Protein Structures

DOI: 10.1002/anie.201204435

Structure and Charge-State Dependence of the Gas-Phase Ionization Energy of Proteins**

Alexandre Giuliani, Aleksandar R. Milosavljević, Konrad Hinsen, Francis Canon, Christophe Nicolas, Matthieu Réfrégiers, and Laurent Nahon*

The sequence-to-structure-to-function paradigm postulates that the sequence of a protein determines its folding, which in turn controls its activity.[1] Protein structure and folding have been studied intensively for nearly 50 years,[2] and substantial progress in this field has lead toward an atomistic depiction of the energy landscape for in vivo model proteins.[3] However, the understanding of the interplay between the complex three-dimensional arrangements of proteins and other fundamental physicochemical properties, such as the ionization energy, still remains a challenge.

Since the 1990s and the advent of modern ionization techniques, study of isolated biologically relevant molecules and assemblies has become possible.[4,5] Electrospray ionization (ESI) is a unique method to bring large and fragile biological species intact into the gas phase.[6] Mass spectrometry (MS) has since offered the unique ability of manipulating such ions in the gas phase. In particular, the potential of mass spectrometry to perform mass-selected spectroscopy on isolated biomolecules in the gas phase has been achieved.[7] Furthermore, MS provides additional control on the target charge state, which is not possible otherwise. For instance, Hirsch et al.[8] have reported spectroscopic studies on mass-selected clusters of ions, while Thissen et al.[9] took advantage of the storage of the ions to produce relaxed targets in their electronic ground states.

[*] Dr. A. Giuliani, Dr. F. Canon, Dr. C. Nicolas, Dr. M. Réfrégiers, Dr. L. Nahon
Synchrotron SOLEIL
L'Orme des Merisiers, Saint Aubin, 91192 Gif-sur-Yvette (France)
E-mail: alexandre.giuliani@synchrotron-soleil.fr
Dr. A. Giuliani
UAR1008 CEPIA, INRA
44316 Nantes (France)
Dr. F. Canon
INRA, UMR1324 Centre des Sciences du Goût et de l'Alimentation, Dijon (France)
Dr. K. Hinsen
Centre de Biophysique Moléculaire, UPR4301 CNRS, Orléans (France)
Dr. A. R. Milosavljević
Institute of Physics, University of Belgrade (Serbia)

[**] This work was supported by the Agence Nationale de la Recherche Scientifique, France, under the project no ANR-08-BLAN-0065 and partially by "Pavle Savic" project of bilateral scientific collaboration between Serbia and France. A.R.M. also acknowledges support by the Ministry of Education and Science of Republic of Serbia (Project No. 171020) and the COST Action MP1002 (Nano-IBCT)". The SOLEIL synchrotron radiation facility is acknowledged for providing beamtime under project 20100847.

Infrared gas-phase spectroscopy of protein polyprotonated cations generated by ESI is now routinely performed at free-electron laser facilities.[10,11] The coupling of a linear ion trap (Thermo Scientific LTQ XL) offering ESI capabilities with the vacuum–ultraviolet (VUV) DESIRS beamline[12,13] at the SOLEIL synchrotron radiation facility has been reported recently.[14–16] The association of a quadrupole ion trap and a VUV beamline at Bessy II has also been performed.[17,18] These experimental setups offer the unique ability to study the photoninduced dynamics and electronic properties of protonated protein ions by performing action spectroscopy of trapped ions isolated under vacuum.

Herein, we report on a systematic study on the dependence of the adiabatic ionization energy (IE) on the charge states for three proteins, namely cytochrome C, basic pancreatic trypsin inhibitor (BPTI), and ubiquitin. Our results show that this fundamental physical parameter is governed not only by the charge state but also by the tertiary structure of the proteins. We propose a simple model for the charge-dependent IE with direct predictive capabilities. Interestingly, a mean radius may be derived from IE measurements.

Selected electrosprayed ions of the proteins were stored in the ion trap and submitted to VUV irradiation for a controlled amount of time. For photon energies above the ionization threshold of the ion, photoionization could take place according to:

$$[M+n\,\mathrm{H}]^{n+} + h\nu \rightarrow [M+n\,\mathrm{H}]^{(n+1)+\bullet} + \mathrm{e}^- \quad (1)$$

where $[M+n\,\mathrm{H}]^{n+}$ is the selected precursor and $[M+n\,\mathrm{H}]^{(n+1)+\bullet}$ is the photoionization product. Figure 1 shows the mass spectra obtained for irradiation of charge selected state $+4$ of BPTI as a function of the photon energy. The $[M+4\,\mathrm{H}]^{4+}$ parent ion appears at m/z 1628.6. The photoionization process increases the charge of the ion without affecting its mass. Thus, the radical cation $[M+4\,\mathrm{H}]^{5+\bullet}$ appears at m/z 1302.9. The ionization threshold is obtained by fitting a Wanier-type threshold function[19] to the abundance of the ion formed by photoionization and normalized to both photon flux and the total ion current. For the present example in Figure 1, a value of 12.3 ± 0.1 eV has been determined for the photoionization threshold of $+4$ BPTI ion. The same procedure for determining the ionization threshold has been applied to each charge state that we could produce under ESI conditions for BPTI, cytochrome C and ubiquitin. Charge states ranging from 4 to 8 for BPTI, from 4 to 15 for cytochrome C and from 4 to 9 for ubiquitin, were selected. Figure 2a gathers the ionization energy as a function of the charge states measured for all three proteins.

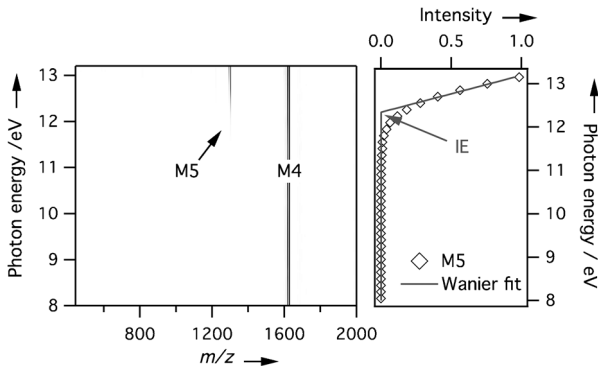

Figure 1. Left panel: plot of normalized mass spectra versus photon energy for the charge state +4 of BPTI. M4 = $[M+4H]^{4+}$; M5 = photoionization product $[M+4H]^{5+\bullet}$. Right panel: The normalized ion abundance for M5 (open symbols). The adiabatic ionization energy (IE) of 12.3 ± 0.1 eV, obtained from a Wanier law[19] fit (straight line) to the data is indicated by the arrow.

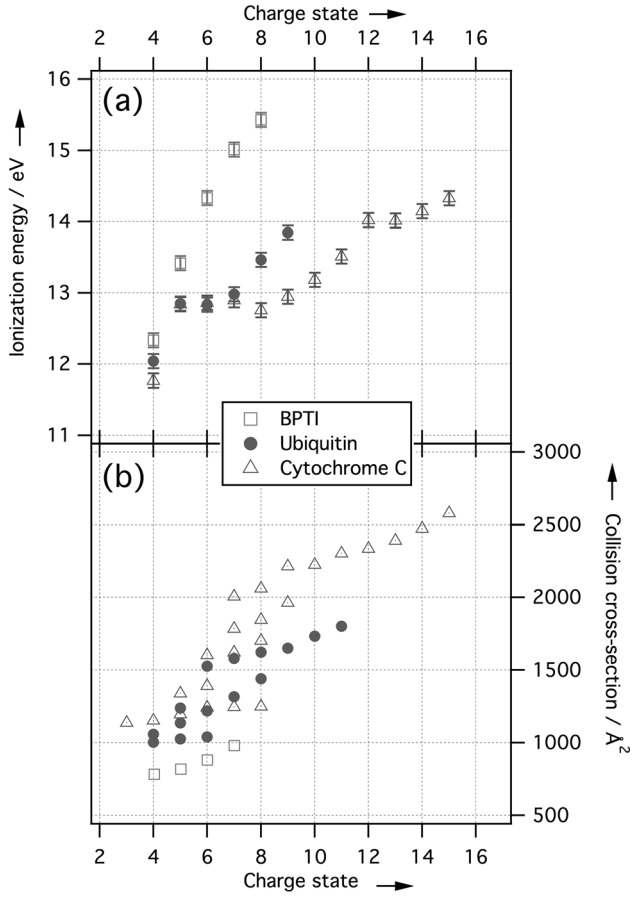

Figure 2. a) Measured ionization threshold as a function of the charge states for cytochrome C, BPTI, and ubiquitin. The error bars appear of similar width as the symbol. b) Literature collision cross-sections from ion mobility measurements for cytochrome C, BPTI, and ubiquitin.[22,26,27]

Budnik and co-workers have reported the first study on charge state-resolved electron impact measurement of the ionization energy for several peptides and small proteins (Substance P, rennin, insulin B-chain, and melitin).[20,21] They observed an increase of the ionization threshold for a given species as the charge was increased. It appeared that for polypeptides in the 1 to 3.5 kDa mass range, the ionization energy values fitted a straight line. In this model, the intercept of 9.8 eV represented the ionization energy of a hypothetical neutral peptide. The slope of 1.1 eV/charge accounted for the increasing attractive Coulombic interaction between the departing electron and the ion core as the charge increased, resulting in an enhancement of the ionization energy. We observe a similar quasi-linear trend over all charge states for BPTI (Figure 2a, □), but this is obviously not the case for cytochrome C (△) and ubiquitin (●). For both cytochrome C and ubiquitin, the ionization energy first increases for the lowest charge states, as would be expected from the linear model.[20,21] However, a plateau is reached surprisingly soon for intermediate charge states, where the ionization energies remain nearly constant. According to the linear model, we would expect an increase in the 2.2–3.3 eV range in Figure 1 between charge state 5 and 7–8, as observed for BPTI, whereas in that range the ionization energy of cytochrome C and ubiquitin is simply constant. For cytochrome C, starting at charge states +8, the ionization energy increases again monotonically but with a 0.24 eV/charge slope, which is much weaker than that observed for BPTI. Ubiquitin exhibits a similar trend in the +7 to +9 charge state range.

Ubiquitin, BPTI, and cytochrome C proteins have been extensively studied by ion mobility.[22–26] Literature collision cross-section data available[22,26,27] for ubiquitin, cytochrome C and BPTI are presented in Figure 2b. The major difference between BPTI on one hand and cytochrome C and ubiquitin on the other lies in their gas-phase tertiary structures. In contrast to the two other proteins, BPTI possesses three disulfide bridges, which provide it with a more rigid structure. This is confirmed by the ion mobility data, for which the collision cross-section of BPTI has been shown to be weakly sensitive to the charge state with a collision cross-section, increasing by about 25% in the 4 to 7 charge state range (see Figure 2b).

Molecular dynamics (MD) calculations of the radius of gyration R_g as a function of the charge state are presented in Figure 3 for BPTI and ubiquitin. The present MD simulations

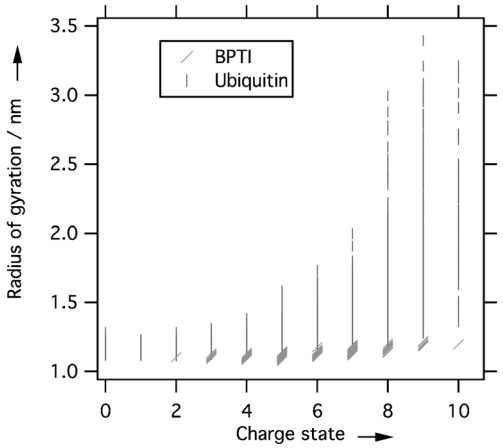

Figure 3. MD calculations of the radius of gyration for charge state ranging from 0 to +10 for BPTI and ubiquitin.

are consistent with the slight dependence of the structure of BPTI to protonation, and remain very compact over this charge state range. For ubiquitin, the MD simulations presented in Figure 3 indicate that the gas phase structure is much more sensitive to protonation: the distribution of the radius of gyration extends towards larger values. This calculated trend is in very good agreement with the outcome of ion mobility experiments (Figure 2b),[22] for which the transition from a compact structure towards an extended structure appears between +5 and +7 charge states. Such unfolding has been interpreted as an extension of the protein structure with the charge state to minimize Coulomb repulsion between the protonated sites of the protein. The situation is absolutely parallel for cytochrome C, for which at low charge states, (+5 and below), the collision cross-sections are almost independent of the charge state, reflecting a compact and nearly spherical gas phase structure. Starting at charge state +5, cytochrome C gradually unfolds. Up to charge state +9, compact structures may co-exist with more extended forms. For charge state +10 and higher, unfolding is gradual and monotonous, producing more extended arrangements.

The correspondence between the collision cross-section data and the measured adiabatic ionization energies as a function of the charge state (Figure 2) is noticeable and clearly reveals a correlation between the ionization energy and the gas-phase tertiary structure of the ions.

The initial modeling for the charge-state dependence of ionization energy proposed by Budnik and co-workers[20,21] was only taking into account the electron–ion Coulomb interaction. Our measurements obviously indicate that the gas-phase structure of the ion in a given charge state should also be considered, as it governs the positive charge distribution experienced by the outgoing electron. We therefore propose a correction to the modeling[20] that takes into account the structure dependence of the ionization energies. Indeed, the ionization energy $E_I(z)$ for charge state z may be described by the addition to the ionization energy of a neutral protein $E_I^0(z=0)$ of an extra energetic term to remove the photoelectron from a spherical protein bearing z positive electric charges on its surface:

$$E_I(z) = E_I{}^0 + z\,e^2/4\pi\varepsilon R_m(z) \qquad (2)$$

Where $R_m(z)$ is the mean radius of the protein cation, ε is the absolute permittivity of the medium, and e the electric charge. Equation (2) assumes a spherical shape for the protein and also a uniform charge distribution on the surface. In this model, when protons are added to a protein, the Coulomb energy is increased and leads to an enhancement of the ionization energy. But if the protein unfolds as a consequence of these additional protons, the potential experienced by the electrons is reduced. A linear trend for the ionization energy as a function of z is observed in Figure 2a for BPTI because its gas-phase structure is locked and remains almost constant (Figure 2b and Figure 3). For both ubiquitin and cytchrome C, the plateau observed in the 4 to 8 charge states range for the ionization energies in Figure 2a corresponds to transition regions where proteins start unfolding. As the ratio $z/R_m(z)$ does not change much in those z regions, the ionization energy does not vary. At higher charge states, the unfolding is less spectacular and the ionization energies increase again with z (Figure 2) but with lower slopes.

It appears from the Equation (2) that the knowledge of a protein's radius of gyration for a given charge state would allow the determination of its ionization energy, providing that the ionization energy for a hypothetical neutral protein ($z=0$) is known. To evaluate this hypothesis, the ionization energies were calculated from Equation (2) using atomic units, the permittivity of the vacuum and the entire population of computed radii of gyration for BPTI and ubiqutin. The results are presented in Figure 4 together with the present IE

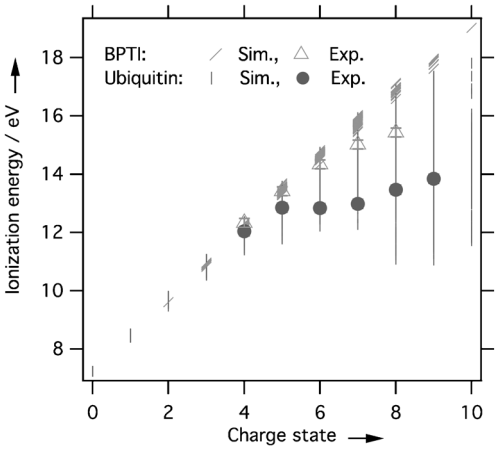

Figure 4. Prediction of the ionization energy based on R_g calculations present in Figure 3 and on Equation (2) (vertical and tilted bars) as compared to the measurements for BPTI and ubiquitin (triangle and full circle). The error bars appear smaller than the symbols on this figure. Sim. = simulation, Exp. = experiment.

measurements. The ionization energies for $z=0$ have been determined for both proteins by calculating the best E_I^0 matching the experimental data at $z=4$ and 5. Those z values were chosen because they correspond to the least-broad R_g distributions. We find for BPTI and ubiquitin, $E_I{}^0$ to be 7.1 eV and 7.3 eV, respectively. These values are much lower than the 9.8 eV reported by Budnik and co-workers for smaller polypeptides.[20] The estimated ionization energies for BPTI in Figure 4 agree well with the measurement up to charge state 6–7. For charge state 8, the calculation overestimates the IE, possibly because the radius of gyration might deviate from the mean radius of the spherical potential. Otherwise, the model agrees also very well with the measurements for ubiquitin. The plateau appearing at charge state +5 to +7 is also observed on the simulation in Figure 4, as well as the linear increase for the low charge.

Conversely, the relationship between ionization energy and the radius of the proteins may allow structural parameters from IE measurements to be derived. Indeed, in Figure 5, the model [Eq. (2)] has been used to derive R_m values from the IE measured for cytochrome C taking $E_I{}^0$ to be equal to the IE of tryptophane,[28] (7.44 ± 0.05 eV), which is the amino acid with the lowest ionization energy.[29] The comparison of the distribution of the radii as a function of the charge state

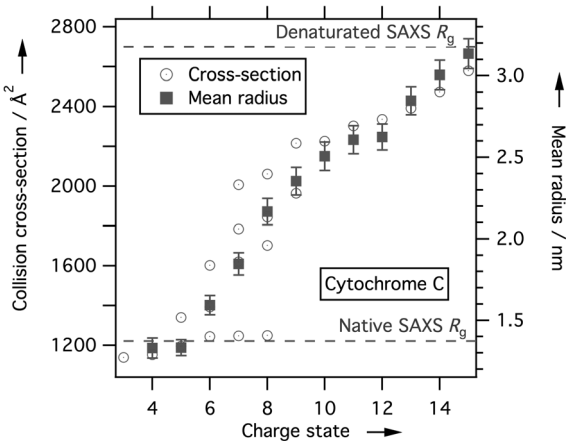

Figure 5. Comparison of the experimental ion mobility collision cross-section[22,27] and the radii for cytochrome C calculated using the model in Equation (2) as a function of the charge state. The dash lines indicate the radius of gyration measured by SAXS for the native[30] and the alcohol denaturated protein.[31]

with literature collision cross-sections over the $z = 4$ to 15 range shows a very good correlation between the two sets of data. In particular, the unfolding trend reported by ion mobility is well-reproduced in the data derived from ionization energy measurements.

Experimental radii of gyration for native cytochrome C in solution have been measured by small-angle X-ray scattering (SAXS).[30] The 13.9 Å experimental data is slightly larger than the 13.3 Å value obtained here for $z = 4$ and 5 (Figure 5). At the low-charge state, the removal of the solvent molecule in the gas phase and intramolecular interactions make the protein structure more compact than in solution, with side chains folding onto the surface. Similar observations have been made from ion mobility experiments.[22] Interestingly, SAXS studies of alcohol denaturated solution of cytochrome C[31] have reported R_g value of 31.7 Å, which is in agreement with the present finding of 31.4 ± 1 Å for $z = 15$. Recently, a SAXS study has been performed for cytochrome C ions stored in a digital ion trap,[32] but without specific m/z selection. In this work, a population of ions ranging from $z = 7$ to 10 was selected and the R_g measured to be 26.3 ± 0.5 Å, which is consistent with R_m data in Figure 5.

In summary, the present work demonstrates the correlation between the ionization energy of a protein and both its charge state and structure. In solution, the charge state distribution of a protein is partly pH-controlled. As a consequence of this study, the ionization energy of a protein is sensitive to each physical parameter that is able to affect either the pH or the structure of the protein. Indeed, denaturing agents, such as surfactants, salts, temperature, pressure, pH, or solvent, impinge on the protein structure and thus also affect the ionization energy of that protein. The ionization energy value for proteins has now to be considered as a quantity averaged over the population of conformers and charge states, and not as a single, defined, and fixed number. These results could open a new perspective in the field of radiation damages. Indeed, the ionization energy is a basic parameter either for direct or indirect damage caused by secondary particles, such as slow electrons[33] produced by the ionization process, which may be estimated using our model. Finally, it is shown that an ionization energy measurement might have the potential to become a new way to derive a mean radius for proteins and thus could lead to a new method for gas phase structural determination of molecular cations.

Experimental Section

The proteins (acquired from Sigma Aldrich) were dissolved in water/acetonitrile (70:30) at 5 μM and infused in the ESI source at 3 to 5 μL min^{-1}. In-source activation was minimized by reducing the cone voltage and the capillary temperature.

A photon shutter allowed precise control of irradiation time.[34] To ensure spectral purity and to avoid second-order light, a gas filter filled with argon was used to remove the higher harmonics of the DESIRS beamline undulator.[35] Monochromatic photons with 20 meV bandwidth were admitted inside the ion trap to interact with the target ions, as described in detail elsewhere.[14] After irradiation, the product-ion mass spectrum was recorded. This process was repeated over the photon energy range of interest to extract the different charge-dependent ionization energies.

Molecular dynamics simulations of BPTI and ubiquitin were performed using the Molecular Modeling Toolkit[36] and the Amber ff99SB force field.[37] The initial conformations were taken from PDB structures 1BPI (BPTI) and 1V80 (ubiquitin), removing hydrogen atoms and all non-protein atoms. For each simulation run, a charge state was chosen at random for each amino acid known to admit multiple charge states in solution. The corresponding hydrogen atoms were added to the structure, which was then subjected to molecular dynamics simulation starting at a low temperature (50 K) that was raised to 300 K over 0.5 ps of simulation time. After another 1.5 ps of equilibration, a trajectory of 50 ps was recorded for analysis, which consisted of calculating the radius of gyration averaged over the trajectory. The radius of gyration were obtained by calculation of the average mass weighted of $|r_i - r_{cm}|^2$. A total of 3000 simulation runs was performed for ubiquitin in order to sample the possible microscopic charge states. For BPTI, all 256 possible microscopic charge states were sampled.

Received: June 7, 2012
Published online: August 21, 2012

Keywords: ionization potentials · mass spectrometry · protein structures

[1] M. G. Rossmann, P. Argos, *Annu. Rev. Biochem.* **1981**, *50*, 497 – 532.
[2] C. B. Anfinsen, *Science* **1973**, *181*, 223 – 230.
[3] A. I. Bartlett, S. E. Radford, *Nat. Struct. Mol. Biol.* **2009**, *16*, 582 – 588.
[4] J. Fenn, M. Mann, C. Meng, S. Wong, C. Whitehouse, *Mass Spectrom. Rev.* **1990**, *9*, 37 – 70.
[5] M. Karas, R. Krüger, *Chem. Rev.* **2003**, *103*, 427 – 440.
[6] J. Benesch, C. Robinson, *Curr. Opin. Struct. Biol.* **2006**, *16*, 245 – 251.
[7] M. Duncan, *Int. J. Mass Spectrom.* **2000**, *200*, 545 – 569.
[8] K. Hirsch, J. Lau, P. Klar, A. Langenberg, J. Probst, J. Rittmann, M. Vogel, V. Zamudio-Bayer, T. Möller, B. Issendorff, *J. Phys. B* **2009**, *42*, 154029.
[9] R. Thissen, J. M. Bizau, C. Blancard, M. Coreno, C. Dehon, P. Franceschi, A. Giuliani, J. Lemaire, C. Nicolas, *Phys. Rev. Lett.* **2008**, *100*, 223001.

[10] G. Arteca, C. Reimann, O. Tapia, *Mass Spectrom. Rev.* **2001**, *20*, 402 – 422.
[11] J. Oomens, N. Polfer, T. D. Moore, L. van der Meer, A. G. Marshphyall, J. R. Eyler, G. Meijer, G. von Helden, *Phys. Chem. Chem. Phys.* **2005**, *7*, 1345 – 1348.
[12] http://www.synchrotron-soleil.fr/portal/page/portal/Recherche/LignesLumiere/DESIRS; accessed on the 22nd of April **2012**.
[13] L. Nahon, N. de Oliveira, G. Garcia, J. F. Gil, B. Pilette, O. Marcouille, B. Lagarde, F. Polack, *J. Synchrotron Radiat.* **2012**, *19*, 508 – 520.
[14] A. R. Milosavljević, C. Nicolas, J.-F. Gil, F. Canon, M. Réfrégiers, L. Nahon, A. Giuliani, *J. Synchrotron Radiat.* **2012**, *19*, 174 – 178.
[15] A. R. Milosavljević, A. Giuliani, C. Nicolas, J.-F. Gil, J. Lemaire, M. Réfrégiers, L. Nahon, *J. Phys. Conf. Ser.* **2010**, *257*, 012006.
[16] A. R. Milosavljević, C. Nicolas, J. Lemaire, C. Dehon, R. Thissen, J.-M. Bizau, M. Réfrégiers, L. Nahon, A. Giuliani, *Phys. Chem. Chem. Phys.* **2011**, *13*, 15432.
[17] S. Bari, O. Gonzalez-Magaña, G. Reitsma, J. Werner, S. Schippers, R. Hoekstra, T. Schlathölter, *J. Chem. Phys.* **2011**, *134*, 024314.
[18] O. González-Magaña, G. Reitsma, S. Bari, R. Hoekstra, T. Schlathölter, *Phys. Chem. Chem. Phys.* **2012**, *14*, 4351.
[19] G. H. Wannier, *Phys. Rev.* **1953**, *90*, 817 – 825.
[20] B. A. Budnik, Y. O. Tsybin, P. Håkansson, R. A. Zubarev, *J. Mass Spectrom.* **2002**, *37*, 1141 – 1144.
[21] B. A. Budnik, R. A. Zubarev, *Chem. Phys. Lett.* **2000**, *316*, 19 – 23.
[22] K. Shelimov, D. Clemmer, R. Hudgins, M. Jarrold, *J. Am. Chem. Soc.* **1997**, *119*, 2240 – 2248.
[23] A. A. Shvartsburg, F. Li, K. Tang, R. D. Smith, *Anal. Chem.* **2006**, *78*, 3304 – 3315.
[24] E. Badman, S. Myung, D. Clemmer, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1493 – 1497.
[25] D. Clemmer, R. Hudgins, M. Jarrold, *J. Am. Chem. Soc.* **1995**, *117*, 10141 – 10142.
[26] S. J. Valentine, A. E. Counterman, D. E. Clemmer, *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 954 – 961.
[27] http://www.indiana.edu/~clemmer/Research/cross%20section%20database/Proteins/protein_cs.htm, accessed April 22, 2012.
[28] F. Gaie-Levrel, G. A. Garcia, M. Schwell, L. Nahon, *Phys. Chem. Chem. Phys.* **2011**, *13*, 6993 – 7005.
[29] D. M. Close, *J. Phys. Chem. A* **2011**, *115*, 2900 – 2912.
[30] S. Akiyama, S. Takahashi, T. Kimura, K. Ishimori, I. Morishima, Y. Nishikawa, T. Fujisawa, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1329 – 1334.
[31] Y. O. Kamatari, T. Konno, M. Kataoka, K. Akasaka, *J. Mol. Biol.* **1996**, *259*, 512 – 523.
[32] B. J. McCullough, A. Entwistle, I. Konishi, S. Buffey, S. S. Hasnain, F. L. Brancia, J. G. Grossmann, S. J. Gaskelln, *Anal. Chem.* **2009**, *81*, 3392 – 3397.
[33] B. Boudaïffa, P. Cloutier, D. Hunting, M. Huels, L. Sanche, *Science* **2000**, *287*, 1658 – 1660.
[34] A. R. Milosavljević, C. Nicolas, J.-F. Gil, F. Canon, M. Réfrégiers, L. Nahon, A. Giuliani, *Nucl. Instrum. Methods Phys. Res. Sect. B* **2012**, *279*, 34 – 36.
[35] B. Mercier, M. Compin, C. Prevost, G. Bellec, R. Thissen, O. Dutuit, L. Nahon, *J. Vac. Sci. Technol.* **2000**, *18*, 2533 – 2541.
[36] K. Hinsen, *J. Comput. Chem.* **2000**, *21*, 79 – 85.
[37] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Genet.* **2006**, *65*, 712 – 725.

