Impact of Vibrational Entropy on the Stability of
Unsolvated Peptide Helices with Increasing Length

Mariana Rossi,* Matthias Scheffler, and Volker Blum

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany
E-mail: rossi@fhi-berlin.mpg.de

Abstract

Helices are a key folding motif in protein structure. The question which factors determine helix stability for a given polypeptide or protein is an ongoing challenge. Here we use van der Waals corrected density-functional theory to address a part of this question in a bottom-up approach. We show how intrinsic helical structure is stabilized with length and temperature for a series of experimentally well studied unsolvated alanine based polypeptides, Ac-Ala<sub>n</sub>-LysH<sup>+</sup>. By exhaustively exploring the conformational space of these molecules, we find that helices emerge as the preferred structure in the length range \( n=4-8 \) not just due to enthalpic factors (Hydrogen bonds and their cooperativity, van der Waals dispersion interactions, electrostatics), but importantly also by a vibrational entropic stabilization over competing conformers at room temperature. The stabilization is shown to be due to softer low-frequency vibrational modes in helical conformers than in more compact ones. This observation is corroborated by including anharmonic effects explicitly through \textit{ab initio} molecular dynamics, and generalized by testing different terminations and considering larger helical peptide models.

Keywords: polyalanine | vibrational entropy | density-functional theory | helix stability crossover | secondary structure

*To whom correspondence should be addressed
Introduction

Polypeptide helices are a key secondary structure motif in a wide range of polypeptides and proteins.\textsuperscript{1–3} It is well known that some amino acids (e.g., alanine) exhibit a stronger helix propensity than others,\textsuperscript{4–11} but the fact that the helical structure is so abundant\textsuperscript{12,13} in peptides and proteins is still intriguing. Helices appear in many of the most important proteins\textsuperscript{14–16} and are also an important ingredient in protein design (e.g., Refs.\textsuperscript{13,17,18}). The ability to quantify the competition of peptide helices with other possible structure elements is thus critical for properties related to biology, but also to rather non-biological chemistry, such as enzyme catalysis in organic solvents,\textsuperscript{19,20} at solid surfaces,\textsuperscript{21} in thin films,\textsuperscript{22} or peptide-driven synthesis of nanoparticles.\textsuperscript{23,24}

From a thermodynamic point of view, there are at least two possible limits in which helices compete with other structure prototypes. Towards high temperature, one expects the transition to a random coil,\textsuperscript{25} which should become entropically favored as the temperature increases. Towards low temperature, however, helices may themselves compete with other enthalpically stable conformations. In fact, in the most interesting regime, namely at ambient temperatures, stability may be determined by a delicate balance between enthalpy and entropy.\textsuperscript{26} Using density-functional theory (DFT) including van der Waals (vdW) dispersion interactions,\textsuperscript{27} we here unravel this balance quantitatively for the emergence of helical structure in a particularly well studied series of unsolvated polyalanine based peptides Ac-Ala\textsubscript{n}-LysH\textsuperscript{+}, \(n=4\text{-}8\).\textsuperscript{28–31} We consider explicitly not just the helical part of conformational space, but actually the much larger, general low-energy conformational space of the peptides, of which helices are a part. In this paper we show from first principles: (i) a comprehensive search of the conformational space for Ac-Ala\textsubscript{n}-LysH\textsuperscript{+}, \(n=4\text{-}8\), involving thousands of possible conformers; (ii) harmonic free energy calculations for a broad range of the most favorable structural candidates; (iii) the role of anharmonicities in the potential energy surface; and (iv) a theoretical comparison to longer model peptides with a different termination, considering only helical motifs, in order to clarify the impact of Lys on the soft vibrational modes. Our work is based on an exhaustive prediction of low-energy conformers using DFT and the PBE\textsuperscript{32} exchange-correlation functional, corrected to account for long-range vdW interac-
tions (here called PBE+vdW). This level of theory treats accurately, and without system-specific empirical parameters, critical length-dependent contributions such as H-bond cooperativity and vdW dispersion interactions, including their effect on vibrational frequencies and in \textit{ab initio} molecular dynamics. Our key finding is that there is a significant vibrational entropic stabilization of helices compared to other, more compact conformers, a contribution that should indeed make a difference in actual proteins as well. This contribution is intrinsic to the helix and should therefore act largely independently, not entangled with environment-dependent terms such as a solvent entropy.

\textbf{Terms that shape the potential energy surface of polypeptide helices}

Beginning with the terms that shape the potential-energy surface (PES), known reasons for helix stability include \cite{37,38} (i) their efficient hydrogen-bond (H-bond) network and increasing H-bond cooperativity with helix length, \cite{31,33–36} (ii) suitably bonded and/or electrostatically favorable termination, for instance the LysH\textsuperscript{+} termination considered here, \cite{28,39–45} and (iii) remarkably, a rather specific favorable contribution of vdW dispersion interactions for α-helices. \cite{31,46} Clearly, the peptide chain length plays a role: Too short chains have too few and too weak hydrogen bonds for helices to compensate the cost of strain in the backbone. \cite{29,43,47–49} In practice, environment effects will necessarily influence helix stability. \cite{49–52} In an aqueous medium, the hydrophobic effect will be prominent, \cite{53} but it is important to note that peptides and proteins are also found (and used) in different environments. In water-poor conditions, like membranes, helices are frequently observed. \cite{54,55} Understanding the relative stability of key structure elements will also be critical in “man-made” chemical environments (e.g., the examples given at the outset of this paper \cite{19–24}). In fact, the longer members ($n \geq 8$) of the polyalanine based peptide series studied here are helices in experiment under vacuum conditions. \cite{28,30,56} These helices are stable \textit{in vacuo} even up to extreme temperatures (not expected in solution), \cite{31,57} or after soft landing on a surface. \cite{58}

The potential energy surface also shapes the entropy, and thus the effect of the temperature $T$. With increasing $T$, the conformational entropy of the backbone will favor an unfolded
state\textsuperscript{25,38,59,60} (so-called “random coil”), while, at low $T$, helices may also compete with other, enthalpically more stable conformers. For instance, gas-phase ion mobility spectrometry (IMS) by Jarrold and coworkers\textsuperscript{26} showed that the Ac-Ala$_4$-Gly$_7$-Ala$_4$H$^+$ polypeptide is helical at $T=400$ K but globular at room temperature. A similar structural change was observed in experiments involving multiply protonated polyalanine peptides in the gas-phase in Ref.\textsuperscript{61} Empirical force-field based simulations by Ma and coworkers\textsuperscript{62} of more than 60 small peptides indicate that the vibrational entropy (harmonic approximation) could stabilize $\alpha$-helices or $\beta$-hairpins over fully relaxed conformer ensembles obtained with random structures generated from high $T$ trajectories – i.e., attempting to mimic the “unfolded structure.” A gain in vibrational entropy due to the existence of low frequency vibrations stemming from helical and $\beta$-sheet motifs in biomolecules has also been studied with empirical theoretical models by K. Chou and coworkers.\textsuperscript{63} Recently, Plowright and coworkers\textsuperscript{64} used DFT including dispersion contributions (the B97-D\textsuperscript{65} exchange-correlation functional) to suggest that, for a small neutral four-residue peptide, $\beta$-sheets and conformers containing 3$_{10}$ helical loops are stabilized by the harmonic vibrational entropy at finite temperatures.

In the present work, we provide independent, unambiguous, and quantitative computational evidence that the vibrational entropy acts to stabilize helical conformers with increasing temperature over more compact, enthalpically competitive structures. What is new compared to previous studies (e.g., Ma \textit{et al.} above\textsuperscript{62}) is that we have access to the full low-energy conformational ensemble, encompassing thousands of conformations that compete with helical structures towards low temperature – for peptides of significant length (up to 110 atoms), based on an accurate, fully quantum-mechanical level of theory that captures all potentially relevant effects (hydrogen bond cooperativity, strain, polarizability changes, etc.) in a completely non-empirical way. We are thus able to trace the reason for the stabilization of helices to their softer low-frequency modes, in comparison to more compact, competing low-energy conformations. We show that this lowering of frequencies persists also for the anharmonic case, evidenced by the vibrational density of states derived from \textit{ab initio} molecular dynamics simulations.

We focus on polyalanine-based peptides, since alanine is known to have a high helix propen-
sity both in solution \(^37,66\) and in vacuo. \(^8\) For Ac-Ala\(_n\)-LysH\(^+\) (\(n=4-20\)) in the gas phase, IMS\(^29\) and first-principles calculations compared to experimental vibrational spectroscopy at room temperature\(^30\) suggest a cross-over from non-helical to helical preferred conformers as a function of polyalanine chain length. For \(n=5\) there is a competition between different conformers, while the \(n=10\) and \(n=15\) conformers are found to be firmly in the helical range.\(^30\) A similar globule to helix transition also occurs for sodiated polyalanine.\(^67\) For idealized polyalanine helix models, various past first-principles studies have dissected helix-stabilizing factors,\(^35,36,49,68\) including an analysis of the role of electrostatics, H-bond cooperativity, and van der Waals interactions for the stability of unsolvated polyalanine-based helices against unfolding in our own group.\(^31\) This class of systems is thus an ideal testing ground to clarify the structural competition of non-helical (compact) and helical conformers as a function of chain length also toward the opposite temperature limit (low temperature, folded state). In the following, we address conformational preference of Ac-Ala\(_n\)-LysH\(^+\) in vacuo for \(n=4-8\), i.e., the length range in which the helical preference at room temperature develops.

**Method**

We are here primarily interested in finding the lowest energy conformers of Ac-Ala\(_n\)-LysH\(^+\), \(n=4–8\). For an accurate description of the underlying potential-energy surface, we turn to first principles, following the same conformational search strategy previously described in Ref.\(^69\) The method which we employ is density-functional theory with the PBE+vdW\(^27,32\) exchange-correlation functional. The vdW dispersion correction\(^27\) consists of a \(C_6[n]/R^6\) term, where the \(C_6\) coefficients depend on the self-consistent electronic density \(n\).

As explained in Ref.,\(^69\) we start searching the conformational manifold by scanning for a wide range of structure candidates using a force field (OPLS-AA\(^70\)) based basin hopping algorithm supplied with the TINKER\(^71\) package. No geometrical constraints are enforced in these searches. We here perform the basin hopping search within a 50 kcal/mol energy window (2 eV), using
15 torsional modes. The procedure leads to at least $10^5$ conformers for each of the molecules in question. In our searches, both the Lys residue and the C-terminal COOH are considered protonated throughout, which is the known gas-phase preference when the N-terminus is capped.\textsuperscript{28,72,73}

The force field guided search for structure candidates has the sole goal of providing an exhaustive list of candidate structures for the second step of our search. For the force field used, this list is converged with respect to the basin hopping parameters, as shown in the SI and in Ref.\textsuperscript{74}

The second step consists of full relaxations of a wide range of low-energy force field conformers using the PBE+vdW functional. We perform these relaxations with the all-electron, localized basis program FHI-aims.\textsuperscript{75,76} Using “light” settings for integration grids and basis sets, we relax 1068, 1000, 800, 800, and 820 conformers for Ac-Ala\textsubscript{n}LysH\textsuperscript{+}, $n$=4, 5, 6, 7, 8, respectively. We then post-relax up to 50 low energy structures using the FHI-aims numerical “tight” settings and “tier 2” basis set, which provide essentially converged energy hierarchies free from basis set superposition error.\textsuperscript{69,75} All energy gradients were converged to magnitudes less than $10^{-3}$ eV/Å in the final relaxed structures.

<table>
<thead>
<tr>
<th>Oxy. Bound to</th>
<th>Oxy. Bound to</th>
</tr>
</thead>
<tbody>
<tr>
<td>O(Ac) NH\textsubscript{4} ($\alpha$)</td>
<td>O(Ac) NH\textsubscript{3} ($3_{10}$)</td>
</tr>
<tr>
<td>O1 NH\textsubscript{6} ($\pi$)</td>
<td>O1 NH\textsubscript{5} ($\alpha$)</td>
</tr>
<tr>
<td>O2 NH\textsubscript{7} ($\pi$)</td>
<td>O2 NH\textsubscript{5} ($\alpha$)</td>
</tr>
<tr>
<td>O3 NH\textsubscript{8} ($\pi$)</td>
<td>O3 NH\textsubscript{6} ($\alpha$)</td>
</tr>
<tr>
<td>O4 OH</td>
<td>O4 NH\textsubscript{7} ($\alpha$)</td>
</tr>
<tr>
<td>O5 NH\textsubscript{3}\textsuperscript{+}</td>
<td>O5 NH\textsubscript{7}\textsuperscript{+}</td>
</tr>
<tr>
<td>O6 NH\textsubscript{3}\textsuperscript{+}</td>
<td>O6 NH\textsubscript{3}\textsuperscript{+}</td>
</tr>
<tr>
<td>O7 NH\textsubscript{3}\textsuperscript{+}</td>
<td>O7 free</td>
</tr>
<tr>
<td>O8 free</td>
<td>O8 NH\textsubscript{3}\textsuperscript{+}</td>
</tr>
<tr>
<td>O(COOH) free</td>
<td>O(COOH) free</td>
</tr>
</tbody>
</table>

Figure 1: Center: Energy hierarchies obtained for Ac-Ala\textsubscript{8}-LysH\textsuperscript{+} with the OPLS-AA force-field, PBE+vdW with “light”\textsuperscript{75} computational settings, and PBE+vdW with “tight”\textsuperscript{75} computational settings. All conformers are fully relaxed at each level of theory. Highlighted in red are two slightly different force field conformers which relax to the lowest energy Family 1 ($\alpha$-helical) conformer in the PBE+vdW functional. Highlighted in violet is the lowest energy OPLS-AA conformer. Left: H-bond patterns of the force-field minimum-energy conformer. Right: H-bond patterns of the PBE+vdW minimum-energy conformer.
In order to illustrate how the conformer energy hierarchies change between the OPLS-AA force-field and the PBE+vdW functional, in Figure 1 we show the OPLS-AA and PBE+vdW energetic hierarchies of all conformers relaxed with the PBE+vdW functional for Ac-Ala8-LysH+. There is a considerable energy rearrangement between the OPLS-AA force field and the PBE+vdW functional (“light” settings, “tier 1” basis set). Importantly, the OPLS-AA force-field predicts a different global minimum structure from the one predicted with the PBE+vdW functional. While the PBE+vdW functional correctly predicts the α-helix as the minimum energy structure, the OPLS-AA force field predicts a mostly π-helical structure. Also shown in Figure 1 is the low energy conformational energy hierarchy calculated with the PBE+vdW functional and FHI-aims “tight” settings. The energy hierarchy of the conformers changes by maximally 20 meV between “light” and “tight” settings. Although the energy hierarchy can strongly vary between the force field and the PBE+vdW functional, for Lys terminated polyalanine our experience is that the lowest energy PBE+vdW conformer is found among the first few hundred relaxations of low-energy force field conformers. We also note that force fields may be prone to structure-specific errors. For example, we observe that OPLS-AA systematically overestimates the energy of 310-helical structures in vacuo. To exclude any unwanted impact of the energetic overestimation of 310-helices, we performed additional constrained basin-hopping searches for all n, in which we constrain the central H-bonds of the molecules to remain 310-helical (see Figure S1 of the SI), followed by individual, unconstrained PBE+vdW relaxations. For n=5, we also tested the influence of other DFT functionals, reported in Figure S2 of the SI. Differences of 50 meV (less than 10 meV per residue) in the energy hierarchy can arise as a result of changing the functional, but as long as the vdW correction is included, the overall energy hierarchy trends are kept.

The PBE+vdW relaxed (“light” settings) conformers were sorted into “families” according to their H-bond pattern. We define a H-bond to be present when an O acceptor atom is closer than 2.5 Å to a H donor atom. Within each H-bond family thus defined, small conformational variations are still possible, e.g., by slight bends of the backbone atoms or different rotamers of the LysH+ side chain. Typically, the lowest-energy PBE+vdW conformer is found among the family members.
arising within 5 kcal/mol ($\approx 0.2$ eV) of the lowest-energy force field conformer.

Figure 2: Numerical convergence of the finite-difference based computed frequencies reported in this paper. (a) Vibrational frequencies of Family 1 of $n=8$ (110 atoms, $\alpha$-helix) in cm$^{-1}$ against the vibrational mode number. (b-d) Numerical changes of frequencies that arise when varying: (b) the atom displacement for calculation of finite differences $\delta$; (c) the force relaxation threshold “max. force”; (d) and the size of the basis set, “tier”. The reference is taken as $\delta=0.0025\text{Å}$, Max. force $10^{-4}$ eV/Å, and tier 2 basis set.

The PBE+vdW structure optimizations are followed by harmonic vibrational and (rigid-body) rotational free energy calculations for the lowest energy conformers of the H-bond families lying within 0.12 eV ($\approx 3$ kcal/mol) of the global energy minimum for each $n$. The harmonic vibrational frequencies and intensities were also computed in FHI-aims, using a finite difference approach. Our primary conclusions regarding free-energy differences in this paper are based on the lowest-frequency vibrations of our conformers. To demonstrate that these modes are accurately converged in our finite-difference approach, Figure 2 shows the differences in predicted frequencies when varying three important numerical choices: the finite-difference displacement length $\delta$ (subfigure b), the force convergence criterion used in the relaxation (subfigure c), and the quantum mechanical basis set (subfigure d). Varying our default choices ($\delta=0.0025\text{Å}$, energy gradients below $10^{-3}$ eV/Å, and tier 2 basis set).
eV/Å, and tier 2 basis set) produces frequency changes below 2 cm$^{-1}$ at most.

For $n=8$, we have also calculated vibrational densities of state beyond the harmonic approximation, taking the Fourier transform of the velocity time autocorrelation function. The trajectories were evaluated with Born-Oppenheimer \textit{ab initio} molecular dynamics, using the PBE+vdW functional. We used a 1 fs time step, and the microcanonical simulations were 21 ps long, with FHI-aims “tight” numerical settings and “tier 2” basis set. The molecules were initially thermalized to approximately room temperature for 5 ps using the Bussi-Donadio-Parrinello$^{77}$ thermostat.

For all molecules containing $n>8$ alanine residues and for the Li$^+$ terminated model structures discussed in this work, no extensive conformational search was performed. These peptides are structure models used specifically for a computer experiment to determine the development of low-frequency vibrational modes with increasing helix length for two different terminations. Their geometries are fully relaxed PBE+vdW structures, using "tight" settings and the tier 2 basis sets.

\section*{Results and Discussion}

\subsection*{Conformational energy hierarchy}

Figure 3 summarizes the energetic ordering of the lowest-energy (PBE+vdW) H-bond families for $n=4$–8 found in the present work. Only the energy of the lowest energy structure belonging to each family is reported, and families are included up to 0.12 eV ($\approx 3$ kcal/mol) of the lowest identified minimum of the PBE+vdW PES. $\alpha$-helical conformers are highlighted in red. The $3_{10}$-helical conformer for $n=4$ is highlighted in blue. We define purely $\alpha$- (or purely $3_{10}$-) helical conformers as those where, counting from the N-terminus, all the backbone CO groups at residues $i$ are either connected to NH groups at residues $i+4$ (or $i+3$) or to the LysH$^+$ side chain (usually the final three or four CO groups at the C terminus). Coordinates and a more detailed analysis of all the geometries shown in Figure 3 are given in the SI.

The conformer associated with the lowest-energy PES minimum for $n=4$ is rather small, connecting almost all its backbone CO groups to the LysH$^+$ termination. The remaining H-bond at
Figure 3: Energy hierarchies (thick horizontal bars), obtained with the PBE+vdW functional, for the low-energy H-bond families of Ac-Ala$_n$-LysH$^+$, $n$=4-8. The “inverted” H-bonds of $n$=5 and 6 are highlighted in yellow and pointed to by an arrow (see text). For each $n$, we include the conformer representatives of the lowest-energy families up to 0.12 eV from the global minimum, as defined by local minima of the PBE+vdW potential-energy surface (PES). We also show their hierarchy after adding the corrections for the harmonic vibrational free energy $F_{\text{harm}}$ at $T$=300 K. Numbers labeling each family, as well as colored structural representations are shown. The placement of the structure pictures in the shaded gray areas is not directly related to the energy axis (y axis), which applies strictly only to the horizontal bars. For these, the $\alpha$-helical conformer of $n \geq 5$ is chosen as the reference (zero) energy. For $n$=4, Family 1 is taken as the reference. $\alpha$-helical conformers are highlighted by red bars. The $3_{10}$-helical conformer for $n$=4 is highlighted by blue bars. The shaded areas in the energy hierarchies indicate the energy difference between the $\alpha$-helix and the nearest non-helical conformers.
the N-terminus is bifurcated, i.e., the oxygen atom is involved into both an $\alpha$- and a $3_{10}$-helical H-bond. This conformer could therefore be classified as the smallest possible $\alpha$-helical prototype in this series. In contrast, the structures that correspond to the lowest-energy PES minima for $n=5$ and 6 are not simple helices. Each contains an “inverted” H-bond where one CO group points to the N-terminus and its connecting NH group points to the C-terminus, producing more compact structures. In Figure 3, these bonds are highlighted in yellow and pointed to by an arrow. For $n=5$, we have previously denoted this conformer as “g-1”.$^{30}$ For $n=7$ and 8, the lowest-energy PES minima correspond to $\alpha$-helices. In each case, they are closely followed by a conformer that we characterize as compact/globular (Families 2 of $n=7$ and 8, with an energy separation of 20 and 60 meV, respectively). Thus, we already observe a cross-over with peptide length to $\alpha$-helical lowest-energy minima of the PES at $n=7$. However, based on the energy hierarchies of the structures of the local PBE+vdW PES minima alone, one would not expect a purely helical ensemble of conformers at room temperature at $n=7$ or 8. Simple Boltzmann factors would indicate a mix of structure candidates. Yet, the experimental work by Kohtani and Jarrold$^{29}$ does suggest a complete room-temperature structural cross-over at $n=8$ at the latest, albeit based on a completely different line of reasoning (water adsorption behavior of “helical” versus “globular” conformers).

For the low-energy conformers in Figure 3, we also compute and show in the same figure the impact of the vibrational free energy at room temperature ($T=300$ K) in the harmonic approximation$^1$. Remarkably, the relative stability of the $\alpha$-helices is systematically enhanced by the vibrational free energy contribution for all $n$. For $n=5$ and 6 the $\alpha$-helical conformers move down in energy with respect to the (non-helical) lowest energy PES minimum. For $n=7$ and 8 the $\alpha$-helices now become the isolated minima. In detail, we observe:

- For $n=8$, the energetic interval between the $\alpha$-helical lowest energy conformer and the nearest globular one (red shaded areas) now amounts to 0.14 eV. The additional Family 1a at 0.13 eV is another $\alpha$-helix with a slightly modified terminating H-bond network. With the vibrational free energy contributions from rigid-body rotations are also included. The differences between these energies for different conformers were found to be of the order of only a few meV, with the maximum difference amounting to 8 meV.
energy included, the energy hierarchy is thus consistent with the experimental claim that \( \alpha \)-helices dominate over all other possible conformers in gas-phase experiments for \( n=8 \).\(^{29,58}\)

- For \( n=7 \), the same qualitative picture emerges. Here, the next remaining conformer (Family 3) at 300 K is a mixed 3\(_{10}/\alpha\)-helix. The competing compact conformers (representatives of Family 2 and 2a) are significantly destabilized by \( F_{harm}(T) \).

- For \( n=6 \), the \( \alpha \)-helix emerges as the room-temperature minimum free energy conformer, but the competing non-helical PES minimum, Family 1, remains close in energy (50 meV).

- For \( n=5 \), the difference between the \( \alpha \)-helix and Family 1 (g-1) decreases by 60 meV at 300 K, compared to the PES minimum, but Family 1 remains overall more favorable.

In essence, there is an uniform stabilization of helical over more compact (globular) structures as the temperature increases. The stabilization tendency increases with peptide length, confirming quantitatively and systematically from first principles the related observations in Refs.\(^{26,62,64}\)

We next demonstrate that it is indeed the vibrational entropy term which is critical, and then pinpoint the physical reason among the low-frequency vibrational modes.

**Origin of the entropic stabilization**

In Figure 4 we analyze the individual quantities composing the vibrational free energy differences between: Families 1 and 2 of \( n=4 \), Families 1 (non-helical) and 2 (\( \alpha \)-helical) of \( n=5 \), Families 1 (non-helical) and 3 (\( \alpha \)-helical) of \( n=6 \), as well as Families 1 (\( \alpha \)-helical) and 2 (compact) for \( n=7 \) and 8. The energy terms plotted are the PBE+vdW PES minimum energy, the harmonic internal energy (containing the zero-point energy) \( \Delta U_{harm}(T) \), and the entropy term \( T S_{harm}(T) \). They are reported in Figure 4 as energy differences, taking the non-helical conformer of each \( n \) as the reference, such that negative slopes correspond to a stabilization of the \( \alpha \)-helices. Upon inspection of Figure 4, we observe a monotonic stabilization of all helical conformers with increasing \( T \).

While for the shortest molecule (\( n=4 \)) there is hardly any observable effect, the stabilization trend is enhanced with increasing length. For \( n=6 \), we predict a cross-over of the lowest-energy structures at \( T \approx 150 \) K. It is clear that among the individual contributions to the vibrational part of the free
Figure 4: Energy differences for each term of the harmonic free energy as a function of temperature for: Families 1 and 2 of $n=4$; Families 1 (non-helical) and 2 ($\alpha$-helical) of $n=5$; Families 1 (non-helical) and 3 ($\alpha$-helical) of $n=6$; Families 1 ($\alpha$-helical) and 2 (compact) of $n=7$; and Families 1 ($\alpha$) and 2 (compact) of $n=8$. In each case, the non-$\alpha$-helical conformer was taken as the reference (i.e. negative slopes mean $\alpha$-helix stabilization). The PBE+vdW energy differences are plotted with a yellow dash-dash-dot line, the harmonic internal energy $\Delta U_{harm}$ as black dash-dot lines, the entropy term $TS_{harm}$ as red dashed lines, and the sum $F_{harm}$ as green full lines.
energy $F_{\text{harm}}(T)$, the entropy term $TS_{\text{harm}}(T)$ is indeed always the most important helix-favoring term. The zero-point-energy ($\Delta U_{\text{harm}}$ at $T=0$), is also favors the helices, but on a smaller energy scale.

\begin{align*}
\Delta Q_i &= k_B T \ln \left\{ 1 - \exp \left[ -\hbar \omega_i / (k_B T) \right] \right\}, \\
\Delta Q_{\text{total}} &= \sum_i \Delta Q_i.
\end{align*}

The contribution of each mode to the helix stabilization free energy at room temperature, for the same conformers treated in Figure 4 for each $n$. The definition of the quantities plotted is $\Delta Q_i = k_B T \ln \left\{ 1 - \exp \left[ -\hbar \omega_i / (k_B T) \right] \right\}$, and $\Delta Q_{\text{total}} = \sum_i \Delta Q_i$. The differences $\Delta$ shown are between the $\alpha$-helical conformer of each $n$, and the non-helical or compact conformer at each $n$. Positive red bars denote a contribution to $\alpha$-helix stabilization with respect to the non-helical, compact conformers and green negative bars denote a contribution to destabilization. For all $n$, the wavenumber corresponding to the highest mode number plotted is approximately 570 cm$^{-1}$.

The contribution of each mode to the helix stabilization free energy at room $T$ can be quantified by calculating the single mode temperature-dependent contribution to the vibrational free energy $\Delta Q_i = k_B T \ln \left\{ 1 - \exp \left[ -\hbar \omega_i / (k_B T) \right] \right\}$, where $\omega_i$ are the vibrational frequencies. The quantity that enters the the computation of $\Delta F$ between different conformers here presented is the difference $\Delta Q_{\text{total}} = \sum_i \Delta Q_i$. In order to analyze the contribution of each vibrational mode $i$ in free energy differences, it is thus convenient to plot $\Delta Q_i / \Delta Q_{\text{total}}$, the sum of which over all modes is normalized to one. These mode-dependent contributions are shown in Figure 5 for $n=5$-8 and for the same conformers (helical vs. non-helical) included in Figure 4. The included modes $i$ are plotted up to approximately 570 cm$^{-1}$ for each $n$. Modes in this frequency range dominate the temperature dependent contribution to the harmonic free energies, such that conformers with softer low-energy
vibrational modes will be effectively stabilized against those that have stiffer low-energy vibrations. In Figure 5, positive bars correspond to softer modes in the \( \alpha \)-helical conformer, and thus a contribution to \( \alpha \)-helical stabilization. We observe that the low-frequency modes contribute uniformly and overwhelmingly to the observed stabilization, reflecting the fact that \( \alpha \)-helices have overall softer vibrational modes than compact conformers. The stabilization contribution “de-localizes” over higher energy modes as the length of the molecules is increased (the maximum frequency shown in all plots of Figure 5 is approximately the same for all \( n \)). The same qualitative conclusions hold for other functionals as well, as we exemplify in Figure S3 of the SI for the BLYP\(^{78}\)+vdW functional and \( n=8 \).

Table 1: Position of the lowest vibrational mode, in cm\(^{-1}\) and calculated with the PBE+vdW functional, for the lowest energy conformer of each family shown in Figure 3, for \( n=4-8 \). We use red characters to indicate the \( \alpha \)-helical conformers, a * symbol for the g-1-like conformers of \( n=5 \) and 6, and blue characters for the 3\(_{10}\)-helical conformer.

<table>
<thead>
<tr>
<th>length/conformer</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>23</td>
<td>22*</td>
<td>20*</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Family 1a</td>
<td>27</td>
<td>23</td>
<td>21</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Family 1b</td>
<td>25</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family 2</td>
<td>17</td>
<td>13</td>
<td>20</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>Family 2a</td>
<td></td>
<td>23</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family 3</td>
<td>27*</td>
<td>20</td>
<td>8</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Family 4</td>
<td>17</td>
<td>18</td>
<td>20*</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

For the PBE+vdW functional, Figure 5 suggests that the very lowest-frequency vibrational modes can be taken as an indicator of the observed helical stabilization with respect to more compact conformers. Table 1 shows the frequencies (in cm\(^{-1}\)) corresponding to the first normal modes of all conformers shown in Figure 3 of the manuscript, as obtained with the PBE+vdW functional. The lowest-energy \( \alpha \)-helices, marked in red in Table 1, show first vibrational normal modes between 8 cm\(^{-1}\) and 13 cm\(^{-1}\). The same is true for families 1a of \( n=7 \) and 8, which only differ in details of the termination. In contrast, the conformers that have first vibrational modes of at least 20 cm\(^{-1}\) are all compact and non-helical, for example: 28 cm\(^{-1}\) for Family 2 of \( n=7 \), and 22 cm\(^{-1}\) for Family 1 (g-1 motif) of \( n=5 \). The g-1 motif has the first vibrational mode lying around 20 cm\(^{-1}\) for \( n=5, 6, \) and 7 (marked with a * symbol in Table 1).
Here, a comment regarding structures that are more elongated than $\alpha$-helices is in order. It is customary to compare the stability of $\alpha$-helices with $3_{10}$ helices and $\beta$-sheets or fully extended structures (e.g., in references\textsuperscript{35,36,79–85} and many others). Following our rationale above, the more extended the structure is, the softer the low vibrational modes will be. This indeed happens for the most extreme case, the fully extended structure (FES), where we find the first vibrational mode to lie around only 2 cm\textsuperscript{-1} (calculated for $n=8$ and 15). In fact, entropically stabilized $\beta$-sheets in neutral polyalanine in the gas-phase have been suggested in Ref.\textsuperscript{44} For $3_{10}$ helices, in our own structure searches for $n=4$-8, we find the first vibrational mode to lie always very close to their $\alpha$-helical counterpart. Accordingly, it is the enthalpic energy difference that favors $\alpha$-helices specifically over $3_{10}$ for all molecules studied. For infinite periodic structures, calculation of phonons and vibrational free energies for $\alpha$-, $3_{10}$-, and $\pi$-helices, and the FES in Ref.\textsuperscript{79} corroborate our results. There, all helices are destabilized with respect to the FES at 300 K. The $\pi$-helix, which is the most compact among the helices studied in Ref.,\textsuperscript{79} is most destabilized by the vibrational entropy term.

**Beyond the harmonic approximation: Ab initio molecular dynamics**

![Figure 6: Vibrational density of states calculated as the Fourier transform of the velocity autocorrelation function of 21ps long microcanonical AIMD simulations at $\langle T \rangle = 300$ K, for Family 1 ($\alpha$-helix) and Family 2 (compact) of Ac-Ala\textsubscript{8}-LysH\textsuperscript{+}. Light colored bars correspond to the harmonic normal frequencies of vibration for the respective conformers.](image)

We next show that the observations discussed in the last section should hold also in a fully anharmonic picture. Even at relatively low temperature, where the overall structure is kept, we expect first the inherent anharmonicity of the local, nearly harmonic PES, then the lowest-barrier transitions between basins (side chain rotations), and then transitions between locally different backbone conformations and H-bond networks\textsuperscript{86,87} to contribute. Unfortunately, a direct calcu-
lation of these terms (e.g., by thermodynamic integration) is computationally prohibitive in DFT. We can, however, use explicit *ab initio* molecular dynamics simulations to gain some insights. In Figure 6, we compare the Fourier-transformed velocity autocorrelation functions of the Family 1 (helical) and Family 2 (compact non-helical) conformers of \( n=8 \), extracted from explicit microcanonical *ab initio* molecular dynamics simulations (21 ps total time, 1 fs time step, initially thermalized to approximately room temperature). The Fourier transform of the velocity autocorrelation function corresponds to a vibrational density of states (VDOS). At \( T=0 \) the VDOS should, in the approximation of classical nuclei, reflect the harmonic vibrational modes, also shown in Figure 6. Compared to these modes, the onset of the calculated VDOS at \( T\approx300 \) K is noticeably shifted towards lower frequencies for both conformers, but the onset frequency for the helix is still significantly lower than for the non-helical structure. Thus, the lower vibrational frequencies of the helix are also carried over to the full (anharmonic) motion of the conformers. In addition, the integral over the VDOS up to 50 cm\(^{-1}\) is 6\% larger for the \( \alpha \)-helical Family 1 than for the compact Family 2, i.e., the general downshift of frequencies in this region is also preserved.

Regarding the local structural stability during the AIMD simulation, we can analyze the detailed evolution of the H-bond network of the molecules, which we here show in Figure 7(a) and (b). The colored bars in those figures show whether a specific kind of H-bond is present at a given time during the simulation. The oxygen atoms participating in the possible H-bonds are labeled from \( O(\text{Ac}) \), belonging to the acetate group of the N-terminus, up to \( O8 \), belonging to the Alanine residue closest to the C-terminus. We consider a H-bond every (C)O donor - NH acceptor pair lying at a distance closer than 2.5\( \text{Å} \). Through this definition, it is possible to count a H-bond even if the charge densities would not characterize, strictly speaking, a bond. It is more important, however, that we do not miss any possible bond, and this is guaranteed by our definition. The respective fractions of the time that *helical* bonds of a specific type exist during the simulation are shown on the right side of the plot. The sum of these fractions for a given oxygen can exceed 100\% because when a bond is bifurcated we count it twice, once for each kind of H-bond.

We observe that the H-bond pattern of Family 2 (compact) stays essentially the same through-
Figure 7: (a) Detailed H-bond network evolution during the 21 ps of AIMD simulation (PBE+vdW, microcanonical ensemble) for the $\alpha$-helical Family 1 of n=8. (b) Detailed H-bond network evolution during the 21 ps of AIMD simulation (PBE+vdW, microcanonical ensemble) of the compact Family 2 of $n=8$. Numbers shown on the right side of the plots reflect the respective fractions of time that helical bonds of a specific type exist during the simulation. Since a bifurcated H-bond is counted twice, the overall sum for each oxygen can exceed 100%.

out the entire simulation. In contrast, Family 1 displays local structural fluctuations in the helical part, occasionally forming short-lived $3_{10}$-like H-bond connections. Similar fluctuations also occur in simulations of longer helices (e.g., Ac-Ala$_{15}$-LysH$^+$ in Ref. 30). It seems plausible that the overall greater “floppiness” of the helix compared to the more compact non-helical H-bond network adds another favorable entropic contribution at room temperature. This direct observation that we make has been conjectured as the motive for the loss in entropy observed in $\alpha$-helical formation, compared to a polyproline-II (PPII) helical structure in Ref. 83. According to our reasoning, the PPII structure, being less compact than the $\alpha$-helices, should indeed show more fluctuations, just like $\alpha$-helices do if compared to more compact conformers.

The role of the termination

Finally, we show that, especially for the short peptides considered here, there are two independent aspects to the stabilization of helical conformers: length and termination. To isolate their role, we examine the character of the lowest-frequency modes as a function of peptide length (also for longer helices) for two different terminations. The first is the LysH$^+$-terminated series, which is the
main subject of this work. We contrast this series with Li\(^{+}\)-terminated polyalanine helices Ala\(_n\)-Li\(^{+}\), a much more rigid termination as we shall see. For the latter, we define \(\alpha\)-helical structures for all \(n\) (fully relaxed in DFT-PBE+vdW), which are at least locally stable when Li\(^{+}\) is placed in contact with the last three residues at the C terminus. In Figure 8(a), we compare the position of the first and second vibrational modes of these \(\alpha\)-helical geometries of Ala\(_n\)-Li\(^{+}\), \(n=5-15\), and 20, with the \(\alpha\)-helical geometries of Ac-Ala\(_n\)-LysH\(^{+}\), \(n=5-15\), and 19. For both peptide series, there is a monotonic decrease of the first and second vibrational frequencies for \(n \geq 7\), but the frequency starting point is much higher for the Li\(^{+}\) termination (26 cm\(^{-1}\) at \(n=7\)) than for the LysH\(^{+}\) termination. A softening trend of the respective modes in neutral polyalanine helices with increasing length has also been observed in Ref.\(^{88}\) In order to characterize this vibrational mode in more detail, Figure 8(b) visualizes the displacement of the backbone atoms of Ac-Ala\(_{19}\)-LysH\(^{+}\) when deforming this molecule along the first vibrational mode. Figure 8(c) shows the relative length changes of the hydrogen bonds in the structure upon deformation along this mode. Subfigures (d) and (e) show the equivalent data for Ala\(_{20}\)-Li\(^{+}\). For both molecules, the vibration spans the helical part of the structures. For the LysH\(^{+}\) terminated molecule, we see that the actual LysH\(^{+}\) termination (connected to the last four CO residues) is clearly involved in the vibration. In contrast, the hypothetical Li\(^{+}\) charged termination constrains especially the C terminus to be much more rigid, as evidenced by the almost zero change of all Li-O distances. Here, the N-terminus is much more involved in the lowest-frequency modes. The same trends are observed for the smaller molecules in both series. Movies containing 3D visualizations of the first vibrational modes for helical and compact structures are contained in the SI. We thus conclude two points:

(i) Helices are entropically favored by allowing delocalized, soft low-frequency modes that we do not observe in competing, more compact conformers of the same LysH\(^{+}\) termination (evidenced by Figure 5).

(ii) For short conformers, the already electrostatically favorable LysH\(^{+}\) termination\(^{40,43}\) is additionally helpful by allowing soft, delocalized modes to include the termination also for short conformers, in contrast to the hypothetical, much harder charged termination by Li\(^{+}\). For long enough
helices these softer low frequency modes should exist regardless of the termination.

Conclusions

In summary, we show from first principles, quantitatively, and for a particularly well studied series of polyalanine peptides how helices emerge with length and temperature as the leading structural pattern from a vast array of possible competing conformers. The crossover to helical stability with length is already apparent based on local structural minima of the potential-energy surface alone, due to the critical role of H-bond networks including their cooperativity as well as that of vdw terms. In addition, the contribution from softer low-frequency vibrational modes acts to stabilize helices with increasing temperature over their more compact competition. The specific experimental claim of exclusively helical conformers at and above n=8 of Ac-Ala\textsubscript{n}-LysH\textsuperscript{+} is thus explained by both enthalpic and entropic effects acting together at finite temperature. Ab initio molecular dynamics simulations corresponding to approximately room temperature suggest that these trends are further strengthened by anharmonic effects.

The emergence of room-temperature helix stability with length in Ac-Ala\textsubscript{n}-LysH\textsuperscript{+} is thus the result of a subtle balance of enthalpic and entropic terms. In a non-vacuum environment, further terms would obviously contribute, but we expect the fact that helices, in general, allow locally softer vibrational modes to hold. Regarding the overall ubiquity of the helical motif in folded peptides and proteins, at least, we here show that low-frequency modes will be a significant quantitative contribution.

Acknowledgement

The authors would like to acknowledge Dr. Carsten Baldauf for numerous helpful discussions and suggestions about the figures and paper in general, including the schematic backbone vibration visualization in Figure 8.
Figure 8: Characterization of the lower vibrational modes of helices considered in this work. (a) Positions of the first and second vibrational modes of $\alpha$-helical geometries of Ala$_n$-Li$^+$, $n$=5-15, and 20 (red) and $\alpha$-helical geometries of Ac-Ala$_n$-LysH$^+$, $n$=5-15, and 19 (black). (b) Displacement caused on the backbone atoms of Ac-Ala$_{19}$-LysH$^+$ by deforming this molecule in the direction of the first vibrational mode, with red shaded areas corresponding to positive displacements and green shaded areas to negative displacements (hydrogen atoms not shown). (c) Relative change in H-bond distance when displacing 1 (normalized) unit of the first normal mode for Ac-Ala$_{19}$-LysH$^+$. We number the carbonyl groups from the N-terminus to the C-terminus. (d) Same as (b) for Ala$_{20}$-Li$^+$. (e) Same as (c) for Ala$_{20}$-Li$^+$ (in this case the last three carbonyl groups do not make H-bonds, but are connected to the Li$^+$ ion).
Supporting Information Available

Additional computational details, the XYZ geometries of conformers discussed in this paper, detailed information (H-bond network, Ramachandran plots, relative energies) about the structures discussed in this paper, and movies illustrating the first vibrational modes of helices and compact conformers can be found in the SI. This material is available free of charge via the Internet at http://pubs.acs.org.

References


[71] Ponder, J. *Tinker - software tools for molecular design.* In this work we used versions 4.2 and 5.1 of the program and the force-fields’ versions distributed within the package.


In this manuscript, we show that polypeptide helices are stabilized with respect to more compact conformers by vibrational free energy at finite temperatures. The TOC Figure illustrates this stabilization for the α-helical conformer and a compact conformer of Ac-Ala₈-LysH⁺, by showing potential energy and vibrational free energy (300 K) hierarchies.