Article:
Batchelor, M orcid.org/0000-0001-6338-5698 and Paci, E orcid.org/0000-0002-4891-2768 (2018) Helical Polyampholyte Sequences Have Unique Thermodynamic Properties. Journal of Physical Chemistry B, 122 (49). pp. 11784-11791. ISSN 1520-6106

https://doi.org/10.1021/acs.jpcb.8b08344

Helical Polyampholyte Sequences Have Unique Thermodynamic Properties

Matthew Batchelor and Emanuele Paci*
Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK
Correspondence to e.paci@leeds.ac.uk

Abstract

Helices are the most common structural pattern observed in structured proteins. Polypeptide sequences that form helices in isolation have been identified and extensively studied. These are generally rich in alanine, the amino acid with strongest helical propensity. Insertion of charged or polar amino acids has been shown to be necessary to make alanine-rich peptides soluble, and sometimes even increase the helicity of the peptides. More recently sequences that contain mostly charged residues (E-R/K rich) have been found in naturally occurring proteins that are highly helical, soluble and extended regardless their length. Artificial sequences composed mostly or exclusively of charged amino acids have been designed that are also highly helical, depending on the specific pattern of oppositely charged residues. Here we explore the thermodynamic properties of a number of 16-residue long peptides with varying helical propensity by performing equilibrium simulations in a broad range of temperatures. We observe quantitative differences in the peptides' helical propensities that can be related to qualitative differences in the free energy landscape, depending on the ampholytic patterns in the sequence. The results provide hints on how the specific physical properties of naturally occurring long sequences with similar patterns of charged residues may relate to their biological function.

Introduction

Protein sequences have an average amino acid composition that is relatively well maintained across species, with amino acids with a charged side chain (D, E, K, R) occurring with a frequency of about ~23%. Sequences unusually rich in charged amino acids generally assume expanded conformations lacking secondary or tertiary structure. In fact, 75% of intrinsically disordered proteins (IDPs) are strong polyampholytes, *i.e*., polypeptides containing a significant proportion (>35%) of both positive and negatively charged residues.[1-2] However, sequences that contain up to 80% charged amino acids have been found to constitute domains in some large proteins and form stable helices with intriguing properties.[3-4]

Alpha helices are the most prevalent and arguably most well studied protein secondary structure element. When removed from the protein context, and any stabilizing tertiary contacts, some peptides with specific sequence patterns are able to maintain significant helicity. Alanine-rich sequences that contain either isolated or sparsely distributed interacting polar residues are helical. Examples include $(AAAAX)_n$ with X = E or K,[5] R,[6] or Q,[7-8] with helicity increasing with *n*. For polyalanine-based peptides the helicity is thought to stem from the intrinsic helix-compatibility (or helical propensity) of alanine. More recently discovered helices with sequences that are made up almost entirely of amino acids with charged side chains are what we here term strong polyampholyte helices. Such peptides, besides containing large proportions of both negatively (specifically E) and positively charged (K, R) amino acids, have oppositely charged residues grouped in regular, repeating patterns. Naturally occurring polyampholyte helical sequences, have variable sequence patterns that roughly match an alternating pattern of blocks of ~3-4 same-charge residues. Ion pair interactions between side chains across turns of the helix are thought to contribute to their stability. The stabilizing influence of this E–K/R repeating pattern has been established from observations using short peptide isomers of E_8K_8: $E_4K_4E_4K_4$ was shown to be highly helical whereas $(E_2K_2)_4$ was ahelical.[9] There is evidence that some of these naturally occurring sequences form continuous monomeric helices in solution, over a wide range of pH and salt concentrations, even at lengths of over 100 residues.[10-13]

In this paper we draw comparisons between these different peptide sequences and, on the face of it, their identical resulting helical structures, the manner and mechanism for helix promotion and the different thermodynamic properties that they give rise to. We also explore the helical properties of mixed polyampholyte/polyalanine sequences that lie between the two extremes, and the larger region of polyampholyte sequence-space that is non-helical and generally prefers extended disordered conformational ensembles. Some intriguing hypotheses on the origin of the extraordinary properties of helical polyampholyte sequences have been previously formulated, based on thermodynamic and kinetic properties of simulations of computational models.[14-16]

Computational models are limited by the quality of force fields. Traditional force fields were developed to stabilize known protein structures. Most commonly used force fields have been recently modified to remove a bias towards excessive helical stability,[17] and novel force fields have been produced to reproduce the experimental properties of unstructured polypeptide sequences.[18-19] We previously reported[14-15, 20] results from simulations using a range of force fields. Simulations showed that long naturally occurring helical polyampholytes assume extended helical conformations while alanine-rich ones assume helical bundled conformations.[14] Concerning the origin of the stability of extended helices, simulations suggest that for polyampholyte

sequences the loss in entropy upon helical formation is much lower that for "normal" (*i.e.*, alanine-rich) sequences, and collapse of long helices (>20 amino acids) is not required to minimize their free energy. However, different force fields can provide different origins for such high entropy of the helical state: some show interconversion between α and π helices (CHARMM19/FACTS[21] or GROMOS[22]), and others show a dynamic network of salt bridges on the outside of the α helix (CHARMM36[17, 20] or AMBER2003[16]). In simulations with different force fields we observed a peculiar response to mechanical force, namely the application of a force does not increase the rate of unfolding as for structured proteins, but causes a progressive elongation to which the helix responds with an approximately constant recoiling force.[14-15]

Here we report thermodynamic properties estimated using the OPLS/ABSINTH[19] force-field and replica-exchange Monte Carlo simulations for a wide range of peptide sequences, each sixteen residues in length, and linked, where possible, to ensemble average properties known from experimental studies.

Methods

The thermodynamic properties of the peptides were determined using OPLS/ABSINTH[19] implicit solvation model and replica exchange Monte Carlo simulations using CAMPARI.[23] ABSINTH is recognized as one of the most reliable force fields for the investigation of disordered peptides as it does not overestimate compactness[24] and extent of secondary structure.[19] Replica-exchange simulations were performed at 20 different temperatures in the range 250–600 K. All simulations were started from random conformations, equilibrated for 10^6 Monte Carlo steps and were continued for at least 10^8 further steps or until the radius of gyration and helicity did not vary more than 1% when the simulation was extended by 10%. (The convergence of the helicity along the simulation in one specific case is demonstrated in Figure S1.) Simulations were performed within a sphere 40 Å in radius and ions (Na^+, Cl^-) were explicitly added to neutralize the system when charged side chains were present. N- and C- termini were acetylated and methylated, respectively. Simulation trajectories were analyzed using Wordom.[25] The secondary structure of residues within each peptide was assigned for each timeframe using DSSPcont[26] criteria. This was then used to calculate the helicity of the peptide overall. Helicity was defined as the fraction of residues assigned as 3_{10} (G), α (H), or π (I) helix, but 3_{10} and π content was negligible in all cases.[27] Structures were clustered based on DRMS values, using a cutoff value of 2 Å between clusters.[25] A salt bridge was considered to be present if the distance between any of the oxygen atoms of acidic residues and the nitrogen atoms of basic residues was less than 4 Å.

Results

The maximum helicity of each peptide, *i.e.*, the one observed at the lowest temperature of 250 K, and the helicity at 290 K are given for each of the 16-residue peptides studied in Table 1. The list includes results for some homopolypeptides (A_{16}, G_{16}, K_{16}, E_{16}, D_{16}, R_{16}). A_{16} is highly helical, alanine being the amino acid with strongest helical propensity.[28] Experimentally, limited insertion of charged or polar amino acids is necessary to make polyalanine peptides soluble.[5] The overall helicity of A_{16} was not affected by substitution of three A for K or Q. K_{16} is moderately helical, while D_{16} and G_{16}, which contain residues with low helical propensity,[28] do not form helices in the simulations at any temperature. E and R, despite being residues often found in helices, exhibit negligible helicity as homopolypeptides.

The fraction of helical content as a function of temperature is shown in Figure 1 for some of the peptides listed in Table 1. At the lowest temperature (250 K) A_{16}, $(E_4K_4)_2$ and $(E_4R_4)_2$ are highly helical (~80%). The polyampholyte sequences lose helicity faster with increasing temperature than A_{16}, with $(E_4R_4)_2$ having a higher melting temperature than $(E_4K_4)_2$. Replacement of Glu by Asp significantly reduces the helicity of the E_4K_4 or E_4R_4 motifs (Table 1), as expected.[29] The result for E_8K_8 is intriguing, being at most 60% helical but highly stable. Structural analysis shows that its C-terminal part forms a particularly stable helix while the N-terminus wraps around the helix, with E residues in positions 3–6 making persistent salt bridges with K residues in positions 9–10 (Figure S2). Representative structures and individual residue helicities for E_8K_8 and other select sequences and temperatures are given in Figure S3–S8.

Thermal curves of the end-to-end distance (r_{NC}) and the radius of gyration (R_{gyr}) are shown in Figure 2. At the lowest temperature, all sequences with helical propensity form extended helices with a length around 24 Å (*i.e.*, the length of an extended 16-residue helix). A_{16} becomes shorter on average and more compact as temperature increases and helicity is lost before expanding at even higher temperatures. Helical polyampholyte sequences ($(E_4K_4)_2$ and $(E_4R_4)_2$) sample longer and less compact conformations as temperature increases and helicity decreases. The r_{NC} and R_{gyr} converge to a value that depends not only on the amino acid composition of the sequence and the excluded volume of the amino acids, but also on the arrangement of the amino acids. E_8K_8 remains relatively compact up to the highest temperature reported in Figure 2 (600 K). The salt bridges mentioned above that stabilize the C-terminal helical region persist even at temperatures at which the helicity is completely lost.

The distribution of end-to-end distances and radii of gyration for three of the peptides with strong helical propensity (A_{16}, $(E_4K_4)_2$ and $(E_4R_4)_2$) are shown in Figure 3. At the lowest temperature (Figure 3A and 3D) both are narrowly distributed around values typical of a 16-residue extended helix (r_{NC} ~ 24 Å and R_{gyr} ~ 7.4 Å). Polyalanine and polyampholyte sequences differ, however, in that more compact structures are sampled for A_{16}, while only less compact conformations are sampled for polyampholytes. At temperatures where the sequences are ~40% helical (Figure 3B and 3E), ~65% of A_{16} conformations sampled have R_{gyr} < 7.4 Å, while ~99% of $(E_4K_4)_2$ or $(E_4R_4)_2$ conformations have R_{gyr} > 7.4 Å. In other words, while for polyalanine the helical state is the least compact, for polyampholyte sequences it is the most compact. At a temperature of 380 K when helicity drops below 5% for all peptides, the distribution of R_{gyr} broadens for A_{16}, while it moves towards larger R_{gyr} for the two polyampholyte sequences (Figure 3C and 3F).

Similar helicity between peptides means that the number of amino acids estimated to be in a helical conformation on average is approximately the same, which is precisely the case for A_{16} and $(E_4R_4)_2$ at 250 K. Analysis of the occurrence of specific patterns of secondary structure shows remarkable differences. The frequency of occurrence of conformations with a given number of residues in a helical conformation is shown in Figure 4A. For A_{16} conformers containing helix, there are most likely to be 10 residues in a helical conformation, with the likelihood decreasing in a regular manner down to a minimum of 4 and up to a maximum of 15. For the two helical polyampholytes we observe an intriguing pattern where residues in helical conformations seem to occur most likely in multiples of ~5. Also interesting is the fact that conformations with no residues in a helical conformation are twice as likely for A_{16} than for $(E_4K_4)_2$ or $(E_4R_4)_2$. The frequencies of occurrence of contiguous stretches of residues in a helical conformation are shown in Figure 4B. The variation in frequency from 4 to 15 residues is again relatively smooth for A_{16}. For polyampholyte

helical peptides the length of helices is mostly 4–6 amino acids, which suggests helices do not propagate beyond but exist independently (*i.e.*, appear as independent blocks that join up intermittently). The potential of mean force associated with the helix hydrogen bond distance (Figure 4C) complements the results in Figures 4A and 4B. At a temperature at which they are ~40% helical, all three peptides show a minimum at ~4 Å, resulting from hydrogen bonding with an ideal alpha helical value. Polyampholyte peptides have two additional local minima analogous to solvent separated and fully solvated states for ion pairs in solution.[30]

It is also of interest to explore how systematic replacement with alanine of charged residue pairs in helical polyampholytes affects the properties of these peptides. Starting from the E_4K_4 pattern, sequences were modified by exchanging E_i–K_{i+4} pairs by A_i–A_{i+4} until the sequence was composed solely of alanine. The stability of the helical state is maximal (Figure 5A), higher than that of A_{16}, when two charged residue pairs are replaced by alanines. If one or three charged-residue pairs are changed to Ala then the thermal stability is as high as for A_{16}. The radius of gyration as a function of temperature (Figure 5B) shows that the size increases with the content of charged amino acids. On comparison of peptides at different temperatures but when their helicities are all about 40%, the population of a collapsed state decreases with increasing content of charged amino acids. The pattern $(E_2A_2K_2A_2)_2$ appears to have sufficient hydrophobicity to compensate for the effect of charges towards expanded conformations. A significant component of helix hairpin-like structures (Figure S6) were observed giving rise to a small r_{NC} at low temperature.

The helical polyampholyte sequences ($(E_4K_4)_2$ and $(E_4R_4)_2$) are compatible with the formation of E–K or E–R salt bridges across helical turns. One factor previously highlighted as a possible reason of the high stability of sequences with specific patterns of oppositely charged residues is that of the multiplicity of ion pairs possible, which, it was suggested, does not imply a large loss in entropy upon formation of a helix. The occurrence of salt bridges for the two polyampholyte sequences $(E_4K_4)_2$ and $(E_4R_4)_2$ at the temperature where they are most helical (~80%, 250 K) is shown in Figure 6. Interestingly the salt bridge that occurs most often is only present in 9% of the conformations sampled for $(E_4K_4)_2$ and 3% for $(E_4R_4)_2$. All salt bridges detected are less populated for the more helical $(E_4R_4)_2$. This suggests that formation of more salt bridges may not be the major contributor to a more stable helix for these polyampholyte motifs.

As part of a previous study,[15] we modeled some related but shorter (10 residue) polyampholyte sequences using the CHARMM19 force field and FACTS implicit solvation, and compared with helicities available from the literature.[31] Those simulations did not show the differences in helicity between the E_4K_4 and E_4R_4 motifs observed in experimental studies.[31] We repeated these simulations using the OPLS/ABSINTH model and found much better agreement (Figure S9). The ABSINTH model employed here definitely reproduces a higher helical propensity in helical polyampholyte sequences when Lys is replaced by Arg. Similarly, in polyalanine-based peptides, replacement of an (*i*, *i* + 4) Glu–Lys pairing with (*i*, *i* + 4) Glu–Arg has been reported to increase helicity.[32] It should also be noted that the presence of an N-terminal YS tag appears to decrease the stability of helical conformations. Attempts were made to model the effects of increasing salt concentration on the helical polyampholytes. Moderate concentration of salt (NaCl) affect helicity moderately, in agreement with the experiment. [9] Increasing salt content to 0.9 M significantly increased the helicity of both $(E_4K_4)_2$ and $(E_4R_4)_2$ (Figure S10); this is contrary to the expectation that higher salt levels would reduce helical content.[9] Further work is required to explore the origin of this disagreement; possibly

in the way the model treats ions explicitly in the context of an implicit aqueous environment.

Discussion

We have presented results from simulations of 16-residue long amino acids to extend our understanding of the determinants of the helical propensity of patterned, strong polyampholyte sequences. The force field employed here, together with fully converged replica-exchange Monte Carlo simulations, provides some insights, while providing some qualitative agreement with experimental results. Indeed, the conclusions below only hold to the extent that the models used accurately reflect the real systems, but some observation may be of general interest. The model reflects the natural propensity of certain amino acids to form helical structure: not only polyalanine but also polyampholyte sequences with specific charge patterns are strongly helical. While the model appears to overestimate the helicity relative to experiment at the same temperature, by considering the temperature as a tuner of the intensity of the interactions, a small rescaling of the energy would provide fuller agreement with experiment. What is certainly in agreement is that the sequences A_{16}, $(E_4K_4)_2$ and $(E_4R_4)_2$ have a remarkable helical propensity. Also, in agreement with circular dichroism measurements, sequences with different patterns of oppositely charged pairs E–K (or E–R), are much less helical or not helical at all.

The similarity of the average helicity of a peptide hides some interesting features. For A_{16} (250 K, ~80% helical), the two most likely structures are those with zero and 10 residues in a helical conformation (Figure 4A); more or less helical structures are populated with a probability that decreases slowly and monotonically. Comparing the distributions of radius of gyration, one can identify two co-existing but thermodynamically distinct states, one compact and perfectly non-helical, and one that contains between 4 and 15 residues in a helical conformation and that has on average the length and radius of gyration of an extended helix.

For the polyampholyte sequences $(E_4K_4)_2$ and $(E_4R_4)_2$, conformations are most likely to have 5, 10 or 15 residues in a helical conformation; continuous helices are between 4 and 6 residues long and rarely longer. What appears clearer as temperature increases, is that conformers are favored that are considerably more expanded than the canonical helix conformation but still partly helical (most likely with 5 or 10 residues in a helical conformation). In other words, there appear to be two states, one fully helical, the most compact, and one that is expanded but with considerable helical structure. If radius of gyration and helicity are taken as meaningful reaction coordinates, the two states appear to be separated by a smaller free energy barrier. In other words, similarly helical peptides (alanine-rich and strong polyampholytes) appear to have considerably different free energy landscapes.

In a study that exploited the same ABSINTH implicit solvation model with a focus on polyampholyte peptide sequences,[1] the conformational propensity of charge-neutral sequences made only of E and K was characterized as a function of a patterning parameter that is lowest for a well-mixed sequence (*e.g.*, $(EK)_8$) and highest for sequences where same-charge residues are grouped together (*e.g.*, E_8K_8). The sequence specific distribution of oppositely charged residues was shown to be a crucial determinant of the conformational properties of polyampholytes together with the fraction of charged residues. The results presented here for shorter peptides (16 residues instead of 50 described by Das and Pappu[1]) similarly show that well-mixed

sequences behave like Flory random coils at all temperatures, while segregation of charges favor more compact conformations. This is evident from Figure 2 at temperatures below 400 K (the average helicity over a broader range of temperatures is shown in Figure S11). Das and Pappu[1] reported the propensity of sequences with long stretches of identical charges to form more compact hairpin-like structures. For the shorter sequences analyzed here this is the case for E_8K_8 where the negatively charged N-terminal region wraps back onto the strongly helical positively charged C-terminal half of the chain, and long-range salt bridges contribute to its remarkable stability (Figure S2 and S3).

What was not previously reported is that specific patterns of oppositely charged amino acids impart a strong helical propensity to polyampholyte peptide sequences. Sequences with a pattern of four negatively charged Glu residues, followed by four positively charged residues (Lys or Arg) have helical propensity, which is stronger for the Arg-containing motif than for Lys. This agrees with experimental CD measurements of natural and synthetic polyampholyte peptides. The E_4K_4 or E_4R_4 motifs have persistent blocks of helix: at temperatures with the same overall helicity, they show more conformers with some helical content than polyalanine does (or polyalanine exhibits more completely ahelical conformers, Fig. 4A). In conformers that contain non-helical residues, these sections are all extended/expanded, so the tendency is always to maintain directionality and avoid collapse. E_4K_4 or E_4R_4 motifs allow multiple alternative salt bridges patterns; specific salt bridges are not highly populated, and the stability of the helical state appears to be in large part due to an entropic contribution for only weakly constrained side chain conformations (rotamer flexibility).

The properties summarized above provide some explanation of the thermodynamic factors that make it possible for naturally occurring long sequences with similar patterns of charged residues to assume an extended but soluble helical conformation. They also provide novel information that may relate to their biological function. Perturbations such as increased temperature or exposure to mechanical force[14] lead to local loss of helicity but not to collapse. This also means that the helical structure can reform faster, assisted, if not driven, by electrostatic interactions, providing a sort of self-healing mechanism that allows them to act as a robust ruler or spacer between functional domains.

In conclusion, we have described the thermodynamic properties, over a range of temperatures, for model peptides that sample the sequence space between strong polyampholyte and poly-alanine sequences. Despite similar helical propensity, these sequences differ in their thermodynamic and structural features.

Supplementary Information
In Supplementary Information are shown 11 supplementary figures; supplementary methods include a sample input file for CAMPARI.

Acknowledgments
We thank Andreas Vitalis for insightful discussions and support with CAMPARI (http://campari.sourceforge.net). This work was partly supported by BBSRC (BB/I007423/1).

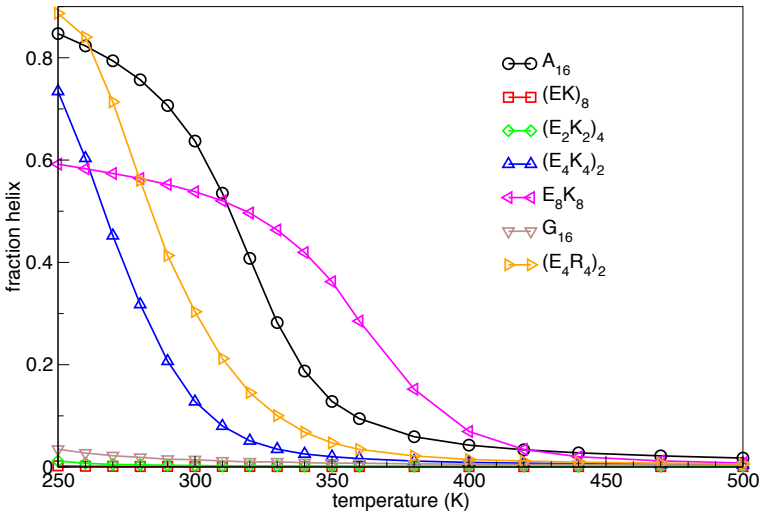

Figure 1. Helicity. At the lowest temperature A_{16}, $(E_4K_4)_2$ and $(E_4R_4)_2$ are highly helical. Polyampholyte sequences lose helicity faster with increasing temperature, and $(E_4R_4)_2$ is at all temperatures more helical than $(E_4K_4)_2$. E_8K_8 is mostly helical in its C-terminal part and loses helicity only at high temperature.

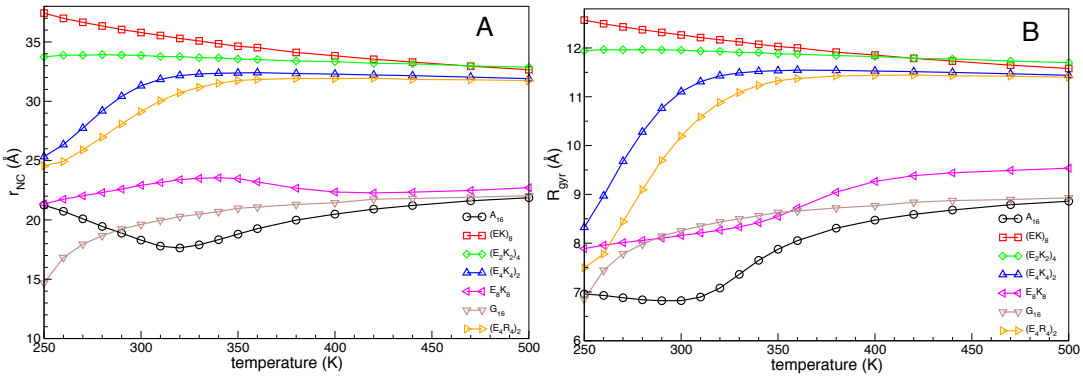

Figure 2. (A) Average end-to-end distance and (B) radius of gyration at a function of temperature. At the lowest temperature, all sequences with helical propensity are extended with a length around 24 Å (length of a perfect 16 amino acid helix). A_{16} gets shorter when it loses helicity (hydrophobic collapse) before expanding at high temperature. Helical polyampholyte sequences ($(E_4K_4)_2$ and $(E_4R_4)_2$) elongate as they lose helicity.

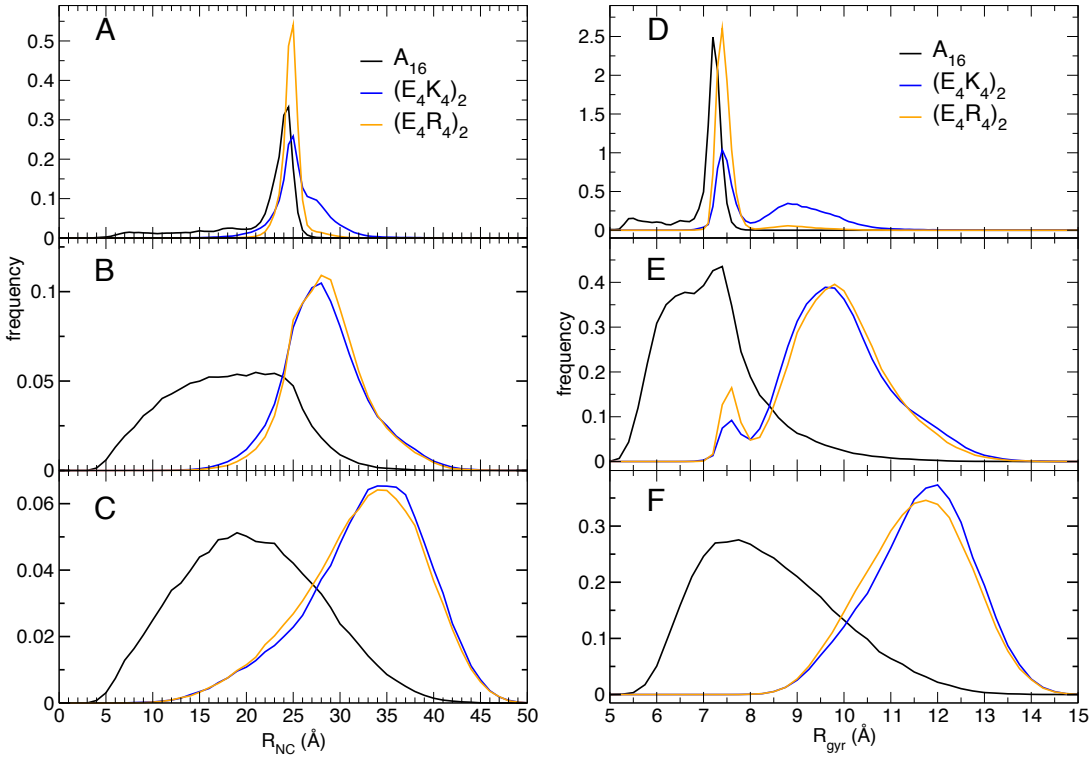

Figure 3. Histograms of r_{NC} (left) and R_{gyr} (right). (A,D) At the lowest temperature (250 K). (B,E) At a temperature in which all peptides are ~40% helical. (C,F) At 380 K when all peptide are <5% helical.

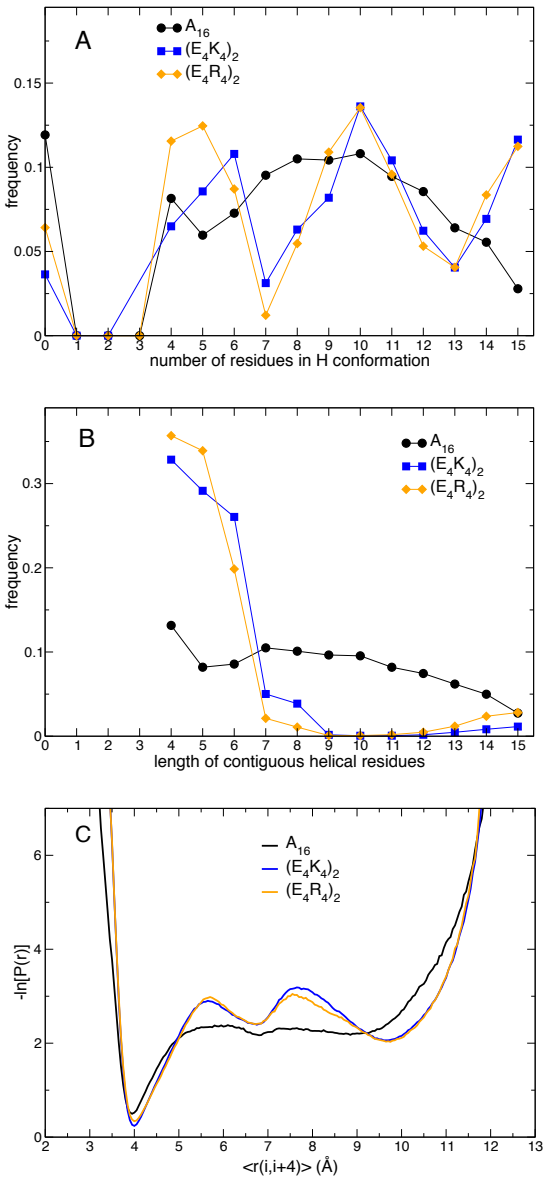

Figure 4. Different nature of helices. (A) Frequency of conformations with a given number of residues in a helical conformation at the lowest temperature (250 K). For A_{16} the maximum frequency is ~10 but any number is similarly likely between the minimum (4) and the maximum (15). For the two helical polyampholytes we observe an intriguing pattern where residues in helical conformations seem to occur most likely in multiples of ~5. (B) Frequencies of occurrence of contiguous stretches of residues in helical conformation. This is again relatively smooth for A_{16}. For polyampholyte helical peptides the length of helices is mostly 4–6 amino acids, which suggests helices do not propagate beyond that length or merge but exist independently. (C) Negative logarithm of the probability of the helix hydrogen bond distance. This is the distance between the amide nitrogen of an amino acid and the carbonyl carbon of the amino acid four residues earlier, averaged over all possible pairs of residues. Profiles are shown for temperatures at which peptides are ~40% helical.

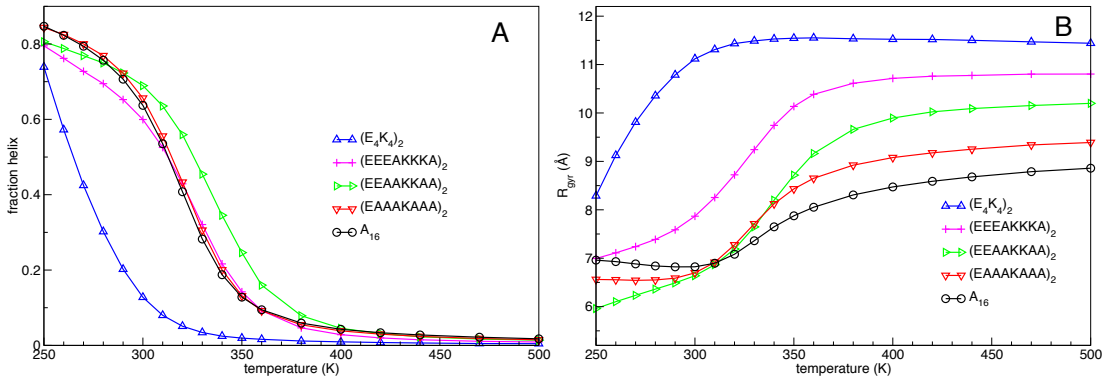

Figure 5. Helicity (A) and radius of gyration (B) of peptides where, starting from $(E_4K_4)_2$, alanines are introduced preserving the 4-4 pattern of charged residues. The peptide with the pattern EEAAKKAA appears to have the highest melting temperature.

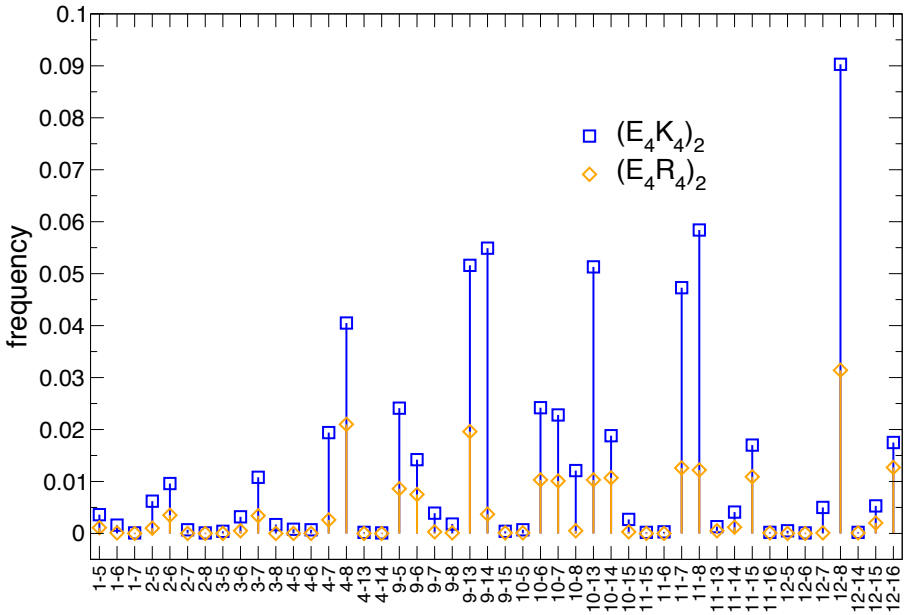

Figure 6. Frequency of salt bridges at 250 K for two polyampholyte helical sequences. Salt bridges are more frequent for $(E_4K_4)_2$ than for $(E_4R_4)_2$ despite the latter having larger helical content. Strength of salt bridges does not appear to correlate with helicity. Residues numbered 1–4 and 9–12 are E, those numbered 5–8 and 13–16 are K or R.

Table 1. List of 16mers studied and average helicity at 250 and 290 K. In parentheses are the standard errors estimated by averaging over contiguous blocks of 10^7 conformations.

Peptide (16mers)	Fraction of residues in helical conformation		Experimental data if available (% helix)
	Simulation (250 K)	Simulation (290 K)	
A_{16}	0.849 (0.004)	0.708 (0.007)	Not soluble.[7]
$(AAAAK)_3A$	0.875 (0.006)	0.81 (0.01)	~72% at 274 K.[5]
$(AAAAQ)_3A$	0.83 (0.01)	0.71 (0.02)	$(AAQAA)_3Y$ ~40% 274 K, ~20% 300 K.[7-8]

G_{16}	0.04 (0.02)	0.02 (0.01)	Not soluble (unstructured).[33]
$(E_4K_4)_2$	0.74 (0.02)	0.20 (0.02)	~35% 278 K.[4] ~65% at 278 K (YS- or G-tag).[3, 9]
$(K_4E_4)_2$	0.01 (0.01)	0.002 (0.003)	22% at 278 K (G-tag).[3]
$(E_4R_4)_2$	0.891 (0.009)	0.44 (0.03)	
$(R_4E_4)_2$	0.02 (0.02)	0.003 (0.004)	
$(D_4K_4)_2$	0.28 (0.02)	0.06 (0.01)	
$(D_4R_4)_2$	0.36 (0.02)	0.21 (0.02)	
$(EK)_8$	0.002 (0.003)	0.001 (0.001)	$G(EK)_{12}G$, 22% at 278 K.[3]
$(E_2K_2)_4$	0.012 (0.008)	0.003 (0.003)	Not helical.[3, 9]
E_8K_8	0.592 (0.006)	0.55 (0.01)	
$(EA_3KA_3)_2$	0.84 (0.01)	0.72 (0.02)	
$(E_2A_2K_2A_2)_2$	0.81 (0.01)	0.73 (0.01)	
$(E_3AK_3A)_2$	0.80 (0.02)	0.67 (0.04)	
$(A_4K_4)_2$	0.875 (0.005)	0.839 (0.008)	
$(E_4A_4)_2$	0.39 (0.02)	0.30 (0.01)	
K_{16}	0.25 (0.04)	0.011 (0.006)	polyK not helical at neutral pH.[34]
R_{16}	0.01 (0.02)	0.001 (0.004)	polyR not helical at neutral pH.[35]
E_{16}	0.000 (0.001)	0.000 (0.001)	polyE not helical at neutral pH.[36]
D_{16}	0.000 (0.002)	0.000 (0.002)	

References

1. Das, R. K.; Pappu, R. V., Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (33), 13392-7.
2. Sickmeier, M.; Hamilton, J. A.; LeGall, T.; Vacic, V.; Cortese, M. S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V. N.; Obradovic, Z.; Dunker, A. K., DisProt: the Database of Disordered Proteins. *Nucleic acids research* **2007,** *35* (Database issue), D786-93.
3. Baker, E. G.; Bartlett, G. J.; Crump, M. P.; Sessions, R. B.; Linden, N.; Faul, C. F.; Woolfson, D. N., Local and macroscopic electrostatic interactions in single alpha-helices. *Nat Chem Biol* **2015,** *11* (3), 221-8.
4. Lyu, P. C.; Marky, L. A.; Kallenbach, N. R., The Role of Ion-Pairs in Alpha-Helix Stability - 2 New Designed Helical Peptides. *Journal of the American Chemical Society* **1989,** *111* (7), 2733-2734.
5. Marqusee, S.; Robbins, V. H.; Baldwin, R. L., Unusually stable helix formation in short alanine-based peptides. *Proceedings of the National Academy of Sciences of the United States of America* **1989,** *86* (14), 5286-90.
6. Lockhart, D. J.; Kim, P. S., Internal stark effect measurement of the electric field at the amino terminus of an alpha helix. *Science* **1992,** *257* (5072), 947-51.
7. Scholtz, J. M.; York, E. J.; Stewart, J. M.; Baldwin, R. L., A neutral, water-soluble, .alpha.-helical peptide: the effect of ionic strength on the helix-coil equilibrium. *Journal of the American Chemical Society* **1991,** *113* (13), 5102-5104.
8. Shalongo, W.; Dugad, L.; Stellwagen, E., Distribution of Helicity within the Model Peptide Acetyl(Aaqaa)(3)Amide. *Journal of the American Chemical Society* **1994,** *116* (18), 8288-8293.
9. Lyu, P. C.; Gans, P. J.; Kallenbach, N. R., Energetic contribution of solvent-exposed ion pairs to alpha-helix structure. *Journal of molecular biology* **1992,** *223* (1), 343-50.
10. Baboolal, T. G.; Sakamoto, T.; Forgacs, E.; White, H. D.; Jackson, S. M.; Takagi, Y.; Farrow, R. E.; Molloy, J. E.; Knight, P. J.; Sellers, J. R.; Peckham, M., The SAH domain extends the functional length of the myosin lever. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (52), 22193-8.
11. Knight, P. J.; Thirumurugan, K.; Xu, Y.; Wang, F.; Kalverda, A. P.; Stafford, W. F., 3rd; Sellers, J. R.; Peckham, M., The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. *The Journal of biological chemistry* **2005,** *280* (41), 34702-8.
12. Spink, B. J.; Sivaramakrishnan, S.; Lipfert, J.; Doniach, S.; Spudich, J. A., Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. *Nature structural & molecular biology* **2008,** *15* (6), 591-7.
13. Wang, C. L.; Chalovich, J. M.; Graceffa, P.; Lu, R. C.; Mabuchi, K.; Stafford, W. F., A long helix from the central region of smooth muscle caldesmon. *The Journal of biological chemistry* **1991,** *266* (21), 13958-63.
14. Wolny, M.; Batchelor, M.; Knight, P. J.; Paci, E.; Dougan, L.; Peckham, M., Stable Single alpha-Helices Are Constant Force Springs in Proteins. *The Journal of biological chemistry* **2014,** *289* (40), 27825-35.
15. Batchelor, M.; Gowdy, J.; Paci, E., Effect of external pulling forces on the length distribution of peptides. *Biochimica et biophysica acta* **2015,** *1850* (5), 903-10.
16. Sivaramakrishnan, S.; Spink, B. J.; Sim, A. Y.; Doniach, S.; Spudich, J. A., Dynamic charge interactions create surprising rigidity in the ER/K alpha-helical protein motif. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (36), 13356-61.

17. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* **2012,** *8* (9), 3257-3273.
18. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; MacKerell, A. D., Jr., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017,** *14* (1), 71-73.
19. Vitalis, A.; Pappu, R. V., ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J Comput Chem* **2009,** *30* (5), 673-99.
20. Wolny, M.; Batchelor, M.; Bartlett, G. J.; Baker, E. G.; Kurzawa, M.; Knight, P. J.; Dougan, L.; Woolfson, D. N.; Paci, E.; Peckham, M., Characterization of long and stable de novo single alpha-helix domains provides novel insight into their stability. *Sci Rep* **2017,** *7*, 44341.
21. Haberthur, U.; Caflisch, A., FACTS: Fast analytical continuum treatment of solvation. *Journal of Computational Chemistry* **2008,** *29* (5), 701-715.
22. Shepherd, C. M.; van der Spoel, D.; Vogel, H. J., Molecular dynamics simulations of peptides from the central domain of smooth muscle caldesmon. *Journal of biomolecular structure & dynamics* **2004,** *21* (4), 555-66.
23. Vitalis, A.; Pappu, R. V., Methods for Monte Carlo simulations of biomacromolecules. *Annu Rep Comput Chem* **2009,** *5*, 49-76.
24. Wuttke, R.; Hofmann, H.; Nettels, D.; Borgia, M. B.; Mittal, J.; Best, R. B.; Schuler, B., Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2014,** *111* (14), 5213-8.
25. Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A., Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics* **2007,** *23* (19), 2625-7.
26. Carter, P.; Andersen, C. A. F.; Rost, B., DSSPcont: continuous secondary structure assignments for proteins. *Nucleic acids research* **2003,** *31* (13), 3293-3295.
27. Zagrovic, B.; Jayachandran, G.; Millett, I. S.; Doniach, S.; Pande, V. S., How large is an alpha-helix? Studies of the radii of gyration of helical peptides by small-angle X-ray scattering and molecular dynamics. *Journal of molecular biology* **2005,** *353* (2), 232-241.
28. Pace, C. N.; Scholtz, J. M., A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical journal* **1998,** *75* (1), 422-7.
29. Huyghues-Despointes, B. M.; Scholtz, J. M.; Baldwin, R. L., Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein science : a publication of the Protein Society* **1993,** *2* (1), 80-5.
30. Sadek, H.; Fuoss, R. M., Electrolyte-solvent Interaction. IV. Tetrabutylammonium Bromide in Methanol-Carbon Tetrachloride and Methanol-Heptane Mixtures. *Journal of the American Chemical Society* **1954,** *76* (23), 5897-5901.
31. Sommese, R. F.; Sivaramakrishnan, S.; Baldwin, R. L.; Spudich, J. A., Helicity of short E-R/K peptides. *Protein science : a publication of the Protein Society* **2010,** *19* (10), 2001-5.
32. Merutka, G.; Shalongo, W.; Stellwagen, E., A Model Peptide with Enhanced Helicity. *Biochemistry* **1991,** *30* (17), 4245-4248.
33. Ohnishi, S.; Kamikubo, H.; Onitsuka, M.; Kataoka, M.; Shortle, D., Conformational Preference of Polyglycine in Solution to Elongated Structure. *Journal of the American Chemical Society* **2006,** *128* (50), 16338-16344.
34. Chou, P. Y.; Scheraga, H. A., Calorimetric measurement of enthalpy change in the isothermal helix--coil transition of poly-L-lysine in aqueous solution. *Biopolymers* **1971,** *10* (4), 657-80.
35. Hayakawa, T.; Kondo, Y.; Yamamoto, H., Secondary Structure of Poly-L-arginine and Its Derivatives. *Bulletin of the Chemical Society of Japan* **1969,** *42* (7), 1937-1941.
36. Rialdi, G.; Hermans, J., Calorimetric Heat of the Helix-Coil Transition of Poly-L-glutamic Acid1a. *Journal of the American Chemical Society* **1966,** *88* (24), 5719-5720.

TOC graphics