Unraveling the Stability of Polypeptide Helices: Critical Role of van der Waals Interactions

Alexandre Tkatchenko[1], Mariana Rossi[1], Volker Blum[1], Joel Ireta[2], Matthias Scheffler[1]
[1] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany and*
[2] *Departamento de Química, Universidad Autónoma Metropolitana-Iztapalapa, A.P. 55-534, México D. F. 09340*

Folding and unfolding processes are important for the functional capability of polypeptides and proteins. In contrast to physiological environment (solvated or condensed phases), an *in vacuo* study provides well-defined "clean room" conditions to analyze the intra-molecular interactions that largely control the structure, stability, and folding/unfolding dynamics. Here we show that a proper consideration of van der Waals (vdW) dispersion forces in density-functional theory (DFT) is essential, and a recently developed DFT+vdW approach enables long time-scale *ab initio* molecular dynamics simulations at an accuracy close to "gold standard" quantum-chemical calculations. The results show that the inclusion of vdW interactions qualitatively changes the conformational landscape of alanine polypeptides, and greatly enhances the thermal stability of helical structures, in agreement with gas-phase experiments.

The helical motif is one of the ubiquitous conformations of aminoacids in protein structures, and helix formation is a fundamental step of the protein folding process [1–3]. Simulations of the helix formation and stability are typically carried out in solvent or condensed phases using classical, empirical "force fields". However, different force field parameterizations lead to reliable results only for a narrow class of systems and conditions. A truly bottom-up understanding of polypeptide structure and dynamics would greatly benefit from a first-principles quantum-mechanical treatment, as it becomes increasingly more feasible for large systems and long time scales. In particular, quantum-mechanical simulations *in vacuo* are invaluable for a quantitative understanding of the intra-molecular forces that largely control the structure, stability and (un)folding dynamics of polypeptides.

Recent progress in the experimental isolation and spectroscopies of gas-phase biological molecules has lead to increasingly refined vibrational spectra for the structure of peptides and proteins [4–6]. In fact, *in vacuo* proteins frequently preserve the secondary structure (helices and β-sheets) observed in solution, and recent results [7] have shown that even tertiary and quaternary structures can be transferred into the gas-phase. Joint experimental and *ab initio* theoretical studies can now successfully determine the geometries of small gas-phase peptides [8–10]. Nevertheless, many fundamental questions remain open, such as: (1) How stable is the folded polypeptide helix in comparison to other (meta)-stable structures? (2) How important are the different enthalpic and entropic contributions to the stability of helical conformations? Here we provide *quantitative* insight into the above two questions from first-principles for the fundamental case of polyalanine helices *in vacuo*. Our study reveals the crucial role that vdW interactions play for the helix stability and dynamics, illustrating how entropy is significantly altered as well. We choose polyalanine as a target for our study due to its high propensity to form helical structures [11], and its widespread use as a benchmark system for peptide stability in experiment [12–15] and in theory (see, e.g.,Refs. [16–24] and references therein). As detailed below, we find that vdW interactions stabilize native gas-phase helical forms of alanine polypeptide by a factor of two in relative energy over the fully extended structure on top of the widely used and established Perdew-Burke-Ernzerhof (PBE) [25] density functional. Only the inclusion of vdW forces can fully explain the remarkable stability of charge-capped polyalanine (Ac-Ala$_n$-LysH$^+$) helices up to a temperature of $\approx$ 750 K as recently observed in gas-phase experiments by Jarrold and coworkers [15]. In contrast, PBE simulations without vdW forces lead to unfolded structures at significantly lower temperatures, being in spurious agreement with solution-phase experiments [26].

Among the available first-principles methods, DFT plays a prominent role because it allows for affordable, prolonged simulations of the dynamics of condensed matter and molecular systems. DFT with present-day exchange-correlation functionals provides satisfactory accuracy for most covalently and/or hydrogen-bonded systems, and also treats electrostatic and polarization interactions accurately. However, the weak but important van der Waals dispersion force is missing from standard local, gradient-corrected, and hybrid functionals. Several encouraging strategies have been proposed to include vdW interactions in DFT [27]. As shown in Ref. [28], particularly accurate results can be obtained by a recently developed DFT+vdW approach. This method relies on augmenting the "plain" density functional energy by a pairwise sum of van der Waals $C_6[n]/R^6$ terms with a smooth cutoff towards short interatomic distances. In the DFT+vdW method the dispersion coefficients and vdW radii change at every step of the MD simulation in response to the electronic density $n(\mathbf{r})$, in contrast to fixed parameters used in other empirical vdW correction approaches [16, 29–32] or empirical force fields. This structure-sensitive definition of vdW parameters is crucial to achieve an accuracy of 5.5% in the C_6 coefficients for 1225 systems compared to accurate reference experimental data [28]. Furthermore, due to the direct de-

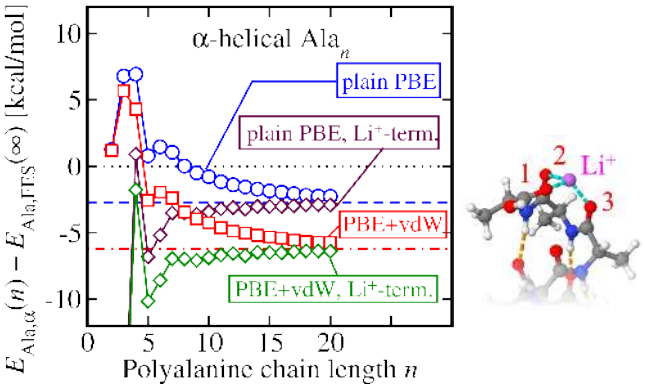

FIG. 1. Energy per added alanine peptide unit for idealized polyalanine α helices as a function of peptide length, referenced to infinite fully extended polyalanine structure. Circles and squares: neutral helices. Diamonds: Li^+-capped helices. A cartoon of the Li^+-capped C terminus structure is shown on the right. The labels 1, 2, 3 indicate the dangling hydrogen bonds saturated by the ionic termination.

pendence of vdW radii on the molecular electron density, the DFT+vdW method describes hydrogen and dispersion bonded systems on an equal footing, achieving mean absolute error (MAE) of 0.3 kcal/mol for a database of 22 quantum chemical "gold standard" CCSD(T) binding energies between organic molecules [28, 33]. Since the performance of DFT+vdW is significantly better than the so-called "chemical accuracy" of 1 kcal/mol (as also shown for short peptides in the supplementary material), fully quantitative simulations of the secondary structure of peptides become possible. The only empirical parameter in the DFT+vdW method is the scaling of vdW radii for every functional. This parameter has a weak dependence on the quantum-chemical reference database. All simulations in this work were performed using the FHI-aims all-electron code [34], which offers scalability with system size [35], and runs efficiently on massively parallel computing platforms.

We first turn to the question of the relative influence of the key qualitative contributions (hydrogen bonds, vdW interactions, and the chosen termination) on the stability of finite polypeptide helices. We focus here on neutral and charge-capped finite polyalanine chains Ala_n (n=5-20). The charge-capping at the C-terminus is necessary to compensate the peptide macrodipole, and obtain stable helical structures in gas-phase experiments [12–14]. To quantify each contribution, we monitor the energy to add one aminoacid residue to a finite polyalanine chain, $E_{\mathrm{Ala}}(n) = E_{\mathrm{tot}}(\mathrm{Ala}_n) - E_{\mathrm{tot}}(\mathrm{Ala}_{n-1})$, as a function of chain length n [20, 21] in Fig. 1. The Ala_n chain is frozen at the geometry of a hypothetical, infinite periodic α-helix [20, 21]. This allows us to study the stabilization towards the well-defined limit of a perfect infinite α-helical structure and compare to direct periodic calculations below. The terminating groups COOH and NH_2 as well as the Ala residues closest to the C terminus are

TABLE I. Stabilization energies (PBE, vdW, and PBE+vdW) per alanine residue for different fully-optimized infinite helix structures (in kcal/mol). The stabilization energies are given with respect to the fully extended structure (FES) of the polypeptide (except "vdW" row). All values are rounded to the first digit.

	FES	α-helix	π-helix	3_{10}-helix
PBE	0	-2.8	-2.2	-2.2
vdW	-3.8	-7.2	-7.4	-5.7
PBE+vdW	0	-6.3	-5.8	-4.2

relaxed for n=5, and then kept at that structure for larger n. For charge-terminated peptides, geometry relaxation does not change the results appreciably, but for neutral peptides short α-helices are directly unstable [10, 24], and would prevent us from comparing charged and neutral systems. The circular symbols in Fig. 1 show the development of $E_{\mathrm{Ala},\alpha}(n)$ from PBE calculation for neutral helices COOH-Ala_n-NH_2 without accounting for vdW effects. There is a significant *cooperative* effect between hydrogen bonds [19, 21, 36], which increases rather slowly with chain length n towards the limit of an infinite periodic chain (dashed line). Including the vdW contribution (squares) more than doubles the α-helical stability to the value reported in Table I. Comparing the curves for plain PBE and PBE+vdW reveals that the effect of vdW interactions is much shorter-ranged than the cooperative effect of H-bonds – the difference between both curves is constant for $n \geq 5$, where the first α-helix hydrogen bond is formed. The simple reason is that for hydrogen bond cooperativity, chains of hydrogen bonds are needed, whose dipolar interactions (including a possible density polarization) strengthen one another. Obviously, no such chains exist yet for short neutral α-helices, which are thus effectively destabilized due to the absence of H-bond cooperativity. By adding a charge near the C terminus (here we use Li^+, but the same conclusions hold for a simple positive point charge), the α-helix is significantly stabilized. This result agrees with the experimental observation that the helix formation requires charge-capping the Ala_n peptides near the C terminus [12–14].

We now turn to the role of vdW interactions for the *relative* stability of periodic polyalanine α, π, and 3_{10} helices. The infinite periodic model eliminates the macrodipole and has been proposed as a fundamental benchmark system beyond small peptides [22]. Table I shows the stabilization energies per residue for fully-optimized infinite α, π, and 3_{10} alanine helices compared to the fully extended structure (FES). The relative energetic stability of helical structures increases dramatically with PBE+vdW in comparison to plain PBE by 126%, 166%, and 90% for the α, π, and 3_{10} helices, respectively. Furthermore, the energy difference between the α and 3_{10} helices is increased from -0.6 kcal/mol per residue in PBE to -2.1 kcal/mol in PBE+vdW. Previous tight-binding calculations arrived at the same conclusion for polyala-

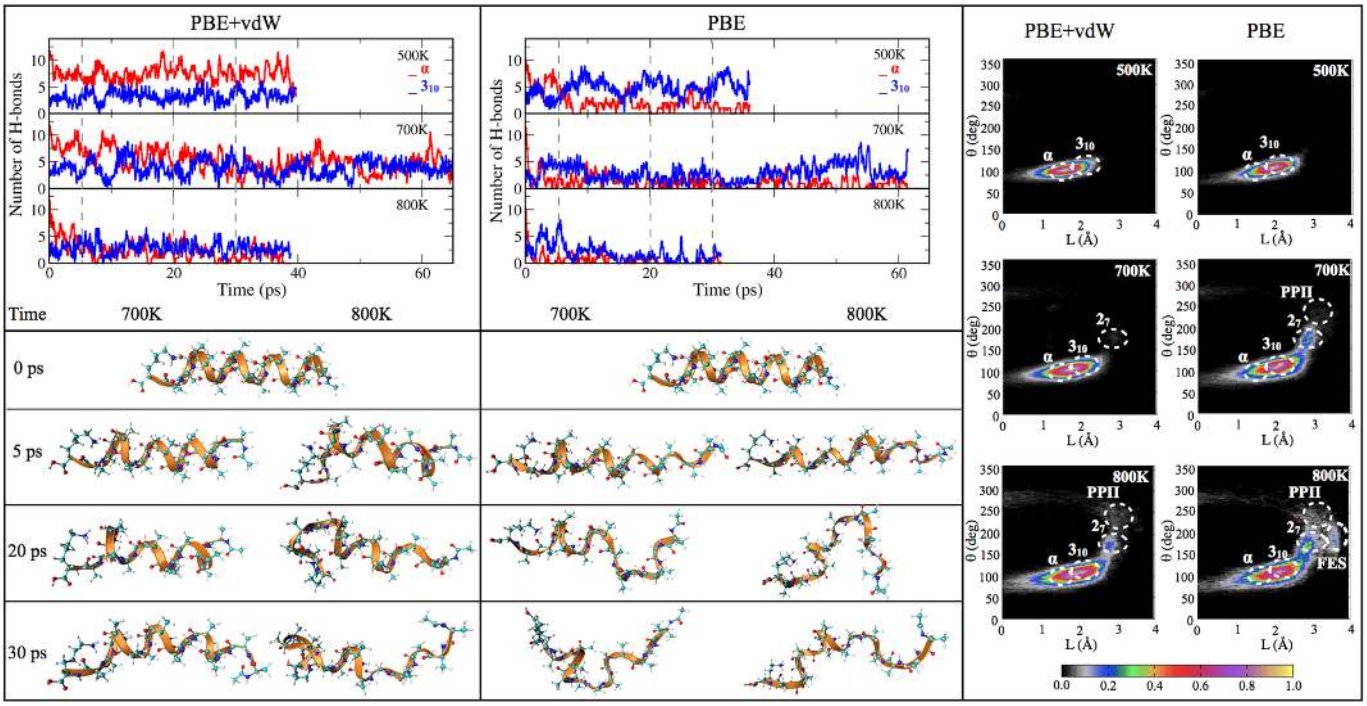

FIG. 2. MD simulations of Ac-Ala$_{15}$-LysH+ peptide at 500, 700, and 800 K. Upper left panel shows the hydrogen bonding pattern throughout the MD trajectory, while lower left panel presents snapshots from MD simulations at 5, 20 and 30 ps. Only α and 3_{10}-helical hydrogen bonds are shown, excluding π, 2_7 and non-bonded residues. Right panel shows the "Ramachandran map" in cylindrical pitch/twist coordinates [20, 22] for all MD simulations. The color code corresponds to the probability (from 0 to 1) of visiting a certain conformation.

nine helices up to 8 residues when vdW interactions were accounted for by using an empirical potential [18]. Since the π-helix has the most compact structure, the vdW interaction is the largest in this case. In contrast, the vdW energy contribution is the smallest for the least-compact fully extended structure. Due to the R^{-6} decay with the distance between two atoms, the largest part of the vdW interaction (95%) comes from within a radius of 5.8 Å and 6.4 Å, for the α and π helices, respectively. This distance range corresponds approximately to the length of the first helical turn for both helices. Hence, vdW interactions are quite localized in polypeptide helices and the accuracy of DFT+vdW for the infinite helices should be of the same order as for small benchmark peptides (see supplementary material).

While the role of interactions underpinning helical stability (van der Waals, hydrogen bonding, and helix termination) has been established above, the missing ingredient to connect our simulations to real experimental systems [12–15] is including temperature effects. We first turn to the free-energy hierarchy of different structure prototypes of Ac-Ala$_{15}$-LysH$^+$. This peptide remains helical in experiment up to $\approx$ 750 K [15]. Let us first analyze the temperature effects on the stability of α-helical Ac-Ala$_{15}$-LysH$^+$ using the harmonic approximation. The potential energy difference between the α-helix and FES for Ac-Ala$_{15}$-LysH$^+$ is -8.7 kcal/mol per residue. The inclusion of zero-point energy increases this value to -8.3 kcal/mol. The free energy difference at 500 K further increases to -5.8 kcal/mol per residue, illustrating the critical role that vibrational free energy plays in destabilizing the α-helix. The missing contribution to the free energy difference between the unfolded random coil and the folded α-helix is the backbone conformational entropy, $\Delta S_{\rm conf}$ [11]. $\Delta S_{\rm conf}$ has been estimated for polyalanine helices from classical force field simulations. Most recent estimates (weakly dependent on the employed force field) are of $\approx$ 2.0 kcal/mol per residue at 500 K [11]. The addition of $\Delta S_{\rm conf}$ reduces the free-energy difference between Ac-Ala$_{15}$-LysH$^+$ α-helix and random coil to -3.8 kcal/mol per residue. However, the α helix remains as the most stable free-energy structure. In contrast to PBE+vdW, in plain PBE the α and 3_{10} structures are similarly stable over the entire range of investigated temperatures, in disagreement with gas-phase experiments.

To analyze the actual dynamics of polyalanine unfolding beyond the harmonic approximation, we performed *ab initio* MD simulations of Ac-Ala$_{15}$-LysH$^+$ with PBE and PBE+vdW at different temperatures (see supplementary material for technical details). At room temperature, PBE+vdW yields a stable α-helix for a 30 picosecond-long AIMD run, while plain PBE yields a predominantly 3_{10}-helical structure. In Fig. 2 we show the number of hydrogen bonds characteristic of a particular helix type (α and 3_{10}) as a function of MD simulation

time up to 65 picoseconds (ps) for 500, 700, and 800 K, starting from a perfect α-helical geometry. For a perfect α-helical structure, there is a total of 12 hydrogen bonds (hydrogen bonds are defined by a maximum CO–NH distance of 2.6 Å, and we do not count hydrogen bonds to the LysH+ termination). PBE+vdW preserves a helical structure throughout the MD simulation at 500 K, with $\approx$ 80% of the H-bonds being of α-helical type at 500 K. Markedly different behavior is predicted by plain PBE. After 10 ps at 500 K, PBE preserves less than two "α-helical" H-bonds on average and the helical part ($\approx$ 50%) of the polypeptide is mainly of 3_{10} nature. Clearly, PBE and PBE+vdW explore a different region of the conformational space as illustrated by the pitch/twist map [20, 22] in the right panel of Fig. 2. At 700 K, PBE+vdW yields a helix of a mixed α-helical and 3_{10}-helical nature, but the overall helical structure is preserved even in MD simulation as long as 65 ps. In contrast, at 800 K, the α-helical structure has essentially disappeared after 10 ps. Once again, PBE yields a very different picture at 700 K and 800 K simulations. In both cases, the α-helical H-bonds largely disappear after 5-7 ps. At 800 K the polypeptide is essentially unfolded after 7 ps and remains unfolded up to 30 ps. It is remarkable that our PBE+vdW results are thus far consistent with the experimental upper limit for helix stability of $T \approx$750 K, while this is not at all the case for plain PBE, not even for 3_{10} instead of α. In addition, also the conformational landscapes with and without vdW interactions differ markedly from one another. For example, plain PBE conformers appear to be significantly more elongated, as illustrated by MD geometry snapshots and pitch/twist plots in Fig. 2. We should remark that longer AIMD simulations would be required to precisely determine the unfolding temperature.

In conclusion, direct *ab initio* simulations reveal that vdW interactions explain the remarkable stability of polyalanine helices *in vacuo* up to high temperatures. We have explicitly evaluated the vibrational energy and entropy using the harmonic approximation, while the anharmonic and multiple-basin effects have been directly addressed using AIMD simulations at different temperatures. The vibrational free energy has been found to play an important role in reducing the stability of the gas-phase α helix, whereas the inclusion of vdW interactions dramatically changes the conformational landscape explored in the *ab initio* dynamics of medium sized polypeptides. The recently developed DFT+vdW method [28] (see Refs. [16, 29–31] for alternative schemes) is the key piece that provides us with a secondary structure peptide energy landscape close to "gold standard" quantum-chemical accuracy, but at the much lower computational cost of DFT-PBE.

[1] T. E. Creighton, *Proteins: Structures and Molecular Properties* (W. H. Freeman, 1992).
[2] L. Pauling, R. B. Corey, and H. R. Branson, Proc. Natl. Acad. Sci. USA **37**, 205 (1951).
[3] C. B. Anfinsen, Science **181**, 223 (1973).
[4] W. Chin, F. Piuzzi, I. Dimicoli, and M. Mons, Phys. Chem. Chem. Phys. **8**, 1033 (2006).
[5] J. P. Simons, Mol. Phys. **107**, 2435 (2009).
[6] F. Bierau, P. Kupser, G. Meijer, and G. von Helden, Phys. Rev. Lett. **105**, 133402 (2010).
[7] J. Benesch and C. V. Robinson, Nature **462**, 576 (2009).
[8] M. P. Gaigeot, Phys. Chem. Chem. Phys. **12**, 3336 (2010).
[9] J. A. Stearns, C. Seaiby, O. V. Boyarkin, and T. R. Rizzo, Phys. Chem. Chem. Phys. **11**, 125 (2009).
[10] M. Rossi *et al.*, J. Phys. Chem. Lett. **1**, 3465 (2010).
[11] R. L. Baldwin, J. Mol. Biol. **371**, 283 (2007).
[12] R. A. Hudgins, M. A. Ratner, and M. F. Jarrold, J. Am. Chem. Soc. **120**, 12974 (1998).
[13] R. R. Hudgins and M. F. Jarrold, J. Am. Chem. Soc. **121**, 3494 (1999).
[14] M. Kohtani, B. S. Kinnear, and M. F. Jarrold, J. Am. Chem. Soc. **122**, 12377 (2000).
[15] M. Kohtani, T. C. Jones, J. E. Schneider, and M. F. Jarrold, J. Am. Chem. Soc. **126**, 7420 (2004).
[16] Q. Wu and W. Yang, J. Chem. Phys. **116**, 515 (2002).
[17] M. Elstner *et al.*, Chem. Phys. **256**, 15 (2000).
[18] H. Liu *et al.*, Proteins: Struct. Funct. Genet. **44**, 484 (2001).
[19] J. J. Dannenberg, Adv. Protein Chem. **72**, 227 (2006).
[20] J. Ireta, J. Neugebauer, M. Scheffler, A. Rojo, and M. Galvan, J. Am. Chem. Soc. **127**, 17241 (2005).
[21] J. Ireta, J. Neugebauer, M. Scheffler, A. Rojo, and M. Galvan, J. Phys. Chem. B **107**, 1432 (2003).
[22] E. Penev, J. Ireta, and J.-E. Shea, J. Phys. Chem. B **112**, 6872 (2008).
[23] J. Kubelka, R. Huang, and T. A. Keiderling, J. Phys. Chem. B **109**, 8231 (2005).
[24] Bour, J. Kubelka, and T. A. Keiderling, Biopolymers **65**, 45 (2002).
[25] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[26] G. E. Job *et al.*, J. Am. Chem. Soc. **128**, 8227 (2006).
[27] A. Tkatchenko *et al.*, MRS Bulletin **35**, 435 (2010).
[28] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[29] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, J. Chem. Phys. **114**, 5149 (2001).
[30] S. Grimme, J. Comp. Chem. **27**, 1787 (2006).
[31] P. Jurecka, J. Cerny, P. Hobza, and D. R. Salahub, J. Comp. Chem. **28**, 555 (2007).
[32] J. Cerny, J. Vondrasek, and P. Hobza, J. Phys. Chem. B **113**, 5657 (2009).
[33] P. Jurecka, J. Sponer, J. Cerny, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[34] V. Blum *et al.*, Comput. Phys. Comm. **180**, 2175 (2009).
[35] V. Havu, V. Blum, P. Havu, and M. Scheffler, J. Comp. Phys. **228**, 8367 (2009).
[36] Guo and M. Karplus, J. Phys. Chem. **98**, 7104 (1994).