

Molecular Dynamics

DOI: 10.1002/anie.201406910

Structure and Dynamics of Oligonucleotides in the Gas Phase**

Annalisa Arcella, Jens Dreyer, Emiliano Ippoliti, Ivan Ivani, Guillem Portella, Valérie Gabelica, Paolo Carloni, and Modesto Orozco*

Abstract: By combining ion-mobility mass spectrometry experiments with sub-millisecond classical and ab initio molecular dynamics we fully characterized, for the first time, the dynamic ensemble of a model nucleic acid in the gas phase under electrospray ionization conditions. The studied oligonucleotide unfolds upon vaporization, loses memory of the solution structure, and explores true gas-phase conformational space. Contrary to our original expectations, the oligonucleotide shows very rich dynamics in three different timescales (multi-picosecond, nanosecond, and sub-millisecond). The shorter timescale dynamics has a quantum mechanical nature and leads to changes in the covalent structure, whereas the other two are of classical origin. Overall, this study suggests that a re-evaluation on our view of the physics of nucleic acids upon vaporization is needed.

wenty years after the first mass spectrometry (MS) experiments showing that under mild electrospray ionization (ESI) conditions DNA can be detected as a duplex, the nature of

[*] Dr. A. Arcella, I. Ivani, Dr. G. Portella, Prof. Dr. M. Orozco Institute for Research in Biomedicine (IRB Barcelona) Joint BSC-CRG-IRB Program on Computational Biology Barcelona (Spain)

E-mail: modesto.orozco@irbbarcelona.org

Prof. Dr. M. Orozco

Department of Biochemistry and Molecular Biology University of Barcelona (Spain)

Dr. V. Gabelica

Mass Spectrometry Laboratory, Department of Chemistry University of Liège (Belgium)

Univ. Bordeaux, IECB, ARNA Laboratory, 33600 Pessac (France)

Inserm, U869, ARNA Laboratory, 33000 Bordeaux (France)

Dr. J. Dreyer, Dr. E. Ippoliti, Prof. Dr. P. Carloni Computational Biophysics

German Research School for Simulation Sciences and Institute for Advanced Simulation IAS-5, Computational Biomedicine Forschungszentrum Jülich (Germany)

Prof. Dr. P. Carloni

Institute for Neuroscience and Medicine INM-9 Computational Biomedicine, Forschungszentrum Jülich (Germany)

[**] This work is supported by the Spanish MINECO (BIO2012-32868), the Spanish National Institute of Bioinformatics (INB) and the

European Research Council (ERC). M.O. is an ICREA-Academia fellow and G.P. is a Sara Borell Fellow. Moreover, the authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN at the Jülich Supercomputing Centre



Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201406910.

such a duplex remains to be elucidated.^[1] Experiments and simulations have suggested that a medium-sized duplex maintains a large amount of stacking and hydrogen bonding for significant periods of time after vaporization, [2] and that when the oligonucleotide reaches the detector the general shape of the DNA duplex resembles that in solution.[3] However, little is known on the detailed structure and dynamics of the duplex under vacuum, and the rules governing the conformational changes induced by dehydration need to be elucidated to take advantage of current and future gasphase technology.^[4] We report herein a combined experimental (ESI-MS) and theoretical (classical and ab initio molecular dynamics) study to describe, for the first time, the equilibrium structure and dynamics of a small DNA hairpin duplex in the gas phase. The results provide a surprising picture of the changes occurring upon vaporization of DNA, and reveal that, contrary to general belief, the DNA duplex in the gas phase is far from being a rigid and chemically inert molecule.

Experiments and calculations were carried out with a short oligonucleotide (d(GCGAAGC)) which is known to fold in water (in the micro- to millisecond timescale) into a hairpin containing a d(GAA) triloop and a short B-DNA duplex stem.^[5] The short length of the oligonucleotide allowed us to increase our sampling capability, decreasing memory effects to a minimum. Ion-mobility data were recorded on a Synapt G1 HDMS mass spectrometer (Waters, Manchester, UK) with electrospray ionization in the negative mode using a 10 µm strand concentration, either in pure H₂O or in aqueous NH₄OAc (150 mm), and infused at 4 μL min⁻¹. Ion mobility was determined with the T-wave IMS cell (IMS = ion-mobility spectrometry) filled with N_2 at 0.532 mbar, and calibrated using helium drift tube IMS oligonucleotide data (see Table S1 and Figure S1 and S2 in the Supporting Information for details). Classical molecular dynamics (cMD) simulations were performed for all possible charge substates of the three major charged states of the hairpin (Q = -2, Q = -3 (the most populated state), and Q =-4 (the least populated state)). Starting gas-phase geometries were generated from randomly selected snapshots obtained from a 100 ns aqueous simulation^[5c] started from the NMR model with PDB code 1PQT^[6] (see Supporting Information; PDB = Protein DataBank). The phosphate titration required to achieve the desired charge state was performed using the localized charge method^[7] and all the reasonable charge substates (derived from the different ways to place the protons: five substates for Q = -2 and four for Q = -3 and Q = -4; these substates are labeled as LC#, where # labels the expected stability of the substate, where #=0 is the most stable). For each of these 13 systems, we generated 9 replicas taking solution snapshots as seeds for replica-exchange simulations (RExMD), [8] covering the temperature range



from 300-800 K, $^{[9]}$ for a total aggregated time of 0.30 milliseconds (see Supporting Information for details on RExMD calculations). To explore slow time-dependent conformational transitions for the Q = -4 state, we performed extended unbiased MD simulations for the lowest energy substates (LC0) at effective ESI-MS temperatures for a simulation time of 0.25 milliseconds (the longest ever published for a DNA molecule; see the Supporting Information). Ab initio quantum molecular dynamics were performed using Car-Parrinello molecular dynamics (CPMD)^[10] on snapshots selected from cMD simulations. Dispersioncorrected density functional BLYP-Grimme[11] with normconservative pseudopotentials of the Martins-Troullier type^[12] for core electrons along with the Kleinman-Bylander approach for the non-local part^[13] were used. A total of more than 0.5 ns-long ab initio quantum mechanical/molecular dynamics trajectories representative of around 0.5 milliseconds of cMD were collected (see Supporting Information for details on CPMD calculations).

The mass spectra of the oligonucleotide is dominated by three peaks corresponding in order of abundance to charge states $Q -3 \gg -2 > -4$ (Figure S2A in the Supporting Information). The average collision cross section (CCS) values are very dependent on the charge state: CCS = 472 \pm 19 Ų (Q=-4) \gg 375 \pm 13 Ų (Q=-3) > 324 \pm 7 Ų (Q=-2), but not on the bias voltage, or the buffer composition (see Figure S2B for typical arrival time distributions and Table S2 for the full list of values). This invariance suggests that the oligonucleotide has lost memory of the solution-phase conformation when entering in the IMS cell and it is exploring the equilibrium gas-phase landscape, rather than just a metastable conformation close to the solution structure. Comparing experimental CCS with estimates obtained from

the cMD average structure at room temperature in water (CCS = 360 Å²), we can conclude that vaporization collapses the structure for Q = -2, does not modify compactness for the Q = -3 state, and clearly enlarges the structure for Q = -4. Similar transitions to very extended structures have been noted before for the oligonucleotide dT_{10} , [14] which in contrast to our system is unable to form Watson–Crick base pairs. This means that the maintenance of CCS (for low charged states) does not necessarily imply maintenance of a canonical B-type helix.

RExMD simulations show that vaporization produces major structural changes even for the lower charge states (see Figure 1 and Figures S3 and S4). Strong unscreened electrostatic interactions slow down conformational mobility and generate conformers that interconvert slowly, even at high temperatures. Despite the structural diversity, the sampled structures show always a similar shape and compactness for a given charge state (see Figure 1). Rising the temperature leads to enlargement of the structure, which is small for lower charges states but is sizeable for Q = -4.

Taking as reference the IMS effective temperatures [7a,9] for each charge state, we conclude that vaporization compresses the structure for the Q=-2 state, in excellent agreement with experimental measurements (predicted CCS (RExMD) = 330 ± 10 Ų; experimental value 324 ± 7 Ų). The sampled Q=-2 structures are severely distorted, without memory of the solution conformation. The canonical H bonds are lost (only a residual 4% is detected), and are substituted by a myriad of unspecific interactions including stacking and non-canonical H bonds, the most prevalent (present in 90% of the snapshots) of which are the phosphate(anion)–phosphate(neutral) contacts (for one typical structure see Figure S5). For the Q=-3 state, CCS values

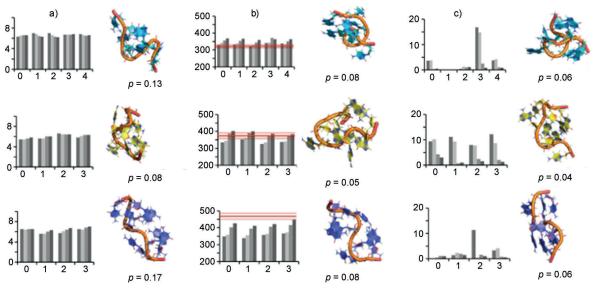


Figure 1. Summary of RExMD results in the gas phase for charge states Q = -2 (top), Q = -3 (middle), and Q = -4 (bottom) for different temperatures (gray bars in each plot from left to right: 300, 400, 500, and 600 K). Column (a) corresponds to root mean square deviation (RMSD in Å, y axis) plotted against LC# (x axis). Column (b) corresponds to collision cross sections (CCS given in Å²; y axis) plotted against different charge substates LC# (x axis). Column (c) corresponds to the percentage of hydrogen bonds (y axis) plotted against LC# (x axis). Red lines correspond to experimental values within the error of the CCS). At the right of each panel are representative structures of the most populated clusters for all charge states at the IMS effective temperature. The populations (p, relative to 1) of the clusters are shown.



derived from RExMD experiments also agree well with experimental estimates (cMD predicted CCS = $350 \pm 20 \text{ Å}^2$; experimental CCS value = $375 \pm 13 \text{ Å}^2$). This agreement confirms that for the major charge state, solution compactness is maintained, even though the sampled structures appear severely distorted with respect to the solution conformation (Figure 1 and Figures S3 and S4). The canonical H bonds are better preserved in the most abundant Q = -3ion than for Q = -2 (or Q = -4), but the percentage of stable canonical H bonds (10%) is still much smaller than that found for lengthier duplexes and again for non-canonical contacts. In particular, the phosphate-phosphate contacts become prevalent.

Finally, for the least populated high charge state (Q = -4) the type of intramolecular interactions detected by RExMD ensembles are similar to those found in the other charge states, but the structures appeared less densely packed (Figure 1 and Figure S4), as detected experimentally. The magnitude of such extension is not well reproduced by the RExMD simulations (at the same temperature: experimental $CCS = 472 \pm 19 \text{ Å}^2$, theoretical $CCS = 360 \pm 100 \text{ A}^2$

20 Å²), suggesting the existence of very slow conformational rearrangements (probably of entropic nature) which are not sampled in our RExMD simulations. Therefore, we extended equilibrium MD simulations to 0.25 milliseconds for the reference LC0 substate of Q = -4 oligo, reaching timescales not far from the experimental ones. This extremely long trajectory shows a very slow cyclic (period of 25–75 μs) conformational change (Figure 2). Such a conformational transition implies the exchange between compact (CCS < 380 Å², as in RExMD simulations) and extended structures $(CCS > 440 \pm 20 \text{ Å}^2)$, which are in much better agreement with the experiment. Formation of such an extended form is linked to a surprisingly modest structural modification related to the loss of residual stem interactions (Figure 2 and Figure S6).

The prevalence of phosphate-phosphate contacts found in cMD simulations raises the possibility of intramolecular proton transfers (PTs), something never considered before in similar theoretical studies. To analyze this, we performed an extensive ab initio molecular dynamics study using as seed structural snapshots from distant portions of the cMD trajectories which display at least a couple of hydrogen bonds involving acidic hydrogen atoms. The analysis of 500 ps of quantum trajectories (corresponding to more than 4.2 million QM-derived structures) shows the magnitude of the proton dynamics (Table 1, Figure 3, and Figures S7 and S8) that affects around 15% (Q = -3), 37% (Q = -2), and 79% (Q = -4) of the hydrogen-bond contacts studied.

PT is extremely fast (BLYP suggests sub-picosecond scale, test calculations (see the Supporting Information for details) using an exact HF exchange functional confirms the subpicosecond to picosecond scale) compared to the milliseconds required in solution. For the "active hydrogen bonds" (those

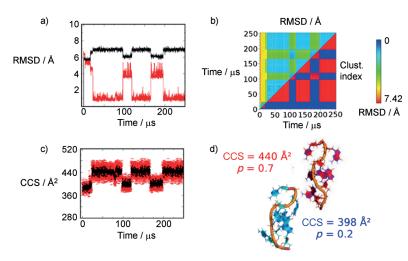


Figure 2. MD structural results for the hairpin at charge state Q = -4 considering the most stable LCO charge substate. Variation of structural descriptors a) RMSD and c) CCS along extended classical MD simulation (250 µs). In (a), the black line denotes the RMSD with respect to the native structure and the red line indicates the RMSD with respect to the gas-phase time-averaged structure. In (c), the black data points represent the CCS values and the red points denote the respective error bars for each point. b) Bidimensional RMSD/clustering plots and d) the representative structures corresponding to the two main clusters indicated in red (high CCS) and blue (low CCS) in the bidimensional plot are shown with their associated normalized probability, following the same color code.

Table 1: Summary of the CPMD analysis of selected H-bond interactions and associated proton transfer (PT) for ensembles of structures at different charge states.

Entry	Parameters	Q = -2	Q = -3	Q=-4
1	% contacts with PT	33	53	79
2	% full transfer	4	7	24
3	% broken contacts	21	_	5
4	% new contacts	17	_	5
5	% conf. changes	-	_	9
6	Av transfers/contact	6.7	6.0	12.0
7	Max transfers/contact	26	12	40
8	% time in acceptor	22	26	42
9	% time in share zone	9	4	7

Values reported correspond to H-bond contacts with phosphates as the donor. A PT has occurred if the proton jumps from the donor region (1.2 Å from the donor nucleus) to the acceptor region (1.2 Å from the acceptor nucleus) and stays bonded to it for at least 150 fs (~20 times the vibrational period of OH stretching). Full transfer is assumed if the proton remains bound for a longer time to the acceptor than to the donor site in the CPMD trajectory. Broken contacts are those where the original hydrogen bond contact is lost. The percentage of lost contacts leading to conformation changes are those where at the end of the CPMD trajectory donor-acceptor heteroatoms have significantly separated from the original conformation. For all the contacts where PT happens, the number of PT events were computed, showing the average and maximum number of events for each contact. The percentage of simulation time that on average protons stay close to the acceptor and in the "share zone" between donor and acceptor in the CPMD trajectory are given. Parameters 1-5 refer to the total number of contacts (followed by CPMD analysis), parameters 6-9 refer to those contacts showing PT transfer (i.e. 33% (Q=-2), 53% (Q=-3), and 79% (Q=-4) of the total number of contacts.

for which at least one PT happens) many PTs occur during individual trajectories. (On average 6 PTs occur for the most stable Q = -3 state and up to 12 PTs for the least stable highly

469



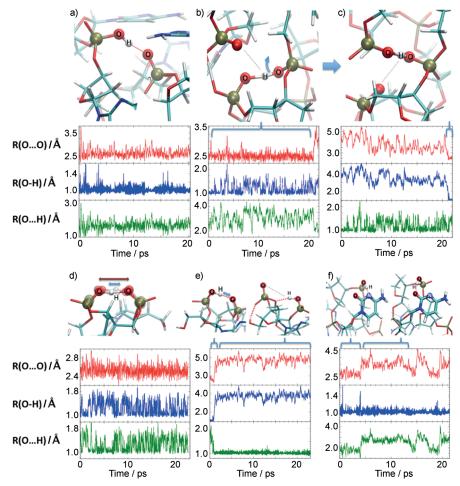


Figure 3. Hydrogen-bond configurations between phosphate groups. Examples of possible hydrogen-bond configurations between phosphate groups, or between phosphate groups and a nucleobase, with corresponding distance profiles describing potential PT processes. a) no PT; b, c) change of the hydrogen-bonding scheme between different phosphate groups without proton transfer; d) the proton is shared between donor and acceptor; e) PT between phosphate groups with subsequent disruption of the hydrogen bond; f) change of hydrogen-bonding scheme between phosphate and cytosine (acceptor moving from N3 to C=O). Atom colors: O=red; N=blue; H=white; P=gold.

charged Q = -4 state. In a few cases up to 40 PT transitions were found in a 20–25 ps trajectory for a single "active hydrogen bond").

Analysis of CPMD simulations highlights six possible proton-dynamic scenarios (Figure 3 and Figures S7 and S8): a) no PT (Figure 3a) and preservation of the H-bond scheme, b) H-bond scheme changes, without PT (Figure 3b, c), c) reversible PT occurs, with the proton being mostly bound to the original donor, d) frequent PT events with the proton being shared between donor and acceptor groups (the fingerprint of a low-barrier H bond is shown in Figure 3d), e) full PT occurs, with the proton being most of the time bound to the acceptor group, and f) as a consequence of donor-acceptor PT, there is a conformational change that implies non-negligible modifications in the geometry of the donor-acceptor groups (Figure 3e). Most of the PTs are reversible, and full transference (i.e. trajectories where the proton remains tightly bound to the acceptor) happens in 4.2% (Q = -2) to 24.1% (Q = -4) of the cases. Low-barrier

hydrogen bonds occur in 3.5-8.7% of the cases. The majority of PTs correspond to proton interchange between phosphates, but a few PTs were detected from a phosphate group to a nucleobase. In two of these cases the proton transferred from the phosphate tightly bound to the remained acceptor N3 group of guanine (Q = -4)or cytosine (Q = -3),respectively (Figure S8). It is worth noting that in most cases (see Table 1, Figure 3, and Figure S7), PT has little influence on the conformation of Hbonded pairs, and contact promiscuity (i.e. hydrogen bond contacts lost or created during the CPMD trajectories) is rare and never induces major structural rearrangements.

Herein, ESI-IMS-MS experiments are combined with classical and ab initio quantum MD simulations to provide a picture of unprecedented quality of the conformational space sampled by a small B-DNA duplex in the gas phase. Caution is needed as the oligonucleotide is smaller than usual, and accordingly the timescale of conformational transitions is probably faster than when a longer piece of DNA is considered. However, information derived is useful to gain a quantitative picture of what can happen to a piece of DNA when transferred for significant periods of time to the gas phase.

Our results suggest that overall, the equilibrium gas-phase ensemble is defined by a large number of microscopically different but macroscopi-

cally similar conformations, defining shapes of similar (Q =-3), smaller (Q = -2), or larger collision cross section (Q =-4) than that of the aqueous structure. The under vacuum conformational landscape is very rugged as a result of the stiffness of unscreened interactions, and conversion between close conformers is slow. However, despite this intrinsic rigidity, large dynamic changes responsible for the extension of the structure in the Q = -4 state are detected in the submillisecond timescale. Quite surprisingly, our extensive CPMD study reveals that proton transfer is common, even for low-charged states, and can involve phosphate-to-phosphate or less frequently phosphate-to-nucleobase processes. The phosphate-to-nucleobase case confirms the likelihood of transient base protonation in multiply charged oligonucleotide anions in charge states produced from native conditions. Our theoretical calculations thereby validate the oligonucleotide fragmentation mechanism proposed by Gross et al. for DNA, [15] and McLuckey and co-workers for RNA, [16] in which proton transfer from phosphate to nucleobase triggers fragmentation. Based on our simulations, the frequency of heteroatomic proton-transfer events is low, but it may be not negligible on the experimental timescale, and might even be more frequent at the higher internal energies typical of fragmentation experiments. We failed however, to find evidence supporting the idea that proton transfer between phosphates and sugars are prevalent.^[17] Although rare, we have detected cases where transfer of the proton to the acceptor group leads to separation of donor and acceptor groups, and accordingly can trigger some structural changes. This suggests that for a millisecond-scale process, PT-driven structural rearrangements may not be negligible and can increase even more the structural variability in the ensemble. Clearly, the previously held ideas that gas-phase structures are just a frozen copy of the solution conformation and that their chemical topology remains unaltered upon vaporization under ESI-MS conditions (although valid for cation-stabilized G-quadruplexes composed of ≥ 20 nucleotides)^[18] need to be revised, at least for small systems with potentially mobile protons. Future work to determine the minimum size and maximum time and temperature tolerable to preserve each type of nucleic acid structure is clearly needed.

Received: July 5, 2014 Revised: October 19, 2014

Published online: November 21, 2014

Keywords: ab initio calculations \cdot DNA \cdot gas phase \cdot mass spectrometry \cdot molecular dynamics

- a) M. J. Doktycz, S. Habibi-Goudarzi, S. A. McLuckey, *Anal. Chem.* 1994, 66, 3416–3422; b) B. Ganem, Y.-T. Li, J. D. Henion, *Tetrahedron Lett.* 1993, 34, 1445–1448; c) D. L. S. K. J. Light-Wahl, B. E. Winger, C. G. Edmonds, D. G. Camp II, B. D. Thrall, R. D. Smith, *J. Am. Chem. Soc.* 1993, 115, 803–804.
- [2] a) P. D. Schnier, J. S. Klassen, E. F. Strittmatter, E. R. Williams, J. Am. Chem. Soc. 1998, 120, 9605–9613; b) V. Gabelica, E. D. Pauw, J. Mass Spectrom. 2001, 36, 397–402; c) V. Gabelica, E. De Pauw, Int. J. Mass Spectrom. 2002, 219, 151–159.
- [3] a) J. Gidden, A. Ferzoco, E. S. Baker, M. T. Bowers, J. Am. Chem. Soc. 2004, 126, 15132-15140; b) A. Burmistrova, V.

- Gabelica, A.-S. Duwez, E. Pauw, *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1777 1786.
- [4] R. Neutze, G. Huldt, J. Hajdu, D. van der Spoel, *Radiat. Phys. Chem.* 2004, 71, 905–916.
- [5] a) A. Ansari, S. V. Kuznetsov, Y. Shen, Proc. Natl. Acad. Sci. USA 2001, 98, 7771-7776; b) H. Ma, C. Wan, A. Wu, A. H. Zewail, Proc. Natl. Acad. Sci. USA 2007, 104, 712-716; c) G. Portella, M. Orozco, Angew. Chem. Int. Ed. 2010, 49, 7673-7676; Angew. Chem. 2010, 122, 7839-7842.
- [6] P. Padrta, R. Stefl, L. Králík, L. Zídek, V. Sklenár, J. Biomol. NMR 2002, 24, 1-4.
- [7] a) A. Arcella, G. Portella, M. L. Ruiz, R. Eritja, M. Vilaseca, V. Gabelica, M. Orozco, J. Am. Chem. Soc. 2012, 134, 6596-6606;
 b) M. Rueda, S. G. Kalko, F. J. Luque, M. Orozco, J. Am. Chem. Soc. 2003, 125, 8007-8014;
 c) M. Rueda, F. J. Luque, M. Orozco, J. Am. Chem. Soc. 2006, 128, 3608-3619;
 d) M. Rueda, F. J. Luque, M. Orozco, J. Am. Chem. Soc. 2005, 127, 11690-11698.
- [8] Y. Sugita, Y. Okamoto, Chem. Phys. Lett. 1999, 314, 141-151.
- [9] D. Morsa, V. Gabelica, E. De Pauw, Anal. Chem. 2011, 83, 5775 5782.
- [10] R. Car, M. Parrinello, Phys. Rev. Lett. 1985, 55, 2471-2474.
- [11] a) A. D. Becke, *Phys. Rev. A* 1988, 38, 3098-3100; b) C. T. Lee,
 W. T. Yang, R. G. Parr, *Phys. Rev. B* 1988, 37, 785-789; c) S.
 Grimme, *J. Comput. Chem.* 2004, 25, 1463-1473.
- [12] N. Troullier, J. L. Martins, Phys. Rev. B 1991, 43, 1993-2006.
- [13] L. Kleinman, D. M. Bylander, *Phys. Rev. Lett.* **1982**, 48, 1425 1428.
- [14] C. S. Hoaglund, Y. Liu, A. D. Ellington, M. Pagel, D. E. Clemmer, J. Am. Chem. Soc. 1997, 119, 9051–9052.
- [15] a) Z. Wang, K. Wan, R. Ramanathan, J. Taylor, M. Gross, J. Am. Soc. Mass Spectrom. 1998, 9, 683–691; b) K. X. Wan, J. Gross, F. Hillenkamp, M. L. Gross, J. Am. Soc. Mass Spectrom. 2001, 12, 193–205; c) K. Wan, M. Gross, J. Am. Soc. Mass Spectrom. 2001, 12, 580–589.
- [16] T.-y. Huang, A. Kharlamova, J. Liu, S. A. McLuckey, J. Am. Soc. Mass Spectrom. 2008, 19, 1832–1840.
- [17] a) R. L. Cerny, M. L. Gross, L. Grotjahn, *Anal. Biochem.* 1986, 156, 424–435; b) J. P. Barry, P. Vouros, A. Van Schepdael, S.-J. Law, *J. Mass Spectrom.* 1995, 30, 993–1006; c) M. G. Bartlett, J. A. McCloskey, S. Manalili, R. H. Griffey, *J. Mass Spectrom.* 1996, 31, 1277–1283.
- [18] a) E. S. Baker, S. L. Bernstein, V. Gabelica, E. De Pauw, M. T. Bowers, *Int. J. Mass Spectrom.* 2006, 253, 225-237; b) V. Gabelica, E. S. Baker, M. P. Teulade-Fichou, E. De Pauw, M. T. Bowers, *J. Am. Chem. Soc.* 2007, 129, 895-904; c) R. Ferreira, A. Marchand, V. Gabelica, *Methods* 2012, 57, 56-63.