole (BED) differential cross-section model from the same authors,[38] which resulted from the melding of two different theories with a switching function: modified Mott theory for 'hard' or small-impact-parameter collisions; and Bethe cross-section theory for 'soft' or large-impact-parameter/glancing collisions. Calculations of the total BEB cross-section employed the procedure detailed in Bull *et al*.[5], which determines orbital parameters using P3 electron propogator theory[40] and performs a simple correction to the orbital kinetic energies obtained from the RHF-reference wave function. The aug-cc-pVTZ basis set.[41] was used for all necessary electronic structure calculations. For the nucleobases, the calculated vertical ionization potentials using this method were within a few percent of published values obtained from photoelectron spectroscopy.[42-44]

BEB theory is known to provide an upper-bound TICS because the model assumes that all energy in excess of the ionization potential is channelled into ionization.[5,30,45] In reality, direct ionization occurs in competition with a number of other processes, including neutral dissociation and autoionization. For example, the BEB model assuming single ionization typically overestimates the ionization cross-section by ~7% for main group (carbon substrate) organic species, and overestimates by ~40% for fluorocarbon species.[5,30]

Results and discussion

Aromatic and non-aromatic cyclic species

Experimentally measured TICS for the six aromatic, five non-aromatic heterocyclic, and four hydrocarbon species are shown in Fig. 2. The experimentally measured and modelled maxima in the TICS curves are listed in Table 1. The present manuscript focuses on model substituents of DNA and RNA, but for interested readers, the TICS for ethylacetamide, the smallest primary amide species that is electronically similar to the peptide bond found in protein chains, has also been investigated, and the data is given in the ESI.† Relative (not absolute) cross-sections for several of the species relevant to DNA and RNA have been measured previously: toluene by quadrupole mass spectrometry;[46] pyridine and cyclohexane by fourier transform mass spectrometry;[47,48] and tetrahydrofuran in a number of studies, including the time-of-flight co-incidence study of Fuss *et al*.[49], who made a series of measurements in 50 eV electron energy increments.

The previous TICS determination for toluene by Vacher *et al*.[46], normalized relative to argon, yielded a maximum of around 15 Å^2 at ~ 60 eV. In comparison, the 60 eV TICS measured in the present work is 16.91 Å^2, with a maximum at 90 eV of 18.13 Å^2. The TICS determined for pyridine by Jiao *et al*.[47], again normalized relative to argon, exhibits a maximum of 15 Å^2 at ~90 eV, albeit with large uncertainty of ~18%, which in this instance is in good agreement with the maximum determined in the present work of 14.20 Å^2 at 95 eV. A similar study by Jiao and Adams[48] on cyclohexane yielded a maximum TICS of 16.1 Å^2 at ~60 eV, which is larger than the 60 eV value of 15.23 Å^2 found in the present work, but in good agreement with the maximum TICS of 15.91 Å^2 at 85 eV. For tetrahydrofuran, the TICS determined in the present work at 50 eV is in excellent agreement with that of Fuss *et al*.[49], although their values at higher energies of 100 eV, 150 eV and 200 eV values are respectively ~11%, ~18% and ~7% larger than reported here. Clearly, the measurements of Fuss *et al*.[49] reach a maximum TICS at significantly higher electron energy than found in both the present work and in a recent

quadrupole mass spectrometry determination by Dampc *et al.*[50]. It should be noted at this point that many of these previous determinations have negelcted any discussion of ion detection

Table 1 Experimental and calculated maximum TICS for experimental species. Experimental errors are ±4%.

Species	σ^{max}_{expt} / Å²	σ^{max}_{PGL} / Å² [a]	σ^{max}_{BEB} / Å² [b]
benzene	15.05	15.14	14.61
toluene	18.13	18.14	17.58
pyridine	14.20	13.92	13.43
aniline	17.42	17.33	16.77
indene	22.69	22.92	21.63
pyrrole	12.00	11.95	12.15
piperidine	15.11	15.13	15.17
1,4,5,6-tetrahydropyrimidine	13.99	13.85	13.87
tetrahydrofuran	11.55	11.45	11.57
2-methyltetrahydrofuran	12.28	12.35	12.79
tetrahydrofurfuryl alcohol	15.32	15.07	15.60
cyclopentane	13.31	13.10	13.43
cyclohexane	15.91	15.69	16.04
cyclohexene	15.50	15.53	15.52
hexane	17.02	17.26	17.30
water	2.15	2.14	2.29
ethylacetamide	13.83	14.10	13.94

[a] Calculated assuming the linear correlation in Bull *et al.*[5].

[b] Tabulated values assume a correction factor detailed in the following text.

efficiency, and in particular the possibility of discrimination against fast ions formed in dissociative processes associated with high kinetic energy release (KER). High-KER ions become increasingly more abundant with increasing electron energy as multiple ionization and Coulomb explosion processes become more probable. An ion extraction and/or detection scheme that discriminates against detection of these ions will therefore lead to an underestimation of the ionization cross-section at higher electron energies, and will often also incorrectly record a maximum in the cross-section at a lower electron energy than the true maximum.

The results of the polarizability correlation model applied to the species under study are summarised in Fig. 3, with numerical values tabulated in Table 1. The experimental data all agree within experimental uncertainty (~±4%) with the linear correlation (gradient of 1.478) between maximum ionization cross-section and polarisability volume estabilished in Bull *et al.* 5. This further confirms the universal applicability of the correlation, in this case demonstrated for open-chain, cyclic and aromatic systems, and provides confidence in applying the model to the actual constituents of DNA and RNA in the following section.

The performance of the BEB model for the molecules studied is illustrated in Fig. 4. Fig. 4a demonstrates the linear correlation between the maximum cross-section predicted by the BEB model and that measured in our experiments. The cyclic non-aromatic species exhibit a linear correlation (gradient, m = 0.935) in common with the open-chain species considerd in Bull *et al.*[5], while the aromatic species have m = 0.895, corresponding to ~11% overestimation by theory when compared with experiment.

Despite the BEB model having being widely applied throughout the literature to a broad range of molecular systems, there has been no comprehensive review across a wide range of species of its capability to reproduce accurately the energy-dependence of the cross-section. This gap in the literature can be partly rationalized by the fact that the majority of TICS studies and applications of BEB theory consider only one or a few molecules, rather than investigating a series of related species such as in this study. Moreover, the level of theory employed in

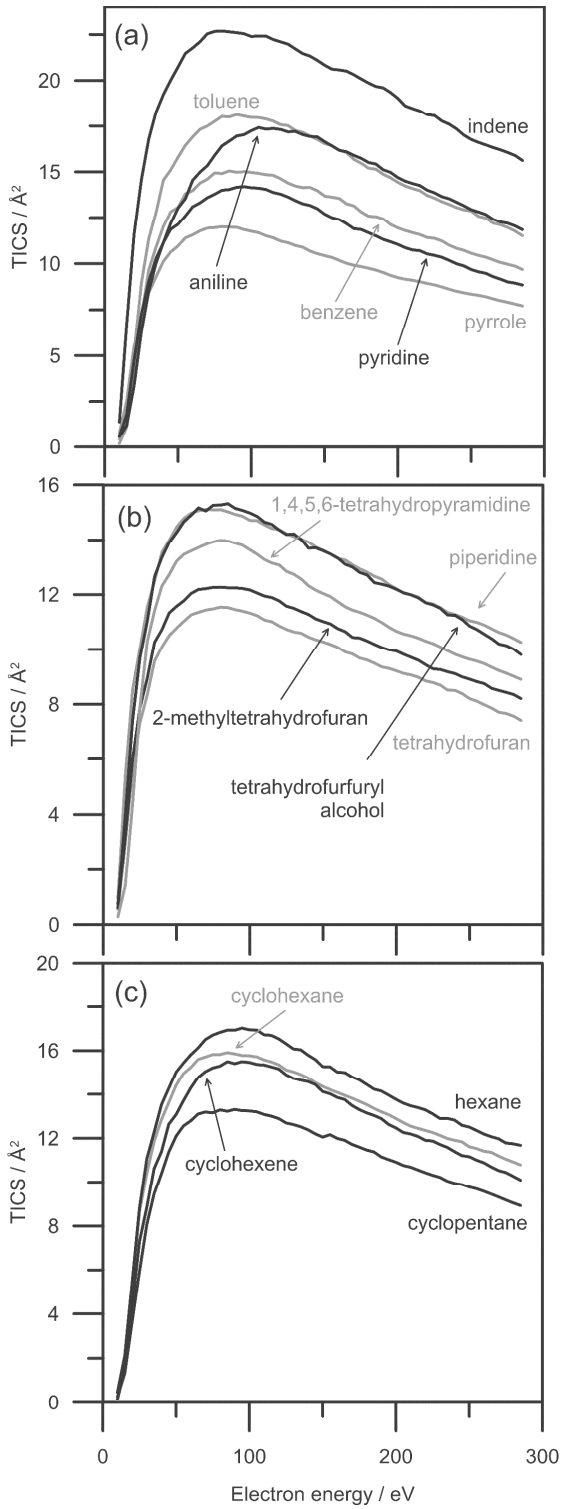

Fig. 2 Experimentally determined TICSs for fourteen of the species studied:
(a) aromatic species; (b) non-aromatic cyclic species; and (c) hydrocarbon species. Errors are ±4%.

carrying out the *ab initio* calculations of the required orbital

parameters is often not well-tested, and also varies greatly between different studies, making comparisons difficult.

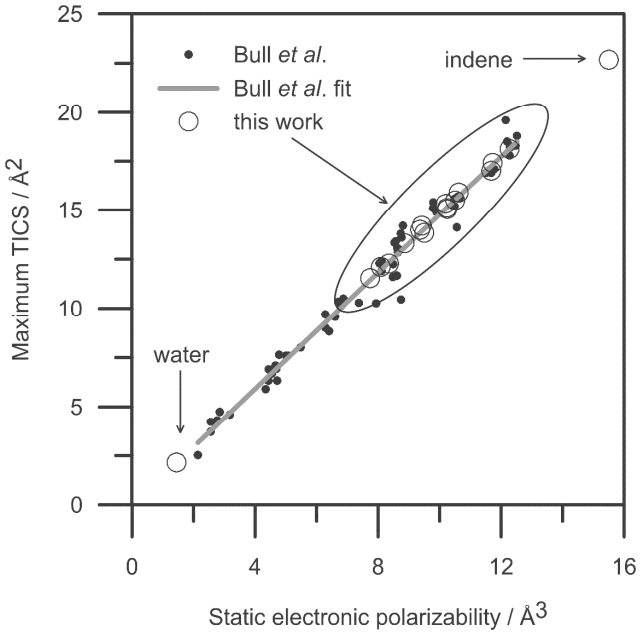

Fig. 3 Performance of the polarizability correlation model used to determine the maximum TICS. The data from the present study is superimposed on that from the previous study of Bull *et al.*[5], which considered 65 different open-chain organic species.

The following discussion now considers an empirical correction to the energy-dependent cross sections predicted by the BEB model to account for the systematic deviation of the model's energy dependence compared with experiment. First, the magnitude of the BEB cross-section at the maximum can be corrected by a factor equal to the slope, *m*, of the linear correlation plot shown in Fig. 5(a). The mismatch between the predicted and measured *shape* of the TICS curve can then be investigated by plotting the quantity $d\sigma = [\sigma(BEB) \times m - \sigma(expt)] / \sigma(expt)$ as a function of electron energy. The 'mismatch function', dσ, has been determined as an average over the aromatic and non-aromatic cyclic species considered in the present work, and is shown for these two cases in Fig. 5(b). It is noted that calculation of these averaged dσ function required careful calibration of the electron energy scale (assuming linear extrapolations due to the FWHM of the electron beam) and small corrections to the BEB electron energy scale because the P3/aug-cc-pVTZ level of theory, on average, always returns very slightly higher vertical ionization potentials than experiment. Bearing in mind that the disagreement between BEB predictions and experiment is largely due to the fact that the BEB model neglects the possibility of neutral dissociation processes, assuming that ionization is the sole fate of sufficiently energised molecules,[5,30,45] the different behaviour of the two dσ functions (taking into account *m*) implies systematically different neutral dissociation behaviour in aromatics and non-aromatics. Individual dσ values for each of the aromatic species are very similar, and each of the non-aromatic species are very similar, thus implying that molecules within each class exhibit similar neutral dissociation processes. The magnitude of dσ in the electron energy regime important for biological low-energy-electron-induced radiation damage indicates that the scaled BEB result overestimates the cross-section by greater than 10% in this region. The 'shape' correction should therefore be applied when employing the BEB model to predict ionization cross-sections for the actual constituents of DNA and RNA. For electron energies above ~75 eV, both dσ

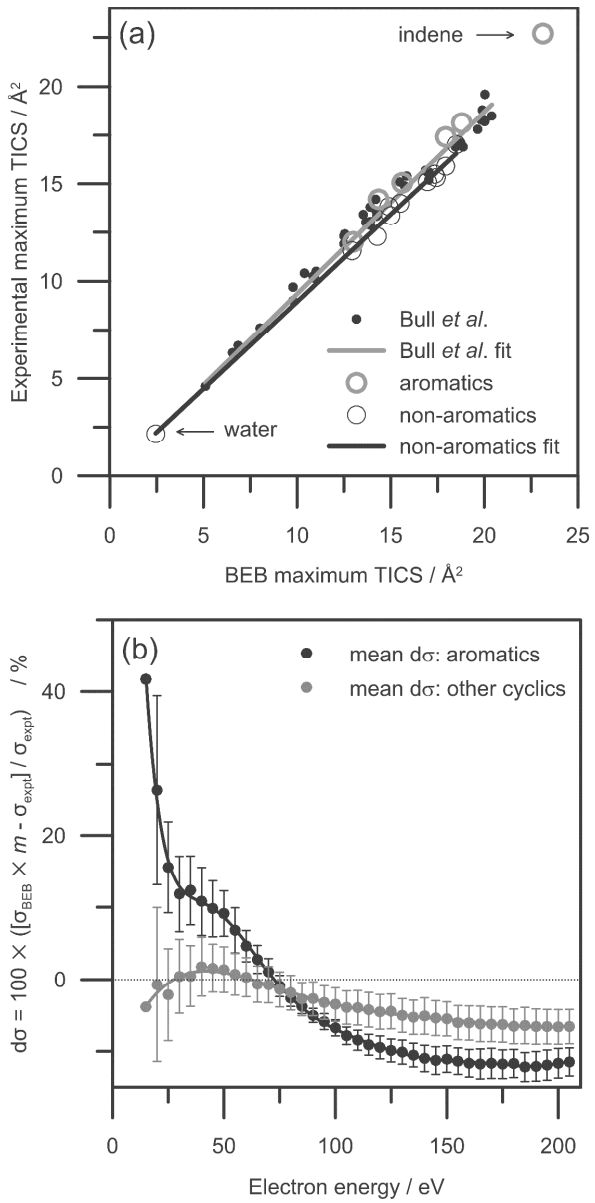

Fig. 4 BEB model calculations assuming single ionization for the species studied: (a) correlation between experimental and calculated TICS; (b) calculated difference function (in terms of percentage) between theory and experiment – see text for details.

functions in Fig. 4b become negative, inferring that maximum-scaled BEB theory systematically underestimates the TICS. This is partially a consequence of neglecting double ionization processes.[30] Finally, owing to the FWHM of the electron beam, the cross-sections in the following section for the actual DNA and RNA substituents should be most reliable in the so-called non-resonant (15 – 100+ eV) regieme. Note the error bars in Fig. 4b as an indication of uncertainities from BEB calculations in the following section.

DNA and RNA constituents

Now that the polarizability model and the BEB model have been thoroughly characterised on a series of similar molecules,

they can be used to calculate TICSs for the actual DNA and RNA nucleobases and sugars. As noted earlier, absolute TICS measurements for these species are difficult due to their low volatility. The results of the BEB calculations, with corrections applied according to the procedure outlined in the previous section, are shown in Fig. 5. Maximum cross-sections predicted by both models are given in Table 2. These isolated-molecule calculations have assumed the nucleobases to exhibit the same tautomers as in the cellular environment. The isolated nucleobases in the gas-phase, especially guanine, can exist in several different tautomeric forms depending on the method of evaporation employed.[43,51,52,53] However, calculations on the four lowest energy tautomers of guanine indicate that the maximum TICS varies between tautomers by less than 0.2 Å^2. Similarly, in the cellular environment, the deoxyribose and ribose sugar units of DNA and RNA sequences, respectively, exist as the five-membered ring deoxyribofuranose and ribofuranose structures. In aqueous solution, these sugars primarily exist in their six-membered deoxyribopyranose and ribopyranose isomers.[54] The linear chain (non-cyclic) isomer is a minor component (< 1% abundance) in most situations. Geometrical structures for all of these isomers are given in the ESI.† Several of the sugar isomers also exhibit two stereoisomeric forms, α and β, depending on the OH group attached to the carbon adjacent to the oxygen bridge, although exhibit identical TICS.

It can be seen from Table 2 that the maximum TICSs predicted by the polarizability and BEB models are in very close agreement for the nucleobases, but not for the four sugars. The BEB model predictions are ~11.4% and 9.4% larger than those of the polarizability correlation model for the ribose and deoxyribose structures, respectively. This overestimation indicates that a different parameter *m* (Fig. 4a) is required compared with the non-aromatic cyclic species considered experimentally in the present work, and may reflect different neutral dissociation behaviour due to the number of –OH functional groups. The BEB calculations given in Fig. 5b have thus been scaled to the maximum TICS predicted from the polarizability correlation model.

Also included in Table 2 are calculated maximum TICS assuming the polarizability correlation for five nucleosides (i.e., nucleobase + sugar). These polymers involve formation of a C-C bond between the nucleobase and sugar and loss of H_2O. Corresponding BEB model cross-sections are not presented, as accurate *ab initio* calculations of the BEB parameters for the nucleosides are too computationally demanding at this time.

If instead of using calculated vertical ionization potentials for the first few orbitals, experimental photoelectron vertical ionization potentials are employed, using data for adenine, cytosine and thymine from Trofimov *et al.*[42], guanine from Zaytseva *et al.*[43], uracil from Holland *et al.*[44], and α-deoxy-D-ribopyranose from Ghosh *et al.*[55], the calculated BEB TICS are essentially unchanged. This result is in agreement with aforementioned review,[5] which noted that all valence orbitals contribute additively to the TICS, and that BEB theory is insensitive to small perturbations in individual orbital parameters.

Comparison of the calculated TICSs for the nucleobases with the results of a number of other literature calculations[20-24] reveals considerable variation between different studies, with maximum TICS spanning about 5 Å^2 (or ~ 20%). The maximum TICSs calculated in this work are, overall, in close agreement with those of Mozejko and Sanche[21] and Huo *et al.*[22]. The Mozejko and Sanche[21] study employed BEB theory without the corrections employed in the present work, and the agreement between their data and this work is apparently due to a fortuitous cancelling between the effects of their low level of *ab initio* theory for the orbital parameters and the neglect of any

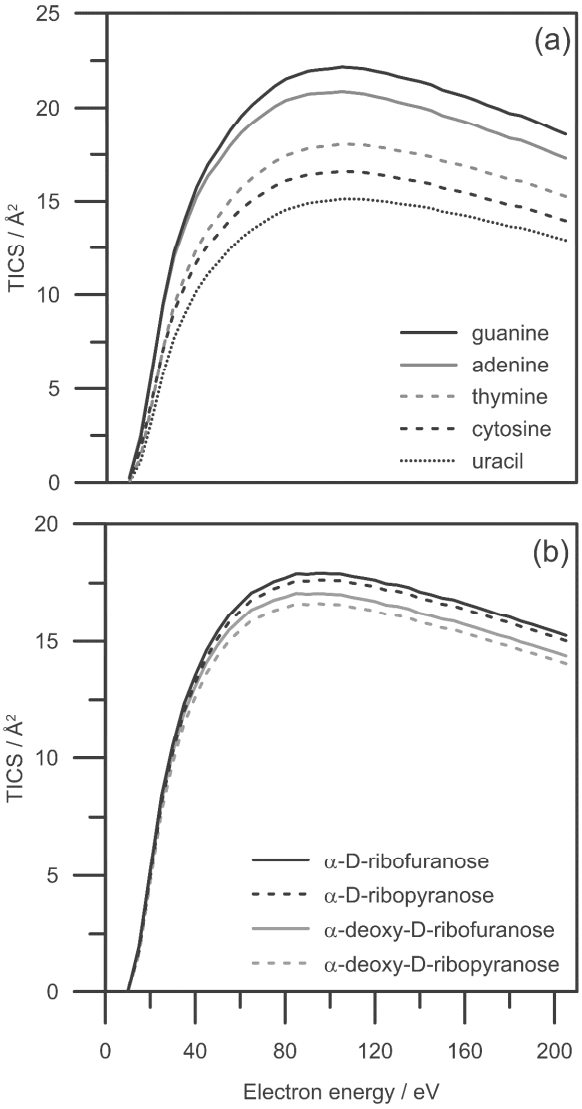

Fig. 5 Calculated TICS for the DNA and RNA constituents: (a) nucleobases; (b) sugar groups.

corrections for neutral dissociation. The study of Huo *et al.*[22] employed a modified variant of BED theory, called improved-BED (iBED),[56] which incorporates additional physics to account for some shortfalls of the parent BED theory. Finally, there have been two publications claiming to report experimentally determined absolute TICS of cytosine and uracil molecules through characterization of molecular beam densities, a procedure which is very difficult to perform reliably.[57,58] The claimed TICS maxima for cytosine of 10(1) Å^2 at 95 eV and uracil of 7.8 Å^2 at 78 eV are in considerable disagreement with the calculations from either this work or any other studies to date, and further experimental studies are clearly needed.

The effects of π-stacking and hydrogen-bonding on the ionization dynamics of nucleobases has been considered recently, both experimentally and theoretically, through the study of uracil (UU), cytosine (CC) and adenine and thymine dimers (AA, AT and TT).[53,59,60] The results indicate that these two non-covalent interactions can modify the vertical ionization potentials by 0.4 – 1.0 eV. To investigate the effect on the ionization cross-section,

polarizability correlation model calculations at the PBE0/Z3Pol level of theory were performed assuming the optimized geometries for the lowest energy hydrogen-bonded (HB1) and π-stacked (ST1) dimers of AA, AT and TT from Bravaya *et al.*[53]. The mean increase or decrease in the TICS relative to the separated monomers, i.e., [σ(dimer) – Σσ(monomers)]/Σσ(monomers), is reported in terms of a percentage in Table 3. These data predict that the π-stacking interaction of *one face* decreases the effective monomer TICS by ~7%, while the hydrogen-bonding interaction increases this quantity by ~4%.

Table 2 Calculated maximum TICS for DNA and RNA constituents, and several nucleosides.

Species[a]	σ^{max}_{Pol} / Å²	E^{max}_{BEB} / eV	σ^{max}_{BEB} / Å²
adenine	20.95	95	20.87
guanine	22.28	95	22.17
cytosine	16.88	105	16.59
thymine	18.11	105	18.05
uracil	15.23	105	15.12
α-D-ribofuranose	17.92	95	19.95
α-D-ribopyranose	17.63	95	19.65
α-deoxy-D-ribofuranose	17.06	85	18.66
α-deoxy-D-ribopyranose	16.62	95	18.21
adenosine[b]	37.44	-	-
guanosine[b]	38.96	-	-
cytidine[b]	33.26	-	-
thymidine[b]	34.37	-	-
uridine[b]	31.70	-	-

[a] All geometrical structures are given in the ESI.†
[b] Polarizabilities calculated using the Z3Pol basis set.

Effect of solvation

As a final consideration, the measured TICS for water is given in Fig. 6, and is in near perfect agreement with the corrected determination of Straub *et al.*[61,62], which has been "recommended" in several literature reviews.[31,62,63] The polarizability correlation produces a maximum TICS in excellent agreement with experiment – see Table 1.

It is well-established that in the gas phase, the DNA and RNA constituent first ionization potentials follow the ordering: phosphate < sugar < nucleobase; whereas in the cellular environment, the ordering becomes: nucleobase < sugar unit < phosphate.[64] Microhydration studies on the individual nucleobases reveal a decrease of ionization potentials by about 0.1 eV per water molecule, up to a maximum of about -0.9 eV (thymine) in the fully solvated nucleobase. Treating nucleobases in bulk solution also requires consideration of long-range electrostatic screening and solvent polarization.[65-67] According to the BEB model, these decreases in nucleobase vertical ionization potentials would imply a small increase in TICS. The situation is different for hydration of a nucleobase as part of a nucleoside or nucleotide (i.e., base + sugar + phosphate) unit; experiment and theory confirm that solvation provides an overall increase in vertical ionization potentials, by up to several electron-volts, in a fully hydrated polymeric DNA chain.[68-70] These effects will consequently result in a decrease in the constituents TICSs, especially at low electron energies. Quantification of solvation influences to cross-sections relevant to DNA and RNA damage are the subject of on-going studies.

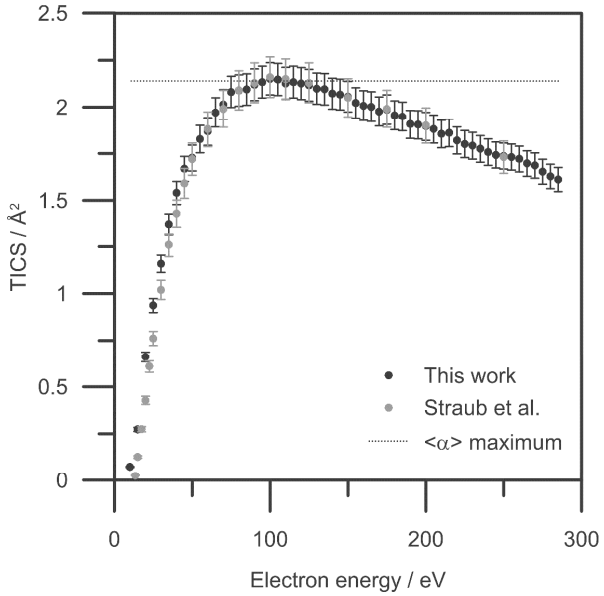

Fig. 6 TICS for water. Literature data are from Straub *et al.*[61].

Conclusions

TICS have been reported for a total of eighteen molecular species, many of which can be considered as basic prototypes for the cyclic moieties in DNA or RNA sequences. In accord with an earlier study, maximum experimental TICS were shown to exhibit excellent correlations with calculated values from two different models. By using the experimental and theoretical data to develop an empirical correction function for the BEB model, the energy-dependent TICSs for the nucleobases and sugar units were calculated. One caveat of the current experiment is that the FWHM of the electron source precludes accurate TICS determination for energies around ionization threshold of 9-10 eV, which is that most relevant in a DNA damage context several microseconds after cellular irradiation. On the shorter picosecond or so timescale, cross-sections at energies of tens and hundreds of electron-volts will contribute to a degree of the total damage.

That many of the species considered in this study have TICS that correlate in an obvious fashion with their structure, molecular mass, and bonding is not surprising. All indications to date infer that the TICS can be considered as a macroscopic or classical-like molecular property, which is relatively insensitive to detailed microscopic or quantum ionization dynamics. For this reason the simple polarizability correlation and functional group additivity models perform very satisfactorily in predicting the maximum in the TICS function.

Finally, it still remains to measure accurate absolute TICS for the actual DNA/RNA constituents in both the gas-phase and condensed phase environments. The primary challenge in the gas-phase is achieving a known, stable, and reproducible number density.

Acknowledgements

Funding is acknowledged from the ERC through Starting Independent Researcher Grant 200733 'ImageMS', from the EPSRC Programme Grant EP/G00224X/1, and from the Marie

Curie Initial Training Network 238671 'ICONIC', the latter two of which provided Postdoctoral Fellowships for JNB. The authors wish to thank the technical staff in Oxford's Department of Chemistry for fabrication of many of the instrumental components. Dr Roma Oakes, from the Open University (United Kingdom), is thanked for providing details of the Z3Pol basis set. A special thank you is extended to Prof. Peter Harland, University of Canterbury (New Zealand), for donating the electron ionization cell and many other pieces of scientific equipment utilized in the assembly of the instrument.

Notes and references

[a] *Chemistry Research Laboratory, Department of Chemistry, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, United Kingdom.*
[b] *Current address: Department of Chemistry, Durham University, South Road, Durham DH1 3LE, United Kingdom. E-Mail: james.bull@eigenket.org*
[c] *E-mail: claire.vallance@chem.ox.ac.uk*

† Electronic Supplementary Information (ESI) available: molecular structures for all of the species considered in this study, the plotted TICS for ethylacetamide, and numerical tabulations of all experimental total ionization cross-sections. See DOI: 10.1039/b000000x/

[1] F. H. Field and J. L. Franklin, Electron Impact Phenomena and the Properties of Gaseous Ions, Academic Press, Inc.: New York, 1957.

[2] T. D. Märk, Electron Impact Ionization, G. H. Dunn, Ed., Springer Verlag, 1985.

[3] J. E. Hudson, M. L. Hamilton, C. Vallance and P. W. Harland, *Phys. Chem. Chem. Phys.,* 2003, 5, 3162.

[4] J. N. Bull and P. W. Harland, *Int. J. Mass Spectrom.,* 2008, **273**, 53.

[5] J. N. Bull, P. W. Harland and C. Vallance, *J. Phys. Chem. A,* 2012, **116**, 767.

[6] B. Boudaïffa, P. Cloutier, D. Hunting, M. A. Huels and L. Sanche, *Science,* 2000, **287**, 1658.

[7] M. A. Huels, B. Boudaïffa, P. Cloutier, D. Hunting and L. Sanche, *J. Am. Chem. Soc.,* 2003, **125**, 4467.

[8] L. Sanche, *Eur. Phys. J. D,* 2005, **35**, 367.

[9] E. Alizadeh and L. Sanche, *Chem. Rev.,* 2012, **112**, 5578.

[10] T. Jahnke, H. Sann, T. Havermeier, K. Kreidi, C. Stuck, M. Meckel, M. Schöffler, N. Neumann, R. Wallauer, S. Voss, A. Czasch, O. Jagutzki, A. Malakzadeh, F. Afaneh, Th. Weber, H. Schmidt-Böcking and R. Dorner, *Nature Phys.,* 2010, **6**, 139.

[11] M. Mucke, M. Braune, S. Barth, M. Förstel, T. Lischke, V. Ulrich, T. Arion, U. Becker, A. Bradshaw and U. Hergenhahn, *Nature Phys.,* 2010, **6**, 143.

[12] S. M. Pimblott and J. A. LaVerne, *Radiat. Phys. Chem.,* 2007, **76**, 1244.

[13] I. Baccarelli, F. A. Gianturco, E. Scifoni, A. V. Solov'yov and E. Surdutovich, *Eur. Phys. J. D,* 2010, **60**, 1.

[14] B. Giese, J. Amaudrut, A.-K. Köhler, M. Spormann and S. Wessely, *Nature,* 2001, **412**, 318.

[15] C. L. Cleveland, R. N. Barnett, A. Bongiorno, J. Joseph, C. Liu, G. B. Schuster and U. Landman, *J. Am. Chem. Soc.*, 2007, **129**, 8408.

[16] S. Kanvah, J. Joseph, G. B. Schuster, R. N. Barnett, C. L. Cleveland and U. Landman, *Acc. Chem. Res.,* 2010, **43**, 280.

[17] S. Uehara, H. Nokjoo and D. T. Goodhead, *Rad. Res.*, 1999, **152**, 202.

[18] H. Nikjoo, D. Emfietzoglou, R. Watanabe and S. Uehara, *Rad. Phys. Chem.,* 2008, **77**, 1270.

[19] M. Müller, M. Durante, H. Stöcker, F. Merz and I. Bechmann, *Eur. Phys. J. D.,* 2010, **60**, 171.

[20] Ph. Bernhardt and H. G. Paretzke, *Int. J. Mass Spectrom.,* 2003, **223**, 599.

[21] P. Możejko and L. Sanche, *Radiat. Environ. Biophys.,* 2003, **42**, 201.

[22] W. H. Huo, C. E. Dateo and G. D. Fletcher, *Rad. Meas.,* 2006, **41**, 1202.

[23] A. Peudon, S. Edel and M. Terrissol, *Rad. Pro. Dos.,* 2006, **122**, 128.

[24] M. Vinodkumar, C. Limbachiya, M. Barot, M. Swadia and A. Barot, *Int. J. Mass Spectrom.,* 2013, **339/340**, 16.

[25] C. Champion, *J. Chem. Phys.,* 2013, 138, 184306.

[26] C. Vallance, S. A. Harris, J. E. Hudson and P. W. Harland, *J. Phys. B: At., Mol. Opt. Phys.,* 1997, **30**, 2465.

[27] D. A. Dahl, *Int. J. Mass. Spectrom.,* 2000, **200**, 3.

[28] T. D. Märk, *Contrib. Plasma Phys.,* 1982, **22**, 257.

[29] J. N. Bull, J. W. L. Lee and C. Vallance, *Phys. Chem. Chem. Phys.,* 2013, **15**, 13796.

[30] J. N. Bull, M. Bart, C. Vallance. and P. W. Harland, *Phys. Rev. A,* 2013, **88**, 062710.

[31] B. G. Lindsay and M. A. Mangan, Photon- and Electron-Interactions with Molecules: Ionization and Dissociation (Chapter 5: Ionization), Y. Itikawa, Ed., Springer, New York, 2003.

[32] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, "Gaussian, Inc.," 2009.

[33] P. W. Harland and C. Vallance, *Int. J. Mass Spectrom. Ion Proc.,* 1997, **171**, 173.

[34] F. W. Lampe, J. L. Franklin and F. H. Field, *J. Am. Chem. Soc.,* 1957, **79**, 6129.

[35] C. Adamo and C. Barone, *J. Chem. Phys.,* 1999, **110**, 6158.

[36] A. J. Sadlej, *Collect. Czech. Chem. Commun.,* 1988, **53**, 1995.

[37] Z. Benkova, A. J. Sadlej, R. E. Oakes and S. E. J. Bell, *J. Comp. Chem.,* 2004, 26, 145.

[38] Y.-K. Kim and M. E. Rudd, *Phys. Rev. A,* 1994, **50**, 3954.

[39] W. Hwang, Y.-K. Kim and M. E. Rudd, *J. Chem. Phys.,* 1996, **104**, 2956.

[40] J. V. Ortiz, *J. Chem. Phys.,* 1996, **104**, 7599.

[41] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.,* 1992, **96**, 6796.

[42] A. B. Trofimov, J. Schirmer, V. B. Kobychev, A. W. Potts, D. M. P. Holland and L. Karlsson, *J. Phys. B: At. Mol. Opt. Phys.,* 2006, **39**, 305.

[43] I. L. Zaytseva, A. B. Trofimov, J. Schirmer, O. Plekan, V. Feyer, R. Richter, M. Coreno and K. C. Prince, *J. Phys. Chem. A,* 2009, **113**, 15142.

[44] D. M. P. Holland, A. W. Potts, L. Karlsson, I. L. Zaytseva, A. B. Trofimov and J. Schirmer, *Chem. Phys.,* 2008, **353**, 47.

[45] H. Nishimura, W. M. Huo, M. A. Ali and Y.-K. Kim, *J. Chem. Phys.,* 1999, **110**, 3811.

[46] J. R. Vacher, F. Jorand, N. Blin-Simiand and S. Pasquiers, *Chem. Phys. Lett.,* 2007, **434**, 188.

[47] C. Q. Jiao, C. A. DeJoseph, Jr., R. Lee and A. Garscadden, *Int. J. Mass Spectrom.,* 2006, **257**, 34.

[48] C. Q. Jiao and S. F. Adams, *J. Phys. B: At. Mol. Opt. Phys.,* 2011, **44**, 175209.

[49] M. Fuss, A. Muñoz, J. C. Oller, F. Blanco, D. Almeida, P. Limão-Vieira, T. P. D. Do, M. J. Brunger and G. García, *Phys. Rev. A,* 2009, **80**, 052709.

[50] M. Dampc, E. Szymanska, B. Mielewska and M. Zubek, *J. Phys. B: At. Mol. Opt. Phys.,* 2011, **44**, 055206.

[51] J. Zhou, O. Kostko, C. Nicolas, X. Tang, L. Belau, M. S. deVries and M. Ahmed, *J. Phys. Chem. A,* 2009, **113**, 4829.

[52] O. Plekan, V. Feyer., R. Richter, M. Coreno, G. Vall-llosera, K. C. Prince, A. B. Trofimov, I. L. Zaytseva, T. E. Moskovskaya, E. V. Gromov and J. Schirmer, *J. Phys. Chem. A,* 2009, **113**, 9376.

[53] K. B. Bravaya, O. Kostko, M. Ahmed and A. I. Krylov, *Phys. Chem. Chem. Phys.,* 2010, **12**, 2292.

[54] L. P. Guler, Y.-Q. Yu and H. I. Kenttämaa, *J. Phys. Chem. A,* 2002, **106**, 6754.

[55] D. Ghosh, A. Golan, L. K. Takahashi, A. I. Krylov and M. Ahmed, *J. Phys. Chem. Lett.,* 2012, **3**, 97.

[56] W. M. Huo, *Phys. Rev. A,* 2001, **64**, 042719.

[57] I. I. Shafranyosh, M. I. Sukhoviya and M. I. Shafranyosh, *J. Phys. B: At. Mol. Opt. Phys.,* 2006, **39**, 4155.

[58] I. I. Shafranyosh and M. I. Sukhoviya, *J. Chem. Phys.,* 2012, **137**, 184303.

[59] A. A. Golubeva and A. I. Krylov, *Phys. Chem. Chem. Phys.,* 2009, **11**, 1303.

[60] O. Kostko, K. Bravaya, A. Krylov and M. Ahmed, *Phys. Chem. Chem. Phys.,* 2010, **12**, 2860.

[61] H. C. Straub, B. G. Lindsay, K. A. Smith and R. F. Stebbings, *J. Chem. Phys.,* 1998, **108**, 109.

[62] Y. Itikawa and N. Mason, *J. Phys. Chem. Ref. Data,* 2005, **34**, 1.

[63] G. G. Raju, Gaseous electronics: tables, atoms, and molecules, CRC Press, 2012.

[64] H. Fernando, G. A. Papadantonakis, N. S. Kim and P. R. LeBreton, *Proc. Natl. Acad. Sci. USA,* 1998, **95**, 5550.

[65] D. Ghosh, O. Isayev, L. V. Slipchenko and A. I. Krylov, *J. Phys. Chem. A,* 2011, **115**, 6028.

[66] L. Belau, K. R. Wilson, S. R. Leone and M. Ahmed, *J. Phys. Chem. A,* 2007, **111**, 7562.

[67] S. K. Kim, W. Lee and D. R. Herschbach, *J. Phys. Chem.,* 1996, **100**, 7933.

[68] E. Cauët, M. Valiev and J. H. Weare, *J. Phys. Chem. B,* 2010, **114**, 5886.

[69] E. Pluhařová, P. Jungwirth, S. E. Bradforth and P. Slavíček, *J. Phys. Chem. B,* 2011, **115**, 1294.

[70] P. Slavíček, B. Winter, M. Faubel, S. E. Bradforth and P. Jungwirth, *J. Am. Chem. Soc.,* 2009, **131**, 6460.