Secondary structure of Ac-Ala$_n$-LysH$^+$ polyalanine peptides (n=5,10,15) in vacuo: Helical or not?

M. Rossi[1], V. Blum[1], P. Kupser[1], G. von Helden[1], F. Bierau[1], K. Pagel[2,*], G. Meijer[1], and M. Scheffler[1]

[1] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195, Berlin, Germany and*
[2] *Freie Universität Berlin, Institut für Chemie und Biochemie, Takustr. 3, 14195 Berlin, Germany*[*]

The polyalanine-based peptide series Ac-Ala$_n$-LysH$^+$ (n=5-20) is a prime example that a secondary structure motif which is well-known from the solution phase (here: helices) can be formed *in vacuo*. We here revisit this conclusion for n=5,10,15, using density-functional theory (van der Waals corrected generalized gradient approximation), and gas-phase infrared vibrational spectroscopy. For the longer molecules (n=10,15) α-helical models provide good qualitative agreement (theory vs. experiment) already in the harmonic approximation. For n=5, the lowest energy conformer is not a simple helix, but competes closely with α-helical motifs at 300 K. Close agreement between infrared spectra from experiment and *ab initio* molecular dynamics (including anharmonic effects) supports our findings.

It is often said that the structure of peptides and proteins can not be understood without the action of a solvent, and this statement is certainly true for the full three-dimensional (tertiary) structure of proteins. However, their *secondary* structure level (helices, sheets, turns) is predominantly shaped by *intramolecular* interactions—most importantly, hydrogen bonds. For these interactions, benchmark experiment-theory comparisons under well-defined "clean-room" conditions *in vacuo* can furnish critical information towards a complete, predictive picture of peptide structure and dynamics. For example, first-principles approaches such as density-functional theory (DFT) with popular exchange-correlation functionals do not account for van der Waals (vdW) interactions. For peptide studies, precise experimental calibration points for different, actively developed theoretical remedies [1–5] would be extremely useful—*if the same structure as in the solution phase can be formed.*

In a seminal ion-mobility spectrometry study more than a decade ago, Hudgins, Ratner, and Jarrold (HRJ) [6] reported the formation of just such a secondary structure motif known from the solution phase (helical) *in vacuo* for a series of designed, charged polyalanine-based peptides Ac-Ala$_n$-LysH$^+$ (n=5-20). While much follow-up work has been done after the original HRJ study (e.g., Refs. [7–14]) the "helical" nature of the exact series Ac-Ala$_n$-LysH$^+$ is so far still established only indirectly by comparing ion-mobility cross-sections to results from force-field based molecular dynamics. A helical assignment for polyalanine is thus plausible and easily accepted, but different structural conclusions are not entirely ruled out. This fact is strikingly evidenced by additional ion-mobility experiments with microsolvation [9], which indicate that the secondary structure is not yet helical for $n < 8$. On the other hand, a spectroscopic study of Ac-Phe-Ala$_5$-LysH$^+$ [12] inferred structures with "helical" H-bond rings (α- or 3_{10}-helix-like) as the dominant conformers.

The key goal of the present work is to unambiguously verify the structure of the n=5, 10, and 15 members of the original HRJ series, both experimentally and theoretically. This is an important task, as *safely* knowing the correct structure is a key prerequisite for any further physical conclusions. Obviously, this is a broadly important statement for a wide part of physics and chemistry —surface science and catalysis, alloys and compounds, semiconductor properties etc.—but for a benchmark system such as the HRJ series, such safe knowledge is particularly crucial. On the theory side, we employ DFT in the PBE [15] generalized gradient approximation corrected for vdW interactions[5] with an accuracy that is critical for the success of our work. We thereby confirm the helical assignment for n=10 and 15, but for n=5, the lowest energy structure is indeed not a simple helix. Even for such a relatively short molecule, there is then an enormous structural variety to navigate, an effort which one must not shun (see below). We verify our findings against experimental infrared multiple photon dissociation (IRMPD) spectra of the vibrational modes in the 1000-2000 cm^{-1} region, which pertain to finite $T \approx$ 300 K. Here, harmonic free-energy calculations show that multiple conformers for n=5 (both helical and non-helical) should coexist, and are supported by calculated vibrational spectra (harmonic and anharmonic) in close agreement with experiment.

For the experiments, the peptides were synthesized by standard Fmoc chemistry. The experimental IR spectra were recorded using the Fourier transform ion cyclotron (FT-ICR) mass spectrometer [16] at the free-electron laser FELIX [17]. Ions were brought into the gas-phase by electrospray ionization (ESI) ($\sim$ 1 mg of peptide in 900 μl TFA/100 μl H$_2$O) and mass selected and trapped inside the ICR cell which is optically accessible. When the IR light is resonant with an IR active vibrational mode of the molecule, many photons can be absorbed, causing the dissociation of the ion (IRMPD). Mass spectra are recorded after 4s of IR irradiation. Monitoring the depletion of the parent ion signal and/or the fragmentation yield as a function of IR frequency leads to an IR spectrum.

All DFT+vdW calculations for this work were performed using the FHI-aims [18] program package for an accurate, all-electron description based on numeric atom-centered orbitals. "Tight" computational settings and accurate *tier 2* basis sets [18] were employed throughout. Harmonic vibrational frequencies, intensities and free energies were computed from finite differences. For Ac-Ala$_5$-LysH$^+$, we computed infrared intensities $I(\omega)$ *beyond* the harmonic approximation from *ab initio* molecular dynamics (AIMD) runs >20 ps (NVE ensemble, with a 300 K NVT equilibration), by calculating the Fourier transform of the dipole auto-correlation function [14, 19, 20] with a quantum corrector factor to the classical line shape [21, 22] proportional to ω^2 (see Ref. [19]). For a direct comparison to experiment, it is important to represent the "density of states" like nature of the measured spectra also in the calculated curves. All calculated spectra (harmonic and anharmonic) are therefore convoluted with a Gaussian broadening function with a variable width of 1% of the corresponding wave number, which accounts for the spectral width of the excitation laser. Further broadening mechanisms, e.g., due to the excitation process, are reflected in the experimental data.[23].

For the DFT+vdW calculations on Ac-Ala$_5$-LysH$^+$, we generate a large body of possible starting conformations using the empirical OPLS-AA force field (as given in Ref. [24] and references therein) in a series of basin hopping structure searches performed with the TINKER [25] package. Our particular choice of force field was not motivated by any other reason than that an input structure "generator" for DFT was needed. That said, the performance of OPLS-AA for gas-phase Alanine dipeptides and tetrapeptides was assessed rather favorably in earlier benchmark work.[26, 27] In the searches, specific constraints on one or more hydrogen bonds could be enforced. In total, we collected $O(10^5)$ nominally different conformers from (i) an unconstrained search, (ii) one hydrogen bond in the Ala$_5$ part constrained to remain α-helical, (iii) two hydrogen bonds in Ala$_5$ constrained to form a 3_{10}-helix, (iv) three hydrogen bonds in the Ala$_5$ part constrained to form a 2_7 helix, or (v) one hydrogen bond in the full peptide constrained to a π-helical form. As is well known [28], conformational energy differences between different types of secondary structure may vary strongly between different force fields and/or DFT. We reduce our reliance on the energy hierarchy provided by the force field by following up with full DFT+vdW relaxations for a wide range of conformers, 134 in total. This range includes the lowest ~0.3 eV for the unconstrained and α-helical searches, and the lowest ~0.15 eV for the 3_{10}-constrained search, as well as the lowest few π- and 2_7-helical candidates. Almost all π-helical geometries found in the force-field relaxed with DFT either into α or 3_{10} helices, and all relaxed 2_7 helices were higher in energy than our lowest-energy conformer by at least
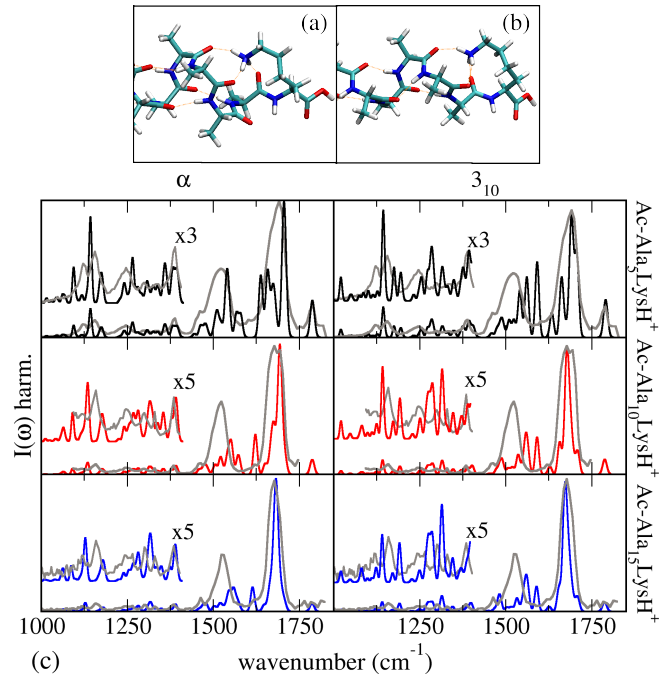
0.26 eV.

FIG. 1: Comparison between experimental [gray full lines in (c)] and calculated, broadened harmonic vibrational spectra [black, red, and blue solid lines in (c)] for α-helices of Ac-Ala$_5$-LysH$^+$(top), Ac-Ala$_{10}$-LysH$^+$ (middle) and Ac-Ala$_{15}$-LysH$^+$ (bottom). Also shown are the terminating H-bond networks used for α (a) and 3_{10} (b). All theoretical results for α- and 3_{10}-helices have been aligned to the free C-O peak at ~ 1790cm^{-1}.

Figure 1 shows the experimental IRMPD spectra for the three lengths of peptides studied (n=5,10,15) and calculated vibrational spectra in the harmonic approximation for two specific types of hydrogen bond networks: α-helical (left) and 3_{10}-helical (right). It is well known that both the choice of the density functional and the neglect of anharmonic effects will lead to characteristic frequency shifts between theoretical and experimental spectra. For a better visual comparison, all calculated spectra in the present work are therefore rigidly shifted to be aligned with the approximate location of the localized free C-O vibration at ~ 1790cm^{-1} in experiment for n=5. For example, this shift amounts to ≈20 cm^{-1} for α-helical conformers. All intensities were uniformly scaled to match the highest peak (Amide-I), but no further scaling factors (frequency or intensity) were employed. Given the limitations of the T=0 harmonic approximation when comparing to room-temperature experimental spectra, the agreement is rather reasonable for the α-helical conformers, while this is much less the case in terms of relative peak positions and fine structure for the 3_{10} helical conformers of n=10, 15. This observation correlates with calculated energy differences, where the 3_{10} helical

conformers are higher in energy than the α-helical ones by 0.41 eV (n=10) and 0.82 eV (n=15). In addition, an OPLS-AA based basin-hopping structure enumeration for n=10 did not reveal any non-α conformers within at least 0.15 eV. For n=15, the employed structure search procedure becomes prohibitive, although direct AIMD simulations show that α-helix conformations are structurally stable even at high T (500 K) for at least several tens of ps [29]. The available evidence thus points to at least predominantly α-helical secondary structure for n=10, 15. In any case, the observed disagreements for 3_{10} confirm the basic structure sensitivity of the measured IRMPD spectra.
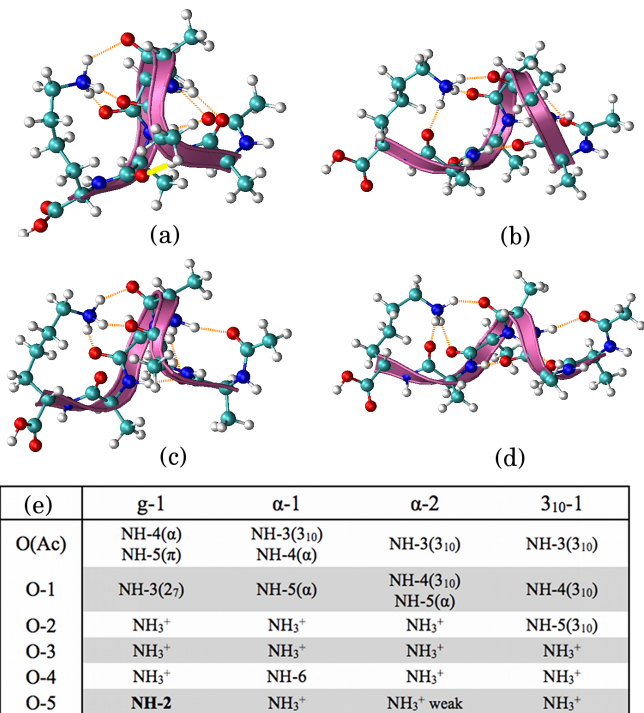
FIG. 2: Geometries of Ac-Ala$_5$-LysH$^+$: (a) g-1; (b) α-1; (c) α-2; (d) 3_{10}-1; and (e) H-bond network of the conformers: (C-)O and N-H groups are numbered starting from the N terminus and ending at the C terminus.

For Ac-Ala$_5$-LysH$^+$, the four lowest-energy conformers from our search and their H-bond networks are shown in Fig. 2. Three of these conformers (labeled α-1, α-2, 3_{10}-1) are "helical", in the sense that they contain two well-separated terminations with the appropriate α- or 3_{10}-like H-bond loops in their Ala$_5$ section. The lowest-energy conformer, labelled g-1, contains only one 2_7 like loop, H-bonds to the NH$_3^+$ end of the Lys side chain, and one H-bond that runs *against* the normal helix dipole, effectively short-circuiting the terminations. In fact, small structural differences in the Lys side chain lead to three nonequivalent conformers with the g-1 H-bond network, only one of which is shown here for simplicity.
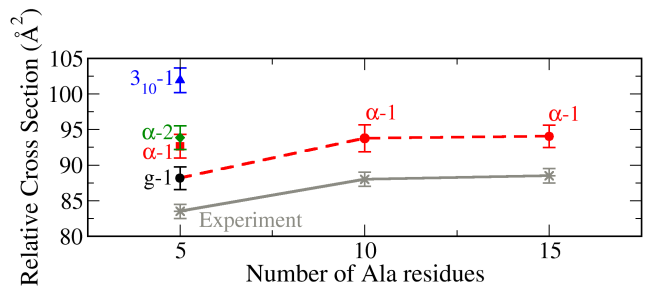
FIG. 3: Ac-Ala$_n$-LysH$^+$, $n = 5, 10, 15$: Calculated empirical ion-mobility cross sections and comparison with experiment (Ref. [6]). The solid and dashed lines serve only as a guide to the eye, and have no physical significance.

TABLE I: Energy differences of the four chosen Ac-Ala$_5$-LysH$^+$ conformers wrt. g-1: Pure DFT-PBE (no vdW), DFT-PBE+vdW (PES only), and harmonic free energy F at 300 K. All energies in eV.

	g-1	α-1	α-2	3_{10}-1
DFT-PBE	0.0	0.04	0.08	0.04
DFT-PBE+vdW	0.0	0.09	0.11	0.19
F(300 K)	0.0	0.01	0.06	0.17

The termination-connecting H-bond of the g-1 conformer also leads to an overall volume of the g-1 conformer that is somewhat smaller than of the α-1, α-2, or 3_{10}-1 conformers. This is quantified in Fig. 3 by way of computed empirical relative ion-mobility cross sections.[30] We show $\Omega = (\Omega_{measured} - 14.50n$ Å$^2)$ as a function of peptide chain length n, the same expression as used by HRJ [6]. Remarkably, the g-1 conformer for n=5 together with α-helical conformers for n=10 and 15 (dashed line) yields exactly the same qualitative behavior as the original data of HRJ. In contrast, our α-1 and α-2 conformers would yield a much shallower drop towards n=5, whereas the 3_{10}-1 conformer ends up too high.

In Table I, we summarize our computed energy hierarchy. In DFT-PBE+vdW, the g-1 conformer is more stable than its closest competitors by 0.1-0.2 eV. On this scale, vdW interactions are important, as seen by comparing to the pure DFT-PBE energy hierarchy (no vdW). On the other hand, finite temperature effects reduce the relative stability of g-1. The calculated harmonic free energy of g-1 and α-1 at 300 K is almost equal, and α-2 is only slightly ($\sim$ 60 meV) less stable; only 3_{10}-1 stays noticeably removed. The expected stability of at least three out of the four conformers is thus similar. We note in passing that the hierarchy for other DFT *functionals* (revPBE, or B3LYP at fixed geometry obtained using the PBE functional) is qualitatively similar, as long as a vdW correction is included.[31]
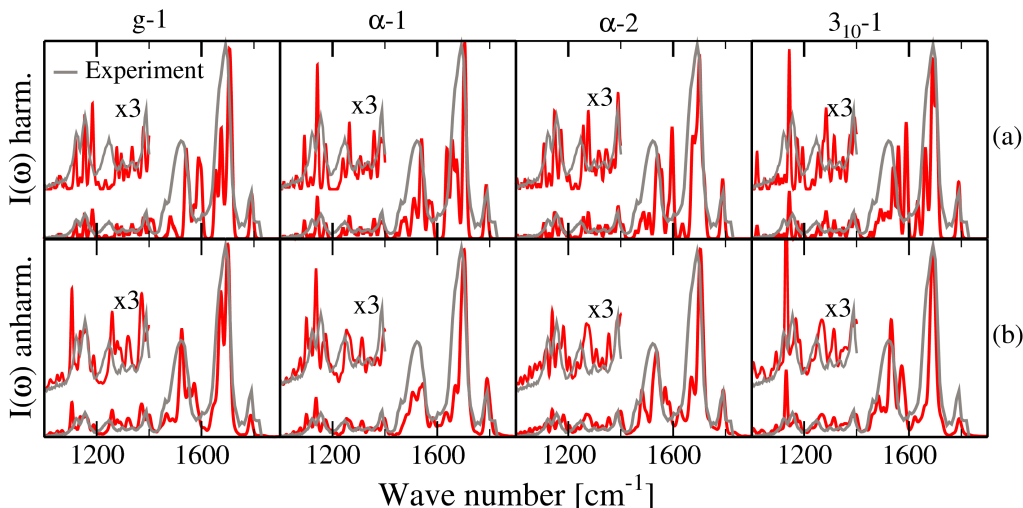
FIG. 4: Ac-Ala$_5$-LysH$^+$: (a) Theoretical harmonic vibrational spectra (red line) for the four chosen conformers compared with experiment (gray line); (b) same for *anharmonic* spectra from AIMD trajectories. All theoretical spectra are shifted rigidly so that the free CO peak ($\sim$ 1790cm^{-1}) is aligned.

Finally, Fig.4 shows computed vibrational spectra for all four conformers compared to experiment. Again, we align the localized free C-O vibration peak to experiment by a rigid shift and scale the intensities to match the height of the Amide-I peak. No further scaling factors are employed. In panel (a), spectra calculated in the harmonic approximation are shown. While there is an overall qualitative similarity of measured and computed spectra for *all* four conformers, it is also clear that none of them fit entirely—peak shifts and incorrect relative intensities (especially those obscuring the gap between Amide-I and -II) abound. This situation changes when spectra computed from *ab initio* molecular dynamics and the dipole-dipole autocorrelation are considered [panel (b)]. For g-1, α-1, and α-2, nearly perfect matches to experiment are obtained: The relative positions of Amide-I and Amide-II are almost exact, the interfering peaks in the gap decrease in intensity, and even the fine structure towards lower wave numbers is well reproduced. As one example, consider the experimental peak at $\approx$1250 cm^{-1} compared to the g-1 conformer. It coincides with a minimum of the harmonic spectrum, while a theoretical peak lies much closer in the anharmonic case. Consistent with the free energy, it is thus possible and plausible that all three conformers contribute to the experimentally observed signal from the ions, which were held in the ion trap at room temperature. In contrast, the features of the theoretical 3$_{10}$-1 conformer do not match quite as well. If 3$_{10}$-1 is present in the ion trap at all, then certainly with a much smaller fraction than g-1, α-1, and α-2.

In summary, we demonstrate a *quantitative* structure prediction for Ac-Ala$_n$-LysH$^+$ (n=5,10,15), with strong support by the good agreement between calculated and measured vibrational spectra. Our calculations provide a direct confirmation for the proposed α-helical nature of Ac-Ala$_n$-LysH$^+$ (n=10,15), while the *lowest energy* conformer of the "classic" HRJ series, n=5, is indeed *not* a simple helix. Importantly, finite-temperature free-energy effects still render α-helical n=5 conformers possible and even rather probable in an experimental molecular beam.

[*] University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK
[1] Q. Wu and W. Yang, J. Chem. Phys. **116**, 515 (2002).
[2] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[3] P. Jurecka, J. Czerny, P. Hobza, and D. R. Salahub, J. Comp. Chem. **28**, 555 (2006).
[4] S. Grimme, J. Antony, T. Schwabe, and C. Mück-Lichtenfeld, Org. Biomol. Chem. **5**, 741 (2007).
[5] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009), some of the present calculations employ an early version of this scheme with fixed C_6 coefficients, which yields the same geometries and energetics within a few meV as the final scheme of Tkatchenko and Scheffler.
[6] R. R. Hudgins, M. A. Ratner, and M. F. Jarrold, J. Am. Chem. Soc. **120**, 12974 (1998).
[7] R. R. Hudgins and M. F. Jarrold, J. Am. Chem. Soc. **121**, 3494 (1999).
[8] M. Kohtani, B. S. Kinnear, and M. F. Jarrold, J. Am. Chem. Soc. **122**, 12377 (2000).
[9] M. Kohtani and M. F. Jarrold, J. Am. Chem. Soc. **126**, 8454 (2004).
[10] M. Kohtani, M. F. Jarrold, S. Wee, and R. A. J. O'Hair, J. Phys. Chem. B **108**, 6093 (2004).
[11] J. A. Stearns, O. V. Boyarkin, and T. R. Rizzo, J. Am. Chem. Soc. **129**, 13820 (2007).
[12] J. A. Stearns, C. Seaiby, O. V. Boyarkin, and T. R. Rizzo, Phys. Chem. Chem. Phys. **11**, 125 (2009).

[13] T. D. Vaden, T. S. J. A. de Boer, J. P. Simons, L. C. Snoek, S. Suhai, and B. Paizs, J. Phys. Chem. A **112**, 4608 (2008).

[14] A. Cimas, T. D. Vaden, T. S. J. A. de Boer, L. C. Snoek, and M.-P. Gaigeot, J. Chem. Theory Comput. **5**, 1068 (2009).

[15] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[16] J. Valle, J. Eyler, J. Oomens, D. Moore, A. Meer, G. von Helden, G. Meijer, C. Hendrickson, A. Marshall, and G. Blakney, Rev. Sci. Instrum. **76**, 023103 (2005).

[17] D. Oepts, A. van der Meer, and P. van Amersfoort, Infrared Phys. Technol. **36**, 571 (1995).

[18] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comp. Phys. Comm. **180**, 2175 (2009).

[19] M.-P. Gaigeot, M. Martinez, and R. Vuilleumier, Mol. Phys. **105**, 19 (2007), and references therein.

[20] X. Li, D. T. Moore, and S. S. Iyengar, J. Chem. Phys. **128**, 184308 (2008).

[21] J. Borysow, M. Moraldi, and L. Frommhold, Molec. Phys. **56**, 913 (1985).

[22] R. Ramirez, T. Lopez-Ciudad, P. Kumar, and D. Marx, Molec. Phys. **121**, 3973 (2004).

[23] J. Oomens, B. G. Sartakov, G. Meijer, and G. von Helden, Int. J. Mass Spectrom. **254**, 1 (2006).

[24] M. L. P. Price, D. Ostrowsky, and W. L. Jorgensen, J. Comput. Chem. **22**, 1340 (2001).

[25] `http://dasher.wustl.edu/tinker/`.

[26] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, and R. A. Friesner, J. Am. Chem. Soc. **119**, 5908 (1997).

[27] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, J. Phys. Chem. B **105**, 6474 (2001).

[28] E. Penev, J. Ireta, and J.-E. Shea, Phys. Chem. B **112**, 6872 (2008).

[29] A. Tkatchenko, M. Rossi, V. Blum, and M. Scheffler (2009), in preparation.

[30] T. Wyttenbach, G. von Helden, J. J. Batka, Jr., D. Carlat, and M. T. Bowers, J. Am. Soc. Mass Spectrom. **8**, 275 (1997), ISSN 1044–0305.

[31] As a test, we verified that the fully relaxed α-1 geometry with the B3LYP functional is very similar to the fully relaxed PBE geometry.