BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Effect of Glutamate Side Chain Length on Intrahelical Glutamate–Lysine Ion Pairing Interactions

Richard P. Cheng,*[,†] Wei-Ren Wang,[†] Prashant Girinath,[‡] Po-An Yang,[†] Raheel Ahmad,[‡] Jhe-Hao Li,[†] Pier Hart,[§] Bashkim Kokona,[§] Robert Fairman,[§] Casey Kilpatrick,[‡] and Annmarie Argiros[‡]

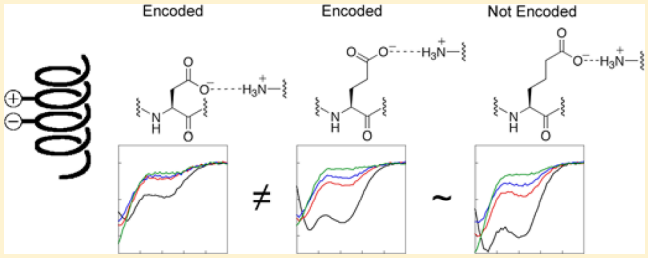
[†]Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan

[‡]Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260-3000, United States

[§]Department of Biology, Haverford College, Haverford, Pennsylvania 19041, United States

Supporting Information

ABSTRACT: Ion pairing interactions between oppositely charged amino acids are important for protein structure stability. Despite the apparent electrostatic nature of these interactions, the charged amino acids Lys, Arg, Glu, and Asp have a different number of hydrophobic methylenes linking the charged functionality to the backbone. To investigate the effect of Glu (and Asp) side chain length on ion pairing interactions, a series of 36 monomeric α-helical peptides containing Zbb-Xaa (i, i+3), (i, i+4), and (i, i+5) (Zbb = Aad, Glu, Asp; Xaa = Lys, Orn, Dab, Dap) sequence patterns were studied by circular dichroism (CD) spectroscopy at pH 7 and 2. Peptides with Glu and Aad exhibited similar helicity and pH dependence, whereas peptides with Asp behaved distinctly different. The side chain interaction energetics were derived from the CD data using the nesting block method coupled with modified Lifson-Roig theory. At pH 7, no Zbb-Xaa (i, i+5) interaction was observed, regardless of side chain length (consistent with the helix geometry). Interestingly, only Lys was capable of supporting Zbb-Xaa (i, i+3) interactions, whereas any Xaa side chain length supported Zbb-Xaa (i, i+4) interactions. In particular, the magnitude of both Zbb^{-}-Lys (i, i+4) and Zbb^{-}-Orn (i, i+4) interaction energies followed the trend Asp > Glu > Aad. Side chain conformational analysis by molecular mechanics calculations showed that the Zbb-Xaa (i, i+3) interactions involved the χ_1 dihedral combination (g+, g+) for the i and i+3 residues, whereas the Zbb-Xaa (i, i+4) interactions were supported by the χ_1 dihedral combination (t, g+) for the i and i+4 residues. These calculated low energy conformers were consistent with conformations of intrahelical Asp-Lys and Glu-Lys salt bridges in a nonredundant protein structure database. These results suggest that Asp and Glu provide natural variation, and lengthening the Glu side chain further to Aad does not furnish additional characteristics that Glu cannot supply.



Most proteins mediate biological functions by relying on a well-defined three-dimensional structure.[1,2] The three-dimensional structure of a protein is predetermined by the primary sequence involving encoded amino acids.[3−5] These encoded amino acids are linked by amide bonds to form a linear polypeptide chain, which folds into the well-defined three-dimensional structure driven by noncovalent forces including electrostatics, hydrogen bonds, hydrophobics, and van der Waals interactions.[5−8] Among these noncovalent forces, electrostatics has gained unexpected attention recently,[9−13] because thermophilic proteins exhibit more stabilizing electrostatic profiles (or charged networks) compared to mesophilic proteins.[14−24] Also, electrostatic interactions can also stabilize collagen triple helices,[25] leading to the design of various stable collagen heterotrimers.[26−28]

The encoded charged residues Asp, Glu, Arg, and Lys, that participate in protein electrostatic interactions, have different side chain lengths (Figure 1). As such, there have been some studies on the effect of altering the side chain length of these charged residues. In particular, the effect of Lys side chain length on helix propensity was investigated,[29] showing decreased helix propensity upon decreasing the Lys side chain length. The effect of Lys side chain length on intrahelical ion pairing interactions has also been investigated.[30,31] Furthermore, increasing the side chain length of Asp to Glu increases helix propensity of the negatively charged amino acid.[32−35] A recent survey on protein salt bridges showed a high preference for interaction between residues separated by fewer than five residues.[13] Also, these local salt bridges were mostly found in helices.[13] High statistical propensity was observed for both Asp and Glu in intrahelical Asp/Glu-Arg (i, i+3) and Lys-Asp/Glu (i, i+4) salt bridges.[13] However, high statistical propensity was observed for only one of the two negatively charged amino acids for intrahelical Glu-Arg (i, i+4), Asp-Lys (i, i+4), and Glu-Lys (i, i+3) salt bridges,[13] implying different

Received: May 20, 2012
Revised: August 15, 2012
Published: August 29, 2012

   dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

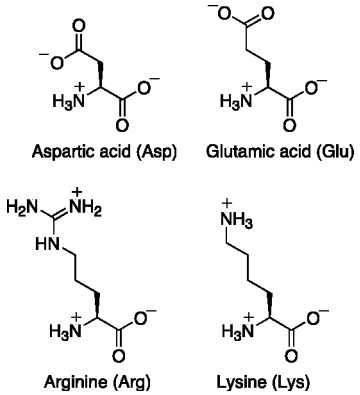


Figure 1. Chemical structure of the natural charged amino acids aspartic acid, glutamic acid, arginine, and lysine.

effects on intrahelical electrostatic interactions upon altering the side chain length of the negatively charged residue.

Experimental studies on Ala-based peptides have provided some information on the difference between Glu and Asp in intrahelical ion pairing interactions. Ala-based peptides with Lys-Glu (i, i+4) and Lys-Asp (i, i+4) sequence patterns exhibited high helical content at pH 7;[36] however, only the helical content of the peptide with Lys-Asp (i, i+4) spacing was significantly attenuated at pH extremes. These results showed that the helicity of peptides with potential intrahelical Lys-Asp and Lys-Glu (i, i+4) interactions exhibited different degrees of pH dependence. In contrast, an Ala-based peptide with the Lys-Glu (i, i+3) sequence pattern showed low helicity at pH 7,[36] with increased helicity at pH extremes. The analogous peptide with the Lys-Asp (i, i+3) sequence pattern exhibited essentially no helicity at pH 7,[36] and small amounts of helicity at pH extremes. Furthermore, the peptide with the Lys-Glu (i, i+3) sequence pattern exhibited higher helicity compared to the peptide with the Lys-Asp (i, i+3) sequence pattern.[36] The higher helicity for Glu-containing peptides compared to Asp-containing peptides was also observed for peptides with potential intrahelical Glu/Asp-Arg and Arg-Glu/Asp interactions involving (i, i+3) and (i, i+4) spacings.[37] Quantitative studies have also been performed to determine the Glu/Asp-Lys (i, i+3) and (i, i+4) intrahelical interaction energetics,[38] showing more stabilizing interactions involving Glu compared to Asp for a given spacing at pH 7 and pH 2.5. This trend is reversed when the orientation of the charged residues is reversed to give Lys-Glu/Asp (i, i+3) and (i, i+4) sequence patterns.[38] These pioneering studies provided a preliminary understanding for the difference between Asp and Glu in forming electrostatic interactions. However, the reason why negatively charged residues Asp and Glu are shorter than the positively charged residues Lys and Arg remains unclear. In other words, the effect of further lengthening the Glu side chain has been unexplored. Accordingly, herein we present a systematic comprehensive study of the effect of Glu side chain length on Glu-Lys intrahelical interactions by circular dichroism spectroscopy, molecular mechanics calculations, and surveys of a nonredundant protein structure database.

MATERIALS AND METHODS

Peptide Synthesis. Peptides were synthesized by solid phase peptide synthesis using Fmoc-based chemistry.[39,40] All peptides were purified by reverse-phase high performance liquid chromatography to greater than 98% purity. The identity of the peptides was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF). Peptides analyzed by sedimentation equilibrium experiments were consistent with monomers. More detailed procedures and peptide characterization data are provided in the Supporting Information.

Circular Dichroism Spectroscopy. CD data was obtained using a 1 mm path length cell. The concentration of peptide stock solutions was determined by the tyrosine absorbance in 6 M guanidinium chloride (ε_{282} = 1220, ε_{280} = 1285, ε_{278} = 1395, ε_{276} = 1455).[41,42] CD measurements were performed at peptide concentrations near 50 μM in 1 mM phosphate, 1 mM citrate, 1 mM borate buffer (pH 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12) at 0 °C. The data were analyzed using Kaleidagraph 3.52 (Synergy Software, CA). Each reported CD value was the mean of at least three determinations. Data are expressed in terms of mean residue molar ellipticity (deg·cm^2·dmol^{-1}) normalized to the number of backbone amide bonds. The mean residue molar ellipticity of the peptides was independent of peptide concentration (50−100 μM). The helical content of two peptides were considered to be significantly different if the CD signal of the less helical peptide was less than 60% of the CD signal for the more helical peptide, and the P-value for comparing the CD signal of the two peptides was less than 10^{-5}. The fraction helix of each peptide (f_{helix}) was derived from the mean residue molar ellipticity at 222 nm and the number of backbone amides (N) using eq 1.[34]

$$f_{helix} = \frac{[\theta]_{222}}{40000\left(1 - \frac{2.5}{N}\right)} \quad (1)$$

Deriving the Interaction Energy Using Modified Lifson-Roig Theory. These calculations were performed following previously published procedures.[31] The modified Lifson-Roig theory[43] that included end-capping effects,[44] and interactions between charged residues and the helix macrodipole[33] was used to derived the fraction helix (f_{helix}) of each peptide in the absence of side chain−side chain interactions.[44] These calculations were performed using compiled computer code written in C++. Literature values for the helix propagation parameter w[29,34,44,45] and the N-terminal capping parameter n[29,34,44] were used for the calculations, and the helix initiation parameter v and the C-terminal capping parameter c for all residues were set to 0.048 [34] and 1,[34] respectively. The contribution of interactions between charged residues and the helix macrodipole was derived based on the method as described by Scholtz.[33] Modifications to the method were made to account for helical states with multiple helical segments.[31]

To obtain the side chain−side chain interaction energetics, the statistical weight for each specific intrahelical side chain-side chain interaction (p) was derived from the experimental f_{helix} based on the nesting block method.[31,33,46] Modifications to the nesting block method were made to explicitly include helix propagation parameters for each amino acid,[29,31,34,44] interactions between charged residues and helix macrodipole,[31,33] and states with multiple helical segments.[31] For every possible state of each peptide, the f_{helix} and probability were calculated and combined. The free energy of each specific side chain−side chain interaction was calculated by $\Delta G = -RT \ln(p)$.[31,33]

Conformational Analysis by Molecular Mechanics. The conformational analysis was performed using the program Discovery Studio 2.1 (Accelrys, CA) on an IBM x3550M2

dx.doi.org/10.1021/bi300655z | Biochemistry 2012, 51, 7157−7172

workstation (CPU: Dual Xeon E5530 2.4 GHz with Quad cores; RAM: 48 G) running the operating system CentOS 5.3. The models were initially created with ideal backbone dihedral angles for α-helix with various combinations of potential low energy side chain dihedrals. Each conformation was then minimized using the CFF forcefield. The nonbond radius of 99 Å, nonbond higher cutoff distance of 98 Å, and nonbond lower cutoff distance of 97 Å were employed to effectively model no cutoffs. Distance dependent dielectric constant of 2 was used as the implicit solvent model. Minimization was performed by steepest descent and conjugate gradient protocols until convergence (converging slope was set to 0.1 kcal/(mol Å)). After minimization, each conformation was re-examined to remove duplicating conformations, because minimization with different starting conformations occasionally resulted in the same final conformation. When the same conformation was represented more than once, only the lowest energy conformation was considered in further analyses.

Survey of Glu-Lys (*i*, *i*+3) and (*i*, *i*+4) Salt Bridges in Protein α-Helices. The survey was performed on the PDBselect (April 2009, 25% threshold),[47,48] a database of nonredundant protein chains. The α-helical conformation was defined based on backbone dihedral angles.[49−52] The helical residues were selected using in-house code written in ActivePerl 5.8.8.819. The backbone dihedral angles were compiled using DSSP.[53] The side chain dihedrals were compiled using Discovery Studio 2.1 (Accelrys, CA, USA). Segments of six or more α-helical residues were considered to avoid end effects.[31,51,52] The occurrence was compiled for Glu/Asp-Lys (*i*, *i*+3), (*i*, *i*+4), and (*i*, *i*+5) sequence patterns. These occurrences were compiled using in-house code written in ActivePerl 5.8.8.819. The helix pair propensity of each residue pattern was calculated by dividing the occurrence of the sequence pattern in α-helices by the expected occurrence for the sequence pattern based on all structures in the database. The pair propensity of each residue pattern was calculated by dividing the occurrence of the sequence pattern in α-helices by the expected occurrence for the sequence pattern based on the α-helices in the database. The expected occurrence and the corresponding standard deviation was obtained by bootstrapping[54] the sequence pattern against the appropriate context across the PDBselect database. The bootstrapping was performed using in-house code written in C++. Dividing the difference between the occurrence and the expected occurrence by the standard deviation gave the *z* value, which was used to obtain the *P* value based on a normal distribution.[55,56] A salt bridge was considered to be present if any of the oxygens of the side chain carboxylate of Asp or Glu was within 3 Å of the nitrogen of the side chain amine of Lys. The occurrence of intrahelical salt bridge was compiled using the program Discovery Studio 2.1 (Accelrys, CA, USA) on a personal computer running the operating system Windows XP (Microsoft, WA, USA).

RESULTS

Peptide Design and Synthesis. The peptides were designed based on monomeric α-helical peptides with oppositely charged residues placed three, four, or five residues apart.[31,57] The peptides were named according to the three-letter code of the negatively charged amino acid (Zbb = Asp, Glu, Aad), the positively charged residue (Xaa = Dap, Dab, Orn, Lys), and the spacing between the two charged amino acids (Figure 2). Since an ideal α-helix would have 3.6 residues per turn, the ZbbXaa3 and ZbbXaa4 peptides could potentially have interactions between oppositely charged residues spaced three and four residues apart, respectively (Figure 2). The ZbbXaa5 peptides have oppositely charged residues spaced five residues apart (Figure 2), which should not interact based on the α-helix geometry and as shown by Scholtz and co-workers.[38] The negatively charged residue Zbb was systematically lengthened from Asp (one methylene) to Glu (two methylenes), and then to (*S*)-2-aminoadipic acid (Aad, three methylenes). To investigate the possibility of length complementarity between the negatively charged and positively charged amino acids, the positively charged residue Xaa was systematically shortened from Lys (four methylenes) to Orn (three methylenes), (*S*)-2,4-diaminobutyric acid (Dab, two methylenes), and (*S*)-2,3-diaminopropionic acid (Dap, one methylene). The overall distribution of the charged residues was designed to stabilize the α-helix macrodipole by placing the negatively charged residue Zbb closer to the N-terminus compared to the positively charged residue Xaa.[31,57] Tyr was incorporated to facilitate concentration determination by UV−vis,[41,42] and the Gly-Gly intervening sequence was included to minimize interference in the circular dichroism (CD) signal by the Tyr chromophore.[58]

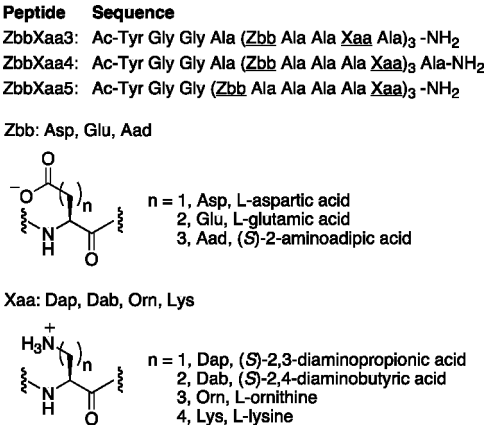


Figure 2. Sequences for the ZbbXaa3, ZbbXaa4, and ZbbXaa5 peptides. The underlined generic 3-letter codes Zbb and Xaa represent the negatively and positively charged residues, respectively.

The peptides were synthesized by solid phase peptide synthesis using Fmoc-based chemistry.[39,40] All peptides were purified by reverse-phase high performance liquid chromatography to greater than 98% purity and confirmed by MALDI-TOF. The peptide concentrations were determined by UV−vis as described by Edelhoch.[41,42] Peptides analyzed by sedimentation equilibrium experiments were consistent with monomers. Therefore, intermolecular interactions are most likely not contributing to the helical content observed by CD.

Circular Dichroism Spectroscopy. The circular dichroism (CD) spectra of the peptides were obtained at pH 7 in the absence of NaCl (Figure 3). Under these conditions, the Glu analogs (including Asp) will be negatively charged and the Lys analogs will be positively charged, and thereby allowing potential intrahelical ion pairing interactions. The CD spectra of GluLys3, GluLys4 and GluOrn4 were similar to the CD spectra of analogous peptides reported by Baldwin[57] and Stellwagen.[30] The magnitude of the CD signal at 222 nm reflects the α-helical content of a peptide.[59] On the basis of the CD spectra, the helical content of AspXaa5 peptides followed the trend: AspLys5 > AspOrn5 > AspDab5 > AspDap5 (Figure 3A). This trend was consistent with the helix propensity of the

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

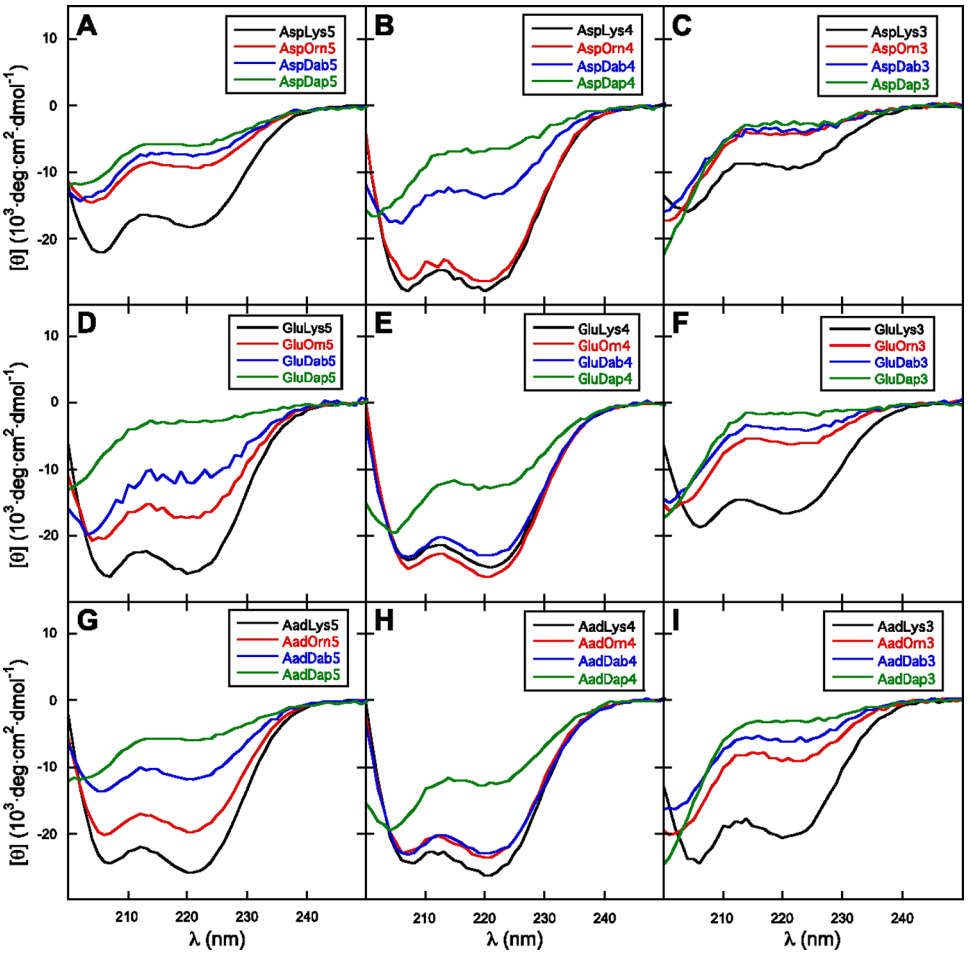


Figure 3. Circular dichroism spectra of the peptides at pH 7 (273 K) in 1 mM of phosphate, borate, and citrate buffer in mean residue ellipticity (panel A: AspLys5, AspOrn5, AspDab5, AspDap5; panel B: AspLys4, AspOrn4, AspDab4, AspDap4; panel C: AspLys3, AspOrn3, AspDab3, AspDap3; panel D: GluLys5, GluOrn5, GluDab5, GluDap5; panel E: GluLys4, GluOrn4, GluDab4, GluDap4; panel F: GluLys3, GluOrn3, GluDab3, GluDap3; panel G: AadLys5, AadOrn5, AadDab5, AadDap5; panel H: AadLys4, AadOrn4, AadDab4, AadDap4; panel I: AadLys3, AadOrn3, AadDab3, AadDap3).

positively charged residue Xaa: Lys > Orn > Dab > Dap.[29] The same trend was observed for GluXaa5 peptides and AadXaa5 peptides (Figure 3D,G). Furthermore, the trend of the helical content for a given Xaa was generally AadXaa5 > GluXaa5 > AspXaa5. This trend was consistent with the helix propensity for Glu and Asp.[32−35]

The helical content of AspXaa4 peptides followed the trend: AspLys4 ∼ AspOrn4 ≫ AspDab4 > AspDap4 (Figure 3B). Peptides AspLys4 and AspOrn4 exhibited significant helicity, whereas AspDab4 and AspDap4 showed diminished helical content. Apparently, the Lys side chain could be shortened by one methylene without affecting the helical content of the AspXaa4 peptides, suggesting significant Asp-Xaa (*i*, *i*+4) interactions involving Lys and Orn. The helical content of GluXaa4 peptides followed the trend: GluLys4 ∼ GluOrn4 ∼ GluDab4 ≫ GluDap4 (Figure 3E). Peptides GluLys4, GluOrn4, and GluDab4 exhibited significant helicity, whereas GluDap4 showed diminished helical content. In contrast to the AspXaa4 peptides, the Lys side chain could be shortened by two methylenes without affecting the helical content of the GluXaa4 peptides, suggesting significant Glu-Xaa (*i*, *i*+4) interactions involving Lys, Orn, and Dab. Apparently, by lengthening the side chain of Asp to Glu, the positively charged residue Xaa side chain could be shortened by one extra methylene (from Orn to Dab) while maintaining high helical content for ZbbXaa4 peptides. These results suggested length complementarity between Asp/Glu and Xaa; shortening Glu requires the lengthening of Xaa to compensate for the side chain length change. However, the helical content of AadXaa4 peptides followed the trend: AadLys4 ∼ AadOrn4 ∼ AadDab4 ≫ AadDap4 (Figure 3H). Only peptides AadLys4, AadOrn4, and AadDab4 exhibited significant helicity, whereas AadDap4 showed diminished helical content. Similar to the GluXaa4 peptides, the Lys side chain could only be shortened by two methylenes without affecting the helical content of the AadXaa4 peptides, suggesting significant Aad-Xaa (*i*, *i*+4) interactions involving Lys, Orn, and Dab, but not Dap. The lack of significant helical content for peptide AadDap4 may be due to the low helix propensity of Dap coupled with the large entropic penalty for the Aad side chain in forming potential Aad-Dap (*i*, *i*+4) interactions. On the basis of these results, Glu and Aad provided similar helical content for peptides with potential intrahelical Zbb-Xaa (*i*, *i*+4) interactions, with no apparent length complementarity between Glu/Aad and the positively charged residues.

The helical content of AspXaa3 peptides followed the trend: AspLys3 > AspOrn3 ∼ AspDab3 ∼ AspDap3 (Figure 3C). All four peptides were relatively nonhelical, suggesting minimal

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

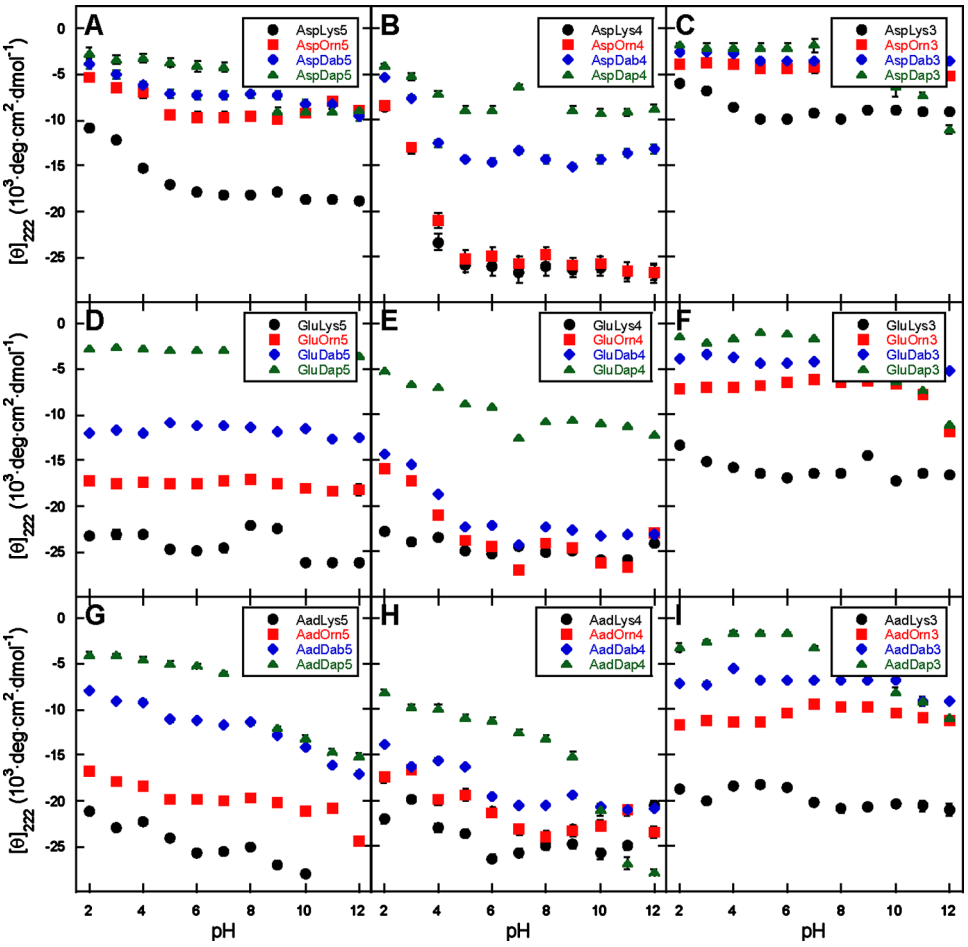


Figure 4. Circular dichroism signal at 222 nm of the peptides at pH 2−12 (273 K) in 1 mM of phosphate, borate, and citrate buffer in mean residue ellipticity (panel A: AspLys5, AspOrn5, AspDab5, AspDap5; panel B: AspLys4, AspOrn4, AspDab4, AspDap4; panel C: AspLys3, AspOrn3, AspDab3, AspDap3; panel D: GluLys5, GluOrn5, GluDab5, GluDap5; panel E: GluLys4, GluOrn4, GluDab4, GluDap4; panel F: GluLys3, GluOrn3, GluDab3, GluDap3; panel G: AadLys5, AadOrn5, AadDab5, AadDap5; panel H: AadLys4, AadOrn4, AadDab4, AadDap4; panel I: AadLys3, AadOrn3, AadDab3, AadDap3).

intrahelical Asp-Xaa (i, i+3) interaction regardless of Xaa side chain length. In contrast, the helical content of GluXaa3 peptides followed the trend: GluLys3 ≫ GluOrn3 > GluDab3 > GluDap3 (Figure 3F). Peptide GluLys3 was significantly helical, whereas the other three GluXaa3 peptides were relatively nonhelical. This result suggested that there could be significant Glu-Lys (i, i+3) interaction. Similar to the GluXaa3 peptides, the helical content of AadXaa3 peptides followed the trend: AadLys3 ≫ AadOrn3 > AadDab3 > AadDap3 (Figure 3I). On the basis of these results, it appeared that only the Lys side chain had sufficient length to support Zbb-Xaa (i, i+3) interaction, but Asp was too short to stabilize the helical conformation through intrahelical Asp-Lys (i, i+3) interactions. Similar to the ZbbXaa4 peptides, Glu and Aad provided similar helical content for ZbbXaa3 peptides with potential intrahelical Zbb-Xaa (i, i+3) interactions, with no apparent length complementarity between Glu/Aad and the positively charged residues.

The CD signal at 222 nm for all peptides were measured between pH 2 and 12 (Figure 4), to provide insight into the nature of the Zbb-Xaa intrahelical interactions. However, the high pH data could not be interpreted without ambiguity, because the Tyr phenol side chain has a pK_a near 10 and aromatic groups may contribute to the CD signal.[58] Since the helix dipole can perturb the pK_a by 1.6 pH units,[60,61] the discussion will focus on data between pH 2 and 7. At pH 2, the Glu analogues (including Asp) are most likely protonated and uncharged, while the Lys analogues are positively charged. Peptide AspLys5 clearly exhibited less helical content at pH 2 compared to pH 7 (Figure 4A), whereas the helical content of peptides AspOrn5, AspDab5, and AspDap5 did not alter significantly with pH. This is most likely due to the loss of favorable Asp^-−helix dipole interaction coupled with minimal change in helix propensity of Asp^- upon protonation to Asp^0.[32,34,35] In contrast, all GluXaa5 and AadXaa5 peptides exhibited similar CD signals between pH 2 and 7 (Figure 4D,G).

The helical content for all AspXaa4 peptides showed varying degrees of attenuation upon lowering the pH from 7 to 2 (Figure 4B), with significant pH dependence for peptides AspLys4 and AspOrn4. Interestingly, less helical content was observed at pH 2 compared to pH 7 for all GluXaa4 peptides except for GluLys4 (Figure 4E), which exhibited minimal pH dependence in its CD signal. For AadXaa4 peptides, there was slight attenuation of helical content upon lowering the pH from 7 to 2 (Figure 4H). In contrast, there was minimal change in CD signal for all ZbbXaa3 peptides except for peptides AspLys3

dx.doi.org/10.1021/bi300655z | Biochemistry 2012, 51, 7157−7172

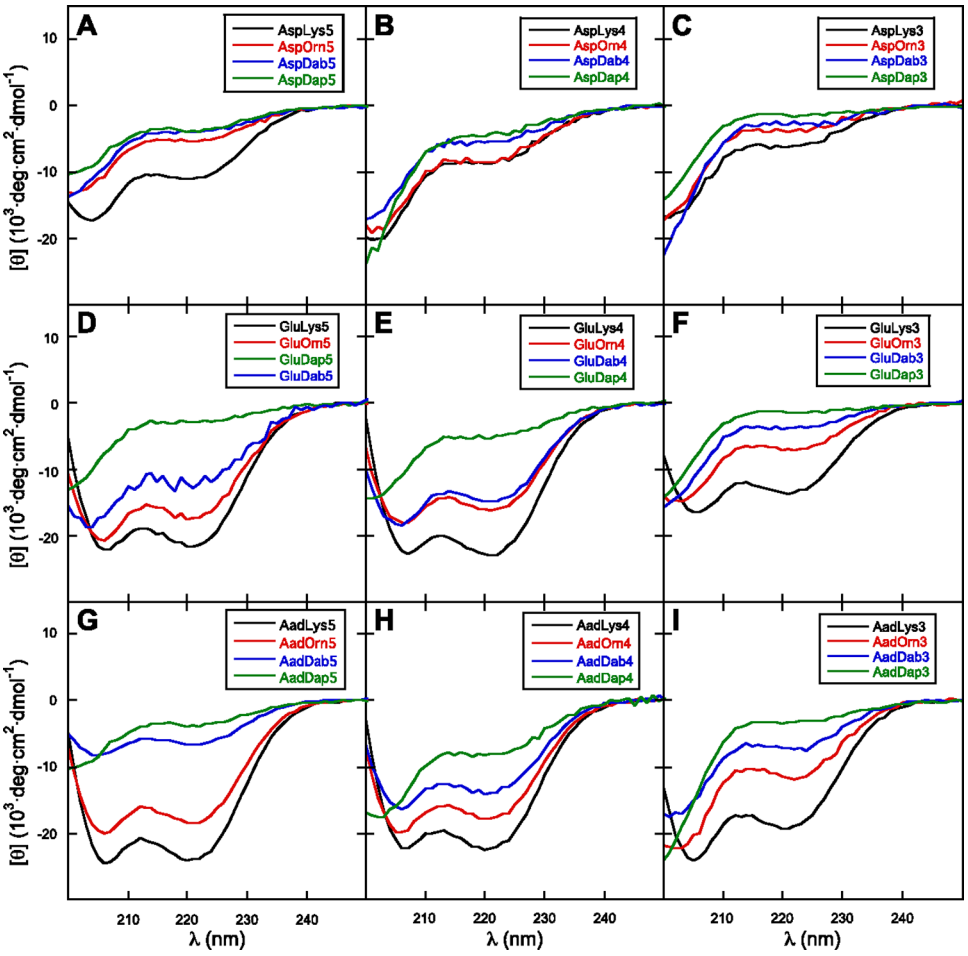


Figure 5. Circular dichroism spectra of the peptides at pH 2 (273 K) in 1 mM of phosphate, borate, and citrate buffer in mean residue ellipticity (panel A: AspLys5, AspOrn5, AspDab5, AspDap5; panel B: AspLys4, AspOrn4, AspDab4, AspDap4; panel C: AspLys3, AspOrn3, AspDab3, AspDap3; panel D: GluLys5, GluOrn5, GluDab5, GluDap5; panel E: GluLys4, GluOrn4, GluDab4, GluDap4; panel F: GluLys3, GluOrn3, GluDab3, GluDap3; panel G: AadLys5, AadOrn5, AadDab5, AadDap5; panel H: AadLys4, AadOrn4, AadDab4, AadDap4; panel I: AadLys3, AadOrn3, AadDab3, AadDap3).

and GluLys3 (Figure 4C,F,I), which exhibited noticeable signal attenuation upon lowering the pH.

The CD spectra of the ZbbXaa3, ZbbXaa4, and ZbbXaa5 peptides at pH 2 were acquired to probe the intrahelical Zbb-Xaa interactions involving neutral Zbb^0 side chains (Figure 5). The trend for the helical content of the ZbbXaa5 peptides was consistent with the helix propensity of the positively charged residue Xaa,[29] similar to the CD results at pH 7 (Figure 3). At pH 2, none of the AspXaa4 or AspXaa3 peptides exhibited significant helical content (Figure 5B,C). In contrast, the helical content for the GluXaa4 peptides followed the trend: GluLys4 > GluOrn4 ~ GluDab4 > GluDap4 (Figure 5E). Similarly, the helical content for the AadXaa4 peptides followed the trend: AadLys4 > AadOrn4 > AadDab4 > AadDap4 (Figure 5H). Both these trends suggested that the side chain length of Lys plays a significant role for Glu^0-Lys and Aad^0-Lys (i, i+4) interactions at pH 2, for which Glu and Aad are most likely neutral. This is in sharp contrast to the results at pH 7 (Figure 3E,H), for which the Lys side chain length does not appear to be critical for Glu^--Lys or Aad^--Lys (i, i+4) interactions. The trends for the helical content for the GluXaa3 and AadXaa3 peptides at pH 2 (Figure 5F,I) were the same as the trends at pH 7 (Figure 3F,I). Overall, it appeared that Aad and Glu provided peptides with similar helical contents, whereas Asp-containing peptides behaved distinctly differently.

Intrahelical Zbb-Xaa Interaction Energetics at pH 7 and 2. The side chain−side chain interaction energy ($\Delta G_{pH7interaction}$ and $\Delta G_{pH2interaction}$) between oppositely charged residues was derived from the CD data using the nesting block method[31,33,46] based on the modified Lifson-Roig theory[33,34,43,44] (Tables 1, 2, and S2). The nesting block method was modified to account for the charged side chain−helix dipole interactions for all states and changes in helix propensity upon altering the pH.[31] At pH 7, the Glu^--Lys (i, i+4), Glu^--Lys (i, i+3), and Asp^--Lys (i, i+3) interaction energies were similar to values reported by other researchers,[33,38,62] but the Asp^--Lys (i, i+4) interaction energy was much higher in magnitude compared to literature values.[38,62] No significant Zbb^--Xaa (i, i+5) interaction was observed for the ZbbXaa5 peptides, consistent with the α-helical structure (Table S2). Several ZbbXaa5 peptides exhibited less helical character compared to expected values without any intrahelical side chain−side chain interactions, suggesting that Xaa-Zbb^- (i, i+1) interactions may be stabilizing nonhelical conformations; analogous helix destabilizing interactions have been reported by Scholtz and co-workers.[33]

dx.doi.org/10.1021/bi300655z | Biochemistry 2012, 51, 7157−7172

Table 1. Energetics[a] of Zbb-Xaa (*i*, *i*+4) Intrahelical Interactions at pH 7 and 2

residue *i*	residue *i*+4	$\Delta G_{pH7interaction}$ (cal·mol^{-1})	$\Delta G_{pH2interaction}$ (cal·mol^{-1})	$\Delta\Delta G^{b}$ (cal·mol^{-1})
Asp	Lys	−736 ± 24	0[c]	−736 ± 24
Asp	Orn	−1090 ± 40	−217 ± 93	−873 ± 101
Asp	Dab	−713 ± 31	−133 ± 43	−580 ± 53
Asp	Dap	−547 ± 34	0[c]	−547 ± 34
Glu	Lys	−459 ± 34[d]	−194 ± 31[d]	−265 ± 46[d]
Glu	Orn	−742 ± 43[d]	−125 ± 34[d]	−617 ± 53[d]
Glu	Dab	−837 ± 36[d]	−281 ± 38[d]	−556 ± 52[d]
Glu	Dap	−781 ± 43[d]	0[c,d]	−781 ± 43[d]
Aad	Lys	−130 ± 24	0[c]	−130 ± 24
Aad	Orn	−243 ± 75	−33.1 ± 38.0	−209 ± 84
Aad	Dab	−531 ± 78	−143 ± 34	−388 ± 85
Aad	Dap	−473 ± 48	−96.2 ± 43.0	−377 ± 64

[a]The energetic values $\Delta G_{pH7interaction}$ and $\Delta G_{pH2interaction}$ were derived from the experimental CD data based on the nesting block method.[31,33,46] [b]$\Delta\Delta G = \Delta G_{pH7interaction} - \Delta G_{pH2interaction}$. The difference in side chain-side chain interaction energy at pH 7 and 2, which reflects the contribution of electrostatics in the Zbb-Xaa (*i*, *i*+4) interaction. [c]The experimental results are within error of the calculated predicted values without any side chain-side chain interaction. [d]Values as previously reported.[31]

Table 2. Energetics[a] of Zbb-Lys (*i*, *i*+3) Intrahelical Interactions at pH 7 and 2

residue *i*	residue *i*+3	$\Delta G_{pH\ 7\ interaction}$ (cal·mol^{-1})	$\Delta G_{pH\ 2\ interaction}$ (cal·mol^{-1})	$\Delta\Delta G^{b}$ (cal·mol^{-1})
Asp	Lys	−146 ± 36	0[c]	−146 ± 36
Glu	Lys	−256 ± 36[d]	−143 ± 32[d]	−113 ± 48[d]
Aad	Lys	−46.9 ± 30.3	−70.1 ± 25.2	23.2 ± 39.4

[a]The energetic values $\Delta G_{pH7interaction}$ and $\Delta G_{pH2interaction}$ were derived from the experimental CD data based on the nesting block method.[31,33,46] [b]$\Delta\Delta G = \Delta G_{pH7interaction} - \Delta G_{pH2interaction}$. The difference in side chain-side chain interaction energy at pH 7 and 2 reflects the contribution of electrostatics in the Zbb-Lys (*i*, *i*+3) interaction. [c]The experimental results are within error of the calculated predicted values without any side chain-side chain interaction. [d]Values as previously reported.[31]

Stabilizing Zbb^{-}-Xaa (*i*, *i*+4) interactions were observed at pH 7, regardless of side chain length of the charged amino acids Zbb and Xaa (Table 1). Interestingly, the Zbb^{-}-Xaa (*i*, *i*+4) interactions became more stabilizing as the positively charged amino acid Xaa side chain length decreased for ZbbXaa4 peptides with significant helical content (i.e., peptides AspLys4, AspOrn4, GluLys4, GluOrn4, GluDab4, AadLys4, AadOrn4, and AadDab4). In particular, the magnitude of the Asp^{-}-Xaa (*i*, *i*+4) interaction energy followed the trend Lys < Orn, whereas the magnitude of the Glu^{-}/Aad^{-}-Xaa (*i*, *i*+4) interaction energy followed the trend Lys < Orn < Dab. Similarly, the Zbb^{-}-Xaa (*i*, *i*+4) interactions became more stabilizing as the negatively charged amino acid Zbb side chain length decreased for ZbbXaa4 peptides with significant helical content. More specifically, the magnitude of both Zbb^{-}-Lys (*i*, *i*+4) and Zbb^{-}-Orn (*i*, *i*+4) interaction energies followed the trend Asp > Glu > Aad, whereas the magnitude of the Zbb^{-}-Dab (*i*, *i*+4) interaction energy followed the trend Glu > Aad. Among ZbbXaa3 peptides, only ZbbLys3 peptides exhibited stabilizing Zbb^{-}-Xaa (*i*, *i*+3) interaction energy (Tables 2 and S2). For the two peptides with significant helicity, GluLys3 and AadLys3, the Zbb^{-}-Xaa (*i*, *i*+3) interaction became more stabilizing with decreasing Zbb side chain length, with essentially no Aad^{-}-Lys (*i*, *i*+3) interaction.

The CD signal and helical content of a given peptide are determined by the helix propensity of the constituting amino acids, the specific sequence, the interaction between the charged residues and the helix dipole, and the intrahelical side chain interactions. As such, the helical content and the interaction energies may appear contradicting without considering the helix propensity. For example, the Asp-Lys (*i*, *i*+4) and Asp-Dab (*i*, *i*+4) interactions are energetically similar, but the helical content of peptide AspLys4 is much higher compared to peptide AspDab4. This apparent discrepancy is due to the significantly different helix propensity for Lys and Dab.[29]

Ion pairing interactions can occur between two formally oppositely charged species. Salt bridges occur when the two oppositely charged functionalities are also hydrogen bonded.[57,63] A hydrogen bond between a formal charge and a neutral functionality is considered to be a hydrogen bond,[63,64] which is typically stronger than hydrogen bonds between two neutral species. Intrahelical Glu-Lys (*i*, *i*+3) and (*i*, *i*+4) interactions have been attributed to electrostatics,[33,37,57] hydrogen bonding,[37,57] and possibly hydrophobics.[31,65] To gain insight into how the side chain length affects these contributing factors, the Zbb-Xaa (Zbb = Asp, Glu, Aad; Xaa = Lys, Orn, Dab, Dap) interactions at pH 7 and 2 were derived (Tables 1, 2, and S2). Changing the pH from 7 to 2 should protonate and neutralize the carboxylate side chain of the negatively charged residue Zbb^{-}. This would disrupt the electrostatic interactions with the negatively charged Zbb^{-} (including Zbb^{-}-Xaa and Zbb^{-}-helix macrodipole) and would also affect the helix propensity of the negatively charged residue.[33] Thus, changes in the CD signal upon lowering the pH from 7 to 2 should reflect changes in both electrostatic interactions and helix propensity. The lack of significant change in the CD signal for most of the ZbbXaa5 and ZbbXaa3 peptides suggests that these effects most likely cancel one another (Figure 4). For the ZbbXaa5 peptides, the loss of favorable electrostatic interaction between Zbb^{-} and the helix macrodipole is compensated by the increase in Zbb helix propensity upon protonation of the Zbb^{-} side chain. For peptides AspLys3 and GluLys3, there is a slight hint of more CD signal at pH 7 compared to pH 2 (Figure 4C,F), reflecting the extra loss of the weak Asp^{-}-Lys (*i*, *i*+3) and Glu^{-}-Lys (*i*, *i*+3) electrostatic interactions upon neutralization of the negatively charged residue, respectively. For all AspXaa4 and GluXaa4 peptides except GluLys4, the loss of Zbb^{-}-Xaa (*i*, *i*+4) interaction is clearly visible upon lowering the pH from 7 to 2 (Figure 4B,E). The AadXaa4 peptides exhibited similar pH dependence compared to the GluXaa4 peptides (Figure 4E,H).

The Zbb^{0}-Xaa interactions at pH 2 (involving the protonated neutral Zbb^{0}) should represent the nonelectrostatic components of the Zbb^{-}-Xaa interaction, namely, the hydrogen bonding[37,57] and hydrophobics.[31,65] As such, the interaction energetics at pH 2 may be considered as an estimate for the upper limit of the strength of the intrahelical side chain−side chain hydrogen bond. Furthermore, the side chain−side chain interaction energies at pH 2 ($\Delta G_{pH2interaction}$) should represent the upper limit for interactions involving Zbb^{0}, because the CD signals were not unambiguously leveled off upon decreasing the pH value to 2. Therefore, there may be some residual electrostatic interactions involving formally charged Zbb^{-}. Nonetheless, the Glu^{0}-Lys (*i*, *i*+3) and (*i*, *i*+4) interaction

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

energies at pH 2 (Tables 1 and 2) were similar to the values reported by other researchers.[38] The ZbbXaa5 peptides did not exhibit any significant Zbb^0-Xaa (i, i+5) interaction at pH 2 (Table S2), consistent with the α-helical structure. For ZbbLys3 peptides, GluLys3 and AadLys3 exhibited significant Zbb^0-Lys (i, i+3) interaction energy at pH 2 (Table 2). In contrast, all ZbbXaa4 peptides exhibited varying degrees of Zbb^0-Xaa (i, i+4) interaction at pH 2 except for AspLys4, AspDap4, GluDap4, and AadLys4 (Table 1). These results suggested that the Asp^-/Aad^--Lys (i, i+4) interactions and Asp^-/Glu^--Dap (i, i+4) interactions at pH 7 completely relied on the formal negative charge on Asp^-/Glu^-/Aad^-. All other Zbb^--Xaa (i, i+4) interactions included nonelectrostatic components such as hydrogen bonding[37,57] or hydrophobics.[31,65]

The difference between side chain−side chain interaction energy at pH 7 and 2 represents the lower limit of the contribution of electrostatics to the Zbb^--Xaa interaction ($\Delta\Delta G$, Tables 1 and 2), due to the apparent incomplete neutralization of the Zbb^- side chain at pH 2. Accordingly, the Asp-Lys (i, i+3) interaction was completely electrostatics (Table 2), whereas electrostatics was only an important component of the Glu^--Lys (i, i+3) interaction. Furthermore, all Asp-Xaa (i, i+4) interactions were predominantly electrostatics (Table 1), and the electrostatic contribution to the interaction at pH 7 followed the trend Orn > Lys > Dab∼Dap. This is clearly reflected in the pH dependence of the CD signal for the AspXaa4 peptides (Figure 4B). The lack of any Asp^0-Dap and Asp^0-Lys (i, i+4) interactions at pH 2 suggests that these interactions at pH 7 are purely electrostatics. In contrast, the electrostatic component of the Glu^--Xaa (i, i+4) interaction energy at pH 7 followed the trend Dap > Orn > Dab > Lys. However, the electrostatic component of the Aad^--Xaa (i, i+4) interaction energy at pH 7 followed yet another trend Dap ∼ Dab > Orn > Lys, apparently decreasing with increasing Xaa side chain length. The residual interaction at pH 2 for these peptides may be due to side chain−side chain hydrogen bonding[37,57] or hydrophobic interactions.[65] The residual side chain−side chain hydrogen bond would serve to bring the two functionalities together, whereas the hydrophobic methylenes on the side chains would shield the helix hydrogen bond to stabilize the helix.[66−68]

The side chain length of Lys could be shortened by one methylene (to Orn) while maintaining significant helicity for AspXaa4 peptides based on the CD data at pH 7 (Figure 3B). Similarly, the side chain length of Lys could be shortened by two methylenes (to Dab) while maintaining significant helicity for the GluXaa4 and AadXaa4 peptides (Figure 3E,H). Apparently, the side chain length of the positively charged amino acid Xaa and negatively charged amino acid Zbb may be for reaching one another to enable favorable Zbb-Xaa (i, i+4) interactions. Side chains too short to reach, such as that for Dap, did not seem to support Zbb-Xaa (i, i+4) interactions. Accordingly, only Lys has a side chain long enough to reach and interact with Zbb for a Zbb-Xaa (i, i+3) interaction. This simple explanation suggests that Zbb-Xaa (i, i+4) interaction should become less stabilizing with decreasing positively charged amino acid Xaa side chain length for a given negatively charged amino acid Zbb. Also, the Zbb-Xaa (i, i+4) interaction should become less favorable with decreasing Zbb side chain length for a given Xaa. However, the interaction energetics at pH 7 based on the same CD data do not agree with these extrapolations (Tables 1, 2, and S2). To resolve this apparent contradiction, conformational analysis on relevant model peptides was performed.

Conformational Analysis of Short Model Peptides. A detailed conformational analysis on model peptides was performed by molecular mechanics calculations. Three series of short hexapeptides were investigated: MSZbbXaa3, MSZbbXaa4, and MSZbbXaa5 (Figure 6). All peptides included one potential side chain−side chain interaction between a negatively charged amino acid (Asp, Glu, or Aad) and a positively charged residue (Lys, Orn, Dab, or Dap) placed three, four, or five residues apart. For each side chain dihedral angle (or χ angle) involving sp^3 carbons, three possible low energy staggered conformations were considered: gauche− (60°, g−), trans (180°, t), and gauche+ (300°, g+) (Figure 7).[69,70] For the dihedral angle involving the sp^2 carboxylate carbon of the negatively charged amino acid, six conformations were considered: 0°, 30°, 60°, 90°, 120°, and 150°. For the χ_1 dihedral of a residue in an α-helix in nature, the t and g+ conformations are known to be more prevalent and more stable than the g− conformation,[69,70] due to significant gauche interactions in the g− conformation and steric clashes with the

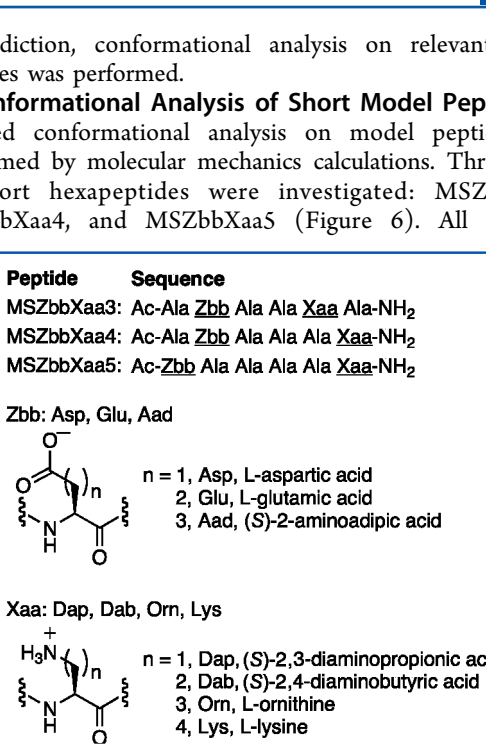


Figure 6. Sequences of the short MSZbbXaa3, MSZbbXaa4, and MSZbbXaa5 model peptides for the conformational analysis by molecular mechanics calculations. The underlined generic three-letter codes Zbb and Xaa represent the negatively and positively charged residues, respectively.

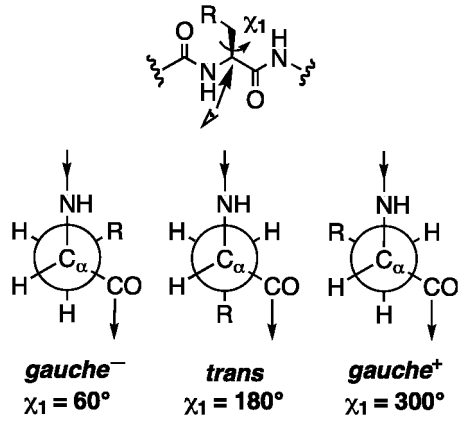


Figure 7. Newman projections of the three low energy χ_1 dihedrals *gauche*−, *trans*, and *gauche*+.

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

Table 3. Summary of Low Energy Conformations from Conformational Analysis of MSZbbXaaN Peptides by Molecular Mechanics Calculations

	lowest energy conformer	conformations within 4 kcal of the lowest energy conformer		
peptide	energy (kcal)	no.[a]	salt bridge[b] (%)	major conformations with salt bridge number (Zbb χ_1, Xaa χ_1)
MSAspLys3	−52.0	11	100	11(*g*+, *g*+)
MSAspLys4	−56.8	3	100	3(*g*−, *g*+)
MSAspOrn3	−45.3	18	0	
MSAspOrn4	−49.6	5	60	2(*t*, *g*+), 1(*g*−, *g*+)
MSAspDab3	−40.8	26	0	
MSAspDab4	−43.5	12	17	2(*t*, *g*+)
MSAspDap3	−35.6	14	0	
MSAspDap4	−37.7	10	0	
MSGluLys3	−50.1	11	100	9(*g*+, *g*+), 1(*t*, *g*+), 1(*g*+, *t*)
MSGluLys4	−50.1	24	67	10(*g*−, *g*+), 4(*t*, *g*+), 2(*t*, *t*)
MSGluOrn3	−43.4	26	19	2(*g*+, *g*+), 2(*t*, *t*), 1(*t*, *g*+)
MSGluOrn4	−45.8	18	56	6(*g*−, *g*+), 5(*t*, *g*+)
MSGluDab3	−39.5	19	0	
MSGluDab4	−43.1	10	50	4(*t*, *g*+), 1(*g*−, *g*+)
MSGluDap3	−34.3	9	0	
MSGluDap4	−35.8	10	0	
MSAadLys3	−55.8	21	100	21(*g*+, *g*+)
MSAadLys4	−54.5	60	100	24(*t*, *g*+), 22(*g*−, *g*+), 14(*g*+, *g*+)
MSAadOrn3	−50.7	14	100	14 (*g*+, *g*+)
MSAadOrn4	−49.6	33	100	21(*t*, *g*+), 9(*g*−, *g*+), 3(*g*+, *g*+)
MSAadDab3	−45.3	16	50	4(*g*+, *g*+), 3(*t*, *t*), 1(*t*, *g*+)
MSAadDab4	−45.2	55	71	32(*t*, *g*+), 5(*g*−, *g*+), 2(*t*, *t*)
MSAadDap3	−38.0	43	12	2(*g*+, *g*−), 2(*t*, *g*−), 1(*g*+, *g*+)
MSAadDap4	−39.5	41	20	7(*t*, *g*+), 1(*t*, *g*−)

[a]The number of conformations within 4 kcal of the lowest energy conformer for each peptide. [b]The percentage of conformations within 4 kcal of the lowest energy conformer with a Zbb-Xaa salt bridge, which is a hydrogen bonded ion pair.[57,63]

nearby backbone. All possible χ angle combinations for the negatively charged residue Zbb and the positively charged residue Xaa with residue spacings of three, four, and five were investigated. A combined total of 84 240 conformations were minimized (with the backbone fixed in an ideal helical conformation). Results from calculations with and without fixing the backbone gave similar results. Furthermore, adding 10 helical Ala residues before and after the short hexapeptide segment with and without fixing the backbone also gave similar results. As such, the results from the short hexapeptides with the backbone fixed during minimization were examined in more detail.

The energy of the model peptides with the same positively charged residue and negatively charged residues (e.g., MSAspLys3, MSAspLys4, and MSAspLys5) could be compared, because the peptides were designed to have the same constituting atoms represented by the same forcefield parameters. The interaction between oppositely charged residues with different spacings were compared by evaluating the energies of the lowest energy conformers for the MSZbbXaa3, MSZbbXaa4, and MSZbbXaa5 peptides (Tables 3 and S3). On the basis of the lowest energy conformers, the MSZbbXaa5 peptide exhibited higher energy (less negative energy, i.e., less stable) compared to the corresponding MSZbbXaa4 and MSZbbXaa3 peptides (Table S3). This is consistent with the lack of experimentally derived Zbb-Xaa (*i*, *i*+5) interaction energy at pH 7 (Table S1), and the geometry of an ideal α-helix. The MSAspXaa4 peptides were significantly lower in energy (i.e., more stable) compared to the MSAspXaa3 peptides, suggesting that Asp-Xaa (Xaa = Lys, Orn, Dab, Dap) interaction is energetically more favorable for (*i*, *i*+4) compared to (*i*, *i*+3) spacing. This is consistent with the experimentally derived interaction energies at pH 7 (Tables 1, 2, and S2). Large energy differences between (*i*, *i*+4) and (*i*, *i*+3) spacing was also observed for the corresponding MSGluOrn and MSGluDab peptides (Table 3). This is also consistent with the significant helical content for peptides GluOrn4 and GluDab4 with large stabilization side chain interactions (Figure 3E and Table 1), and low helical content for peptides GluOrn3 and GluDab3 with minimal side chain interactions (Figure 3F and Table S2). A smaller energy difference was observed for MSGluDap peptides (Table 3), consistent with the low helical content for peptides GluDap3 and GluDap4 (Figure 3E,F). The same energy was obtained for MSGluLys3 and MSGluLys4 (Table 3), consistent with the high helical content and side chain interaction for peptides GluLys3 and GluLys4 (Figure 3E,F; Tables 1 and 2). Small energy differences were observed for MSAadXaa peptides with (*i*, *i*+4) and (*i*, *i*+3) spacings (Table 3). Overall, these computational results are consistent with the CD experimental results (Figure 3) and interaction energies ($\Delta G_{pH7interaction}$, Tables 1 and 2).

Conformations within 4 kcal of the lowest energy conformer for each peptide were then examined (Tables 3 and S3), because room temperature can provide up to 4 kcal·mol^{-1} of thermal energy. In general, the number of conformations increased with increasing side chain length; however, the presence of salt bridges can limit the number of low energy conformers. Nonetheless, this general trend in number of conformations is consistent with higher side chain conformational entropy for longer side chains compared to shorter side

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

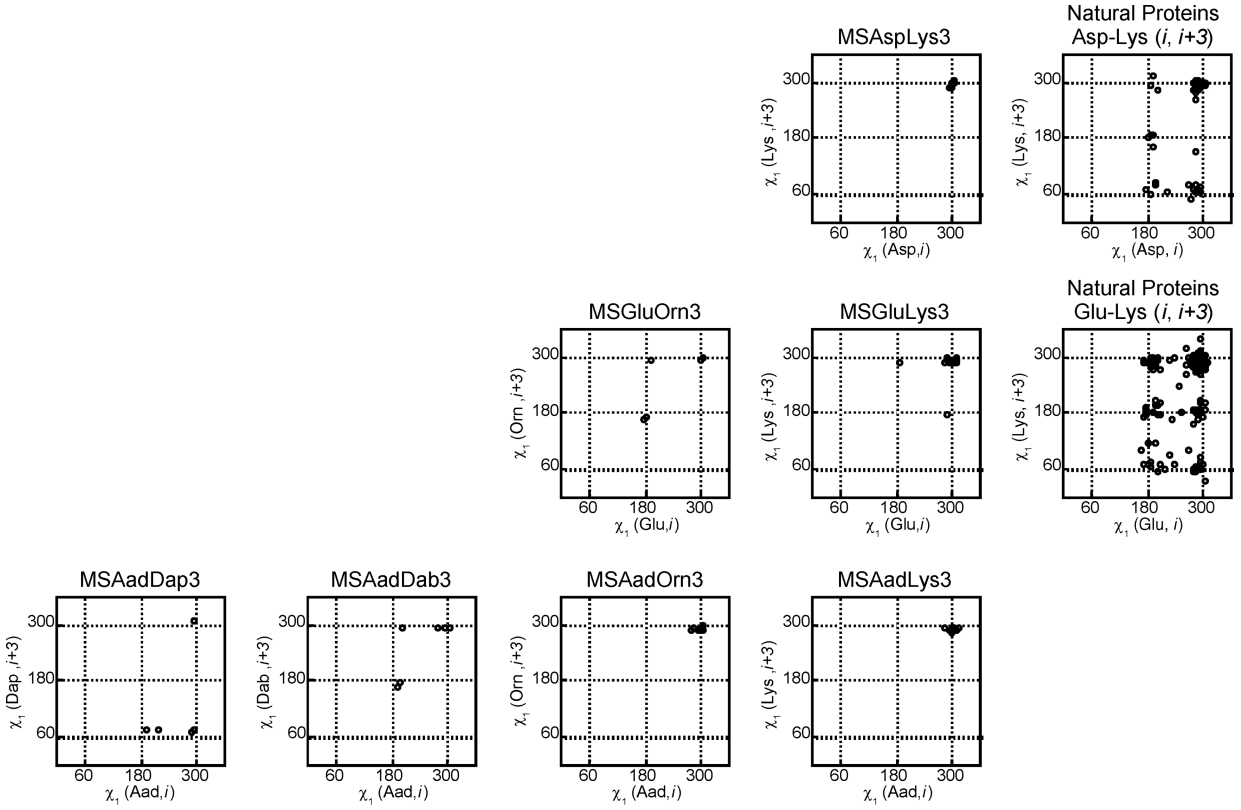


Figure 8. χ_1(Zbb, *i*)−χ_1(Xaa, *i*+3) plots for MSZbbXaa3 peptides with salt bridges within 4 kcal of the lowest energy conformation from conformational analysis by molecular mechanics calculations are in the left four columns (top row from left to right, MSAspLys3; middle row from left to right, MSGluOrn3 and MSGluLys3; bottom row from left to right, MSAadDap3, MSAadDab3, MSAadOrn3, MSAadLys3). The side chain conformation of intrahelical Asp-Lys (*i*, *i*+3) and Glu-Lys (*i*, *i*+3) salt bridges from the survey of natural proteins are in the most right-hand-side column.

chains. Importantly, the weighed average energy (based on Boltzmann distribution) of the peptides followed the same trends as the lowest energy for each peptide (vide supra). Furthermore, none of the MSZbbXaa5 peptides exhibited intrahelical Zbb-Xaa salt bridges (Table S3), consistent with the α-helix geometry. In general, the MSZbbXaa4 peptides exhibited a higher percentage of conformations with intrahelical salt bridges compared to the corresponding MSZbbXaa3 peptides (Table 3), consistent with the CD data at pH 7 (Figure 3). Furthermore, there were more low energy conformers for the MSAadXaa4 peptides compared to MSAadXaa3 peptides except when Dap was the positively charged residue. The energy should reflect the enthalpic component for the peptide conformations, whereas the number of low energy conformers may be viewed as a crude estimate of the entropic component for the peptides. The combination of the two components suggests that Aad-Xaa (*i*, *i*+4) interaction may be more favorable compared to Aad-Xaa (*i*, *i*+3) (as derived experimentally) due to entropic reasons for the longer residues Lys, Orn, and Dab, but due to enthalpic reasons for Dap.

The conformations of the low energy conformers that formed intrahelical salt bridges were then inspected in detail (Table 3, Figures 8 and 9). Since none of the model peptides with Zbb-Xaa (*i*, *i*+5) sequence patterns formed salt bridges, these peptides will not be discussed further. The combination of χ_1 dihedrals is represented in parentheses (Table 3), designating the conformation for the negatively charged residue at position *i* followed by the conformation for the positively charged residue at position *i*+3 or *i*+4. For example, the conformation with *t* for both residues at positions *i* and *i*+3 would be designated (*t*, *t*) for (*i*, *i*+3). For the MSAspXaa3 model peptides, only MSAspLys3 exhibited intrahelical salt bridges for the low energy conformers (Table 3 and Figure 8). Furthermore, these low energy conformers with intrahelical salt bridges all involved the dihedral combination (*g*+, *g*+). For the MSGluXaa3 model peptides, MSGluLys3 and MSGluOrn3 exhibited intrahelical salt bridges for the low energy conformers (Table 3 and Figure 8). The major dihedral combination for model peptide MSGluLys3 with an intrahelical salt bridge was (*g*+, *g*+); the minor combinations were (*t*, *g*+) and (*g*+, *t*). Interestingly, the dihedral combinations for MSGluOrn3 with an intrahelical salt bridge included (*g*+, *g*+), (*t*, *g*+), and (*t*, *t*), with (*g*+, *g*+) and (*t*, *t*) as the major dihedral combinations. For MSAadXaa3 peptides, the dihedral combination (*g*+, *g*+) was present in the low energy conformers with an intrahelical salt bridge regardless of the side chain length of the positively charged residue. In fact, (*g*+, *g*+) was the only dihedral combination for MSAadLys3 and MSAadOrn3, whereas several other combinations were observed for MSAadDab3 and MSAadDap3. According to these results, extending the negatively charged amino acid from Asp to Glu apparently enabled the shortening the positively charged amino acid from Lys to Orn while sustaining intrahelical salt bridge formation. Further extending the negatively charged residue to Aad enabled shortening the positively charged residue all the way to Dap while maintaining intrahelical salt bridge formation.

dx.doi.org/10.1021/bi300655z | *Biochemistry* 2012, 51, 7157−7172

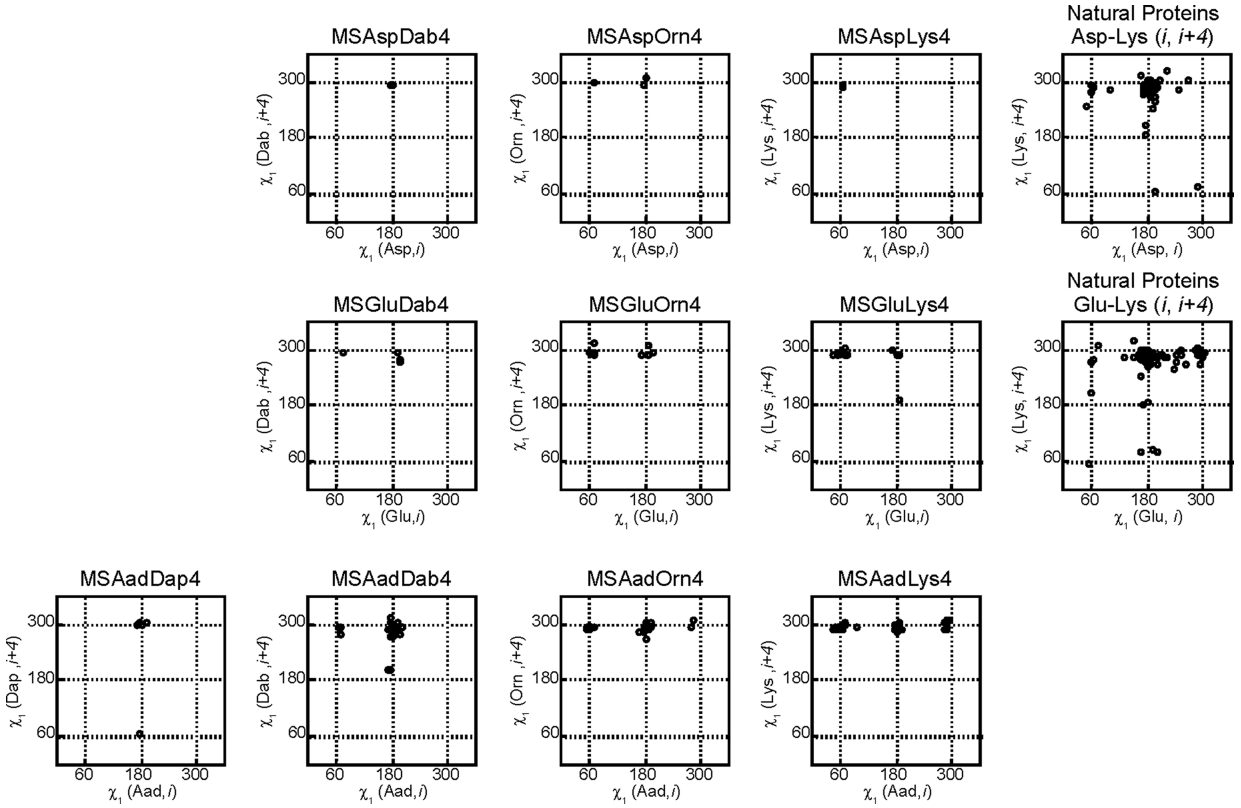


Figure 9. χ_1(Zbb, i)−χ_1(Xaa, i+4) plots for MSZbbXaa4 peptides with salt bridges within 4 kcal of the lowest energy conformation from conformational analysis by molecular mechanics calculations are in the left four columns (top row from left to right, MSAspDab4, MSAspOrn4, MSAspLys4; middle row from left to right, MSGluDab4, MSGluOrn4, MSGluLys4; bottom row from left to right, MSAadDap4, MSAadDab4, MSAadOrn4, MSAadLys4). The side chain conformation of intrahelical Asp-Lys (i, i+4) and Glu-Lys (i, i+4) salt bridges from the survey of natural proteins are in the most right-hand-side column.

Table 4. Survey Results of Oppositely Charged Residues with Various Spacings in Natural Protein Helices[a]

position i	position i +n	n	occurrence[a]	helix pair propensity[b]	pair propensity[c]	salt bridge occurrence[d]	major conformations with salt bridge number (Zbb χ_1, Xaa χ_1)
Glu	Lys	3	1940	1.95 ± 0.06	1.31 ± 0.03	226	146(g+, g+), 22(g+, t), 19(t, g+)
Asp	Lys	3	1054	1.37 ± 0.05	1.44 ± 0.05	49	25(g+, g+), 11(g+, g−)
Glu	Lys	4	1887	2.12 ± 0.07	1.38 ± 0.04	124	97(t, g+), 17(g+, g+)
Asp	Lys	4	918	1.46 ± 0.06	1.35 ± 0.05	95	88(t, g+)
Glu	Lys	5	809	1.36 ± 0.06	0.66 ± 0.02	0	
Asp	Lys	5	460	0.88 ± 0.04	0.76 ± 0.03	2	

[a]The occurrence of oppositely charged residues placed three, four, and five residues apart in helices six residues or longer in the nonredundant protein structure database PDBselect (April 2009, 25% threshold).[47,48] The helix conformation is defined by the backbone dihedral angles (ϕ, ψ) according to the values provided by Balaram and co-workers.[49−52] [b]The occurrence for the various sequence patterns divided by the corresponding expected value. The expected value was obtained by bootstrapping the complete PDBselect database, thereby including the bias for the occurrence of each residue for the helix conformation. Since bootstrapping was performed 100 000 times, this enabled the calculation of a standard deviation for the expected value and thus helix pair propensity. [c]The occurrence for the various sequence patterns divided by the corresponding expected value. The expected value was obtained by bootstrapping the helices in the PDBselect database, thereby removing the bias for the occurrence of each residue for the helix conformation. Since bootstrapping was performed 100 000 times, this enabled the calculation of a standard deviation for the expected value and thus helix pair propensity. [d]The number of the occurrences for each sequence pattern that actually forms salt bridges between the oppositely charged side chain functionalities. The presence of a salt bridge is defined by an N−O distance less than or equal to 3 Å.

For both MSAspXaa4 and MSGluXaa4 model peptides, salt bridges were observed for peptides with Lys, Orn, and Dab as the positively charged residue, but not the shortest residue Dap (Table 3 and Figure 9). The dihedral combination for these model peptides with intrahelical salt bridges involved (t, g+) and (g−, g+) (Table 3 and Figure 9). The dihedral combination (t, t) was also present for peptide MSGluLys4. All MSAadXaa4 model peptides exhibited low energy conformations with salt bridges regardless of side chain length of the positively charged residue. Similar to the MSAspXaa4 and MSGluXaa4 peptides with salt bridges, the conformations of the MSAadXaa4 model peptides involved the dihedral combinations (t, g+) and (g−, g+). Model peptides with the longer Orn and Lys residues also exhibited the dihedral combination (g+, g+). Model peptides with the shorter Dap and Dab residues exhibited extra dihedral combinations (t, g−) and (t, t), respectively. Overall, the major conformations for supporting Zbb-Xaa(i, i+4) interactions were similar regardless of side chain length. Furthermore, length-

dx.doi.org/10.1021/bi300655z | Biochemistry 2012, 51, 7157−7172

ening the negatively charged residue from Glu to Aad enabled the shortening the positively charged residue to Dap while maintaining the intrahelical salt bridge.

These molecular mechanics calculation results appeared to explain the experimental data with some conformational insight. However, it was unclear whether or not these conformational preferences were present in natural proteins. To further confirm the results from this conformational analysis, intrahelical Asp-Lys and Glu-Lys salt bridges in natural proteins were examined for comparison.

Intrahelical Asp-Lys and Glu-Lys Salt Bridges in Protein Structures. A survey was performed on the nonredundant protein structure database PDBselect (April 2009, 25% threshold)[47,48] to study the side chain conformations that support intrahelical Asp-Lys and Glu-Lys interactions in naturally occurring proteins. A total of 4418 protein chains involving 666 086 residues were considered. The α-helical conformation was defined based on backbone dihedrals.[49−52] Only α-helices of six residues or longer were considered to exclude helices with less than one turn.[31,51,52] On the basis of these criteria, there were 17 622 helices involving 236 790 helical residues. The number of occurrences for Asp-Lys and Glu-Lys sequence patterns with (i, i+3), (i, i+4), and (i, i+5) spacings were compiled (Table 4). Helix pair propensity represents how frequently a particular sequence pattern occurs in a helix compared to any structure in the database. This would include biases from both how frequently the amino acids in the sequence pattern occur in helices and how frequently the sequence pattern itself occurs in the context of a helix. In contrast, pair propensity in a helix represents how frequently a sequence pattern occurs in the context of a helix. The propensity for each sequence pattern was derived by dividing the occurrence by the expected occurrence in the appropriate context. The expected occurrence was obtained by bootstrapping the entire database for the helix pair propensity and by bootstrapping the helices in the database for the pair propensity (in helices). Bootstrapping was performed 100 000 times for these cases to obtain standard deviations for the expected occurrence. This enabled the calculation of standard deviations for the statistical propensities and Z values (Tables S4 and S5), which was used to derive the P values (Tables S4 and S5). Helix pair propensities greater than unity represent higher occurrence compared to expected occurrence in all structures, whereas pair propensities greater than unity indicate higher occurrence compared to expected occurrence in helices in the database. As such, helix pair propensities less than unity represent lower occurrence compared to the expected value for all structures, whereas pair propensities less than unity indicate lower occurrence compared to the expected value for helical structures.

The helix pair propensity for Asp-Lys (i, i+n) sequence patterns followed the trend: (i, i+4) > (i, i+3) > (i, i+5) (Tables 4 and S4). Similarly, the helix pair propensity for Glu-Lys (i, i+n) sequence patterns followed the trend: (i, i+4) > (i, i+3) > (i, i+5) (Tables 4 and S4). Both of these trends appeared to be consistent with experimentally derived intrahelical Asp-Lys and Glu-Lys interaction energies (Tables 1, 2, and S2), and mostly consistent with the energy of the low energy conformers from the conformation analysis by molecular mechanics calculations (Tables 3 and S3). The helix pair propensity for sequence patterns involving Glu was higher than that for the corresponding pattern involving Asp. However, the helix pair propensity for a given sequence pattern is determined by both the inherent frequency of occurrence of the amino acid in a helix and the prevalence of the sequence pattern in a helix (i.e., the pair propensity in helix). To remove the bias caused by the inherent frequency of occurrence for the amino acids in a helix, the pair propensity for each sequence pattern was compiled (Tables 4 and S5). Sequence patterns with (i, i+3) and (i, i+4) spacings exhibited pair propensities greater than unity, whereas sequence patterns with (i, i+5) spacing exhibited pair propensities lower than unity. These results suggest that Asp-Lys and Glu-Lys sequence patterns with (i, i+3) and (i, i+4) spacings occur more frequently than expected for helical structures, but Asp-Lys and Glu-Lys (i, i+5) do not occur as frequently as expected, similar to published results by Scheraga and co-workers.[71]

The three-dimensional structures of these occurrences were then examined in detail. Salt bridges were defined by an N−O distance of 3 Å or less. Only a small number of the sequence patterns were manifested as salt bridge interactions in the protein structure (Table 4), similar to findings by Thornton and co-workers.[72] Furthermore, Asp-Lys and Glu-Lys (i, i+5) salt bridges were essentially nonexistent (Table 4). The side chain conformations of the intrahelical salt bridges were then examined further (Table 4; last column in Figures 8 and 9). The most prevalent χ_1 dihedral combination of both Asp-Lys and Glu-Lys (i, i+3) intrahelical salt bridges was (g+, g+), consistent with the conformational analysis by molecular mechanics calculations (Table 3 and Figure 8). The major χ_1 dihedral combination of both Asp-Lys and Glu-Lys (i, i+4) intrahelical salt bridges was (t, g+), which was represented in the conformational analysis of the MSZbbXaa4 peptides (Figure 9). Importantly, the combinations of χ_1 dihedrals for the Asp-Lys and Glu-Lys intrahelical salt bridges in natural protein structures were similar to those from the molecular mechanics calculations (Figures 8 and 9), validating the conformational analysis results.

DISCUSSION

The effect of side chain length and sequence spacing on intrahelical carboxylate-ammonium ion pairing interaction and helical content for short peptides was explored in this study. For the positively charged residues with different lengths, only Lys supported any helix stabilizing Zbb$^-$-Xaa (i, i+3) interaction at pH 7. Furthermore, the Glu-Lys (i, i+3) interaction was the most helix stabilizing among the negatively charged residues (Table 2). It would be reasonable to conclude that the most optimal combination for a Zbb$^-$-Xaa (i, i+3) interaction would involve two and four methylenes for the negatively and positively charged residues, respectively. However, there is no side chain complementarity, meaning the total number of methylenes (in the charged residues) alone is not sufficient to form ideal side chain−side chain interactions in the helical conformation; Lys must be involved for any helix stabilizing Zbb$^-$-Xaa (i, i+3) interaction. Similarly, the most stabilizing Zbb$^-$-Xaa (i, i+4) interaction for the various negatively charged residues were Asp-Orn, Glu-Dab, and Aad-Dab (Table 1). Again, there is no apparent side chain complementarity.

Our studies show that Asp and Glu provide helical peptides with distinctly different characteristics. For example, the helical content of AspLys4 is pH dependent (Figure 4B), whereas the helical content of GluLys4 is pH independent (Figure 4E). Furthermore, the helical content of GluLys3 is much higher compared to AspLys3 (Figure 4C,F), which exhibits minimal

helical content. In contrast, the Aad- and Glu-containing peptides exhibit similar characteristics (Figure 4D−I). Therefore, it appears that nature chose the negatively charged residues Asp and Glu for variation, because increasing the Glu side chain length to Aad does not furnish any characteristics Glu cannot provide.

The positively charged Lys can support significant helical content for peptide GluLys3 at pH 7, whereas shortening Lys to Orn cannot support significant helical content for peptide GluOrn3 at pH 7 (Figure 3F). The same trend is observed for the corresponding Asp peptides, albeit with less difference in helical content (Figure 3C). Furthermore, the helical content of peptide GluLys4 is rather robust and not significantly affected by pH, whereas the helical content of peptide GluOrn4 is attenuated by lowering the pH from 7 to 2 (Figure 4E). In contrast, the helical content of both AspLys4 and AspOrn4 is significantly attenuated by lowering the pH from pH 7 to 2 (Figure 3B). As such, employing Orn instead of Lys would result in similar pH dependence for peptides with Asp-Orn (*i*, *i*+4) spacing and peptides with Glu-Orn (*i*, *i*+4) spacing. Furthermore, peptides with Asp-Orn (*i*, *i*+3) spacing and peptides with Glu-Orn (*i*, *i*+3) spacing would show similar low helical content with minimal Asp-Orn (*i*, *i*+3) and Glu-Orn (*i*, *i*+3) interactions. Therefore, it appears that nature chose Lys over Orn to expose the inherent difference between Asp and Glu. As for the underlying reason for the Arg side chain length, more studies would be necessary to provide further insight into this matter.

This study has focused on the positively charged Lys analogues with an ammonium group. The other positively charged amino acid Arg bears a guanidinium group (Figure 1). The two functionalities have different geometries, charge distribution, and hydrogen bonding capacity. Therefore, Lys and Arg exhibit different geometric patterns in protein ion pairs.[13,73,74] Furthermore, the helix propensity of Lys and the Lys analogues increase with increasing side chain length,[29] whereas the helix propensity of the Arg analogues peak at three methylenes (Arg);[52] further increase in the side chain length lowers the helix propensity.[52] Accordingly, the Arg analogues should behave distinctly differently compared to the Lys analogues in intrahelical ion pairs; however, experimental studies will be necessary to reveal the details.

The helical conformation has also been promoted and stabilized in short peptides by linking properly positioned side chain functionalities through covalent bonds such as metal−ligand complexation,[75] disulfide bonds,[76] amide bonds,[77] bis-amide linkages,[78−80] all-hydrocarbon linkages,[81−83] and bis-thioether linkages.[84] These helical peptides with covalently linked side chains should provide robust helices under a wider range of conditions compared to the helical peptides presented in this study, because a covalent bond is energetically worth more than an ion pairing interaction. Nevertheless, some of the peptides presented in this study such as AspLys4 can provide pH-dependent switches for which the conformation of the peptide changes with pH, complementing the robust helices achieved through covalent linkages.

CONCLUSIONS

Peptides with Zbb-Lys and Zbb-Orn (*i*, *i*+4) sequence patterns exhibited significant helical content at pH 7 regardless of the negatively charged amino acid side chain length (Figure 3). Interestingly, the helical content of peptides with Glu-Lys and Aad-Lys (*i*, *i*+4) sequence patterns were pH independent (Figure 4), whereas the peptide with the Asp-Lys (*i*, *i*+4) sequence pattern was pH dependent. Apparently, Glu and Aad provide similar effects, whereas Asp is distinctly different from the two longer negatively charged residues. Similarly, peptides with Glu-Lys and Aad-Lys (*i*, *i*+3) sequence patterns were significantly helical at pH 7 (Figure 3), whereas the peptide with the Asp-Lys (*i*, *i*+3) sequence pattern exhibited low helical content, again distinguishing Asp from Glu and Aad. Furthermore, peptides with Zbb-Orn (*i*, *i*+3) sequence patterns exhibited low helical content (Figure 3), highlighting the uniqueness of the Lys side chain length. The helical content of these peptides was recapitulated in the conformational analysis by forcefield calculations (Table 3). On the basis of these calculations, the low energy dihedral combinations for Zbb-Xaa (*i*, *i*+3) and (*i*, *i*+4) sequence patterns were (*g*+, *g*+) and (*t*, *g*+) (Table 3), respectively, consistent with surveys of natural protein structures (Table 4). Since Glu and Aad provided similar helical contents, it is not surprising that nature employs only one of these two amino acids, but selected Asp to provide different structural effects. These results should be useful in designing structures with robust helices over a wide pH range for peptide-based materials[85,86] and helix-based pH-dependent peptide switches.[87−96]

ASSOCIATED CONTENT

Supporting Information

Materials and methods section with details of the synthesis and characterization of the peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +886-2-33669789. Fax: +886-2-23636359. E-mail: rpcheng@ntu.edu.tw.

Funding

This work was supported by NYSTAR James D. Watson Investigator Program (R.P.C.), Kapoor funds (R.P.C.), The State University of New York at Buffalo (R.P.C.), National Science Foundation (R.P.C., CHE-0809633; R.F., MCB-0211754), National Taiwan University (R.P.C.), National Science Council (R.P.C., NSC-97-2113-M-002-002-MY2, NSC-99-2113-M-002-002-MY2, NSC-101-2113-M-002-006-MY2).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank the Computer and Information Networking Center at National Taiwan University for the support of the high-performance computing facilities. The authors would like to thank Professor Cheu-Pyeng Cheng (Department of Chemistry, National Tsing Hua University) for helpful discussions, and Ms. Chia-Wen Kuo for proofreading the manuscript.

ABBREVIATIONS USED

Aad, (*S*)-aminoadipate; Ala, alanine; Arg, arginine; Asp, aspartate; CD, circular dichroism spectroscopy; Dab, (*S*)-2,4-diaminobutyric acid; Dap, (*S*)-2,3-diaminopropionic acid; Fmoc, *N*-9-fluorenylmethyloxycarbonyl; Glu, glutamate; Gly, glycine; Lys, lysine; MALDI-TOF, matrix-assisted laser

desorption ionization time-of-flight; Orn, ornithine; Tyr, tyrosine

REFERENCES

(1) Creighton, T. E. (1993) *Proteins. Structure and Molecular Properties*, 2nd ed., W. H. Freeman and Co., New York.

(2) Fersht, A. R. (1999) *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman and Co., New York.

(3) Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science 181*, 223−230.

(4) Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. U. S. A. 103*, 16623−16633.

(5) Dill, K. A., Ozkan, S. B., Shell, M. S., and Weikl, T. R. (2008) The protein folding problem. *Annu. Rev. Biophys. 37*, 289−316.

(6) Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry 29*, 7133−7155.

(7) Makhatadze, G. I., and Privalov, P. L. (1995) Energetics of protein structure. *Adv. Protein Chem. 47*, 307−425.

(8) Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins. *FASEB J. 10*, 75−83.

(9) Pey, A. L., Rodriguez-Larrea, D., Gavira, J. A., Garcia-Moreno, B., and Sanchez-Ruiz, J. M. (2010) Modulation of buried ionizable groups in proteins with engineered surface charge. *J. Am. Chem. Soc. 132*, 1218−1219.

(10) Gong, H., and Freed, K. F. (2010) Electrostatic solvation energy for two oppositely charged ions in a solvated protein system: Salt bridges can stabilize proteins. *Biophys. J. 98*, 470−477.

(11) Zhu, S., and Elcock, A. H. (2010) A Complete thermodynamic characterization of electrostatic and hydrophobic associations in the temperature range 0 to 100 °C from explicit-solvent molecular dynamics simulations. *J. Chem. Theory Comput. 6*, 1293−1306.

(12) Liao, S., and Green, M. E. (2011) Quantum calculations on salt bridges with water: Potentials, structures, and properties. *Comput. Theor. Chem. 963*, 207−214.

(13) Donald, J. E., Kulp, D. W., and DeGrado, W. F. (2011) Salt bridges: Geometrically specific, designable interactions. *Proteins 79*, 898−915.

(14) Elcock, A. H. (1998) The stability of salt bridges at high temperatures: Implications for hyperthermophilic proteins. *J. Mol. Biol. 284*, 489−502.

(15) Yip, K. S. P., Britton, K. L., Stillman, T. J., Lebbink, J., De Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D., and Rice, D. W. (1998) Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem. 255*, 336−346.

(16) Malakauskas, S. M., and Mayo, S. L. (1998) Design, structure and stability of a hyperthermophilic protein variant. *Nat. Struct. Biol. 5*, 470−475.

(17) Xiao, L., and Honig, B. (1999) Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol. 289*, 1435−1444.

(18) Pace, C. N. (2000) Single surface stabilizer. *Nat. Struct. Biol. 7*, 345−346.

(19) Karshikoff, A., and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot roads. *Trends Biochem. Sci. 26*, 550−556.

(20) Kumar, S., and Nussinov, R. (2002) Close-range electrostatic interactions in proteins. *ChemBioChem 3*, 604−617.

(21) Dominy, B. N., Minoux, H., and Brooks, C. L., III (2004) An electrostatic basis for the stability of thermophilic proteins. *Proteins 57*, 128−141.

(22) Thomas, A. S., and Elcock, A. H. (2004) Molecular simulations suggest protein salt bridges are uniquely suited to life at high temperatures. *J. Am. Chem. Soc. 126*, 2208−2214.

(23) Robinson-Rechavi, M., Alibés, A., and Godzik, A. (2006) Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: Lessons from structural genomics of *Thermotoga maritima*. *J. Mol. Biol. 356*, 547−557.

(24) Su, J. G., Chen, W. Z., and Wang, C. X. (2010) Role of electrostatic interactions for the stability and folding behavior of cold shock protein. *Proteins 78*, 2157−2169.

(25) Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2005) Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability. *Biochemistry 44*, 1414−1422.

(26) Gauba, V., and Hartgerink, J. D. (2007) Surprisingly high stability of collagen ABC heterotrimer: evaluation of side chain charge pairs. *J. Am. Chem. Soc. 129*, 15034−15041.

(27) Fallas, J. A., Gauba, V., and Hartgerink, J. D. (2009) Solution structure of an ABC collagen heterotrimer reveals a single-register helix stabilized by electrostatic interactions. *J. Biol. Chem. 284*, 26851−26859.

(28) Xu, F., Zahid, S., Silva, T., and Nanda, V. (2011) Computational design of a collagen A:B:C-type heterotrimer. *J. Am. Chem. Soc. 133*, 15260−15263.

(29) Padmanabhan, S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1996) Helix propensities of basic amino acids increase with the length of the side-chain. *J. Mol. Biol. 257*, 726−734.

(30) Merutka, G., and Stellwagen, E. (1991) Effect of amino-acid ion-pairs on peptide helicity. *Biochemistry 30*, 1591−1594.

(31) Cheng, R. P., Girinath, P., and Ahmad, R. (2007) Effect of lysine side chain length on intra-helical glutamate-lysine ion pairing interactions. *Biochemistry 46*, 10528−10537.

(32) Huyghues-Despointes, B. M. P., Scholtz, J. M., and Baldwin, R. L. (1993) Effect of a single aspartate on helix stability at different positions in a neutral alanine-based peptide. *Protein Sci. 2*, 1604−1611.

(33) Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry 32*, 9668−9676.

(34) Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix propensities of the amino-acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci. 3*, 843−852.

(35) Doig, A. J., and Baldwin, R. L. (1995) N- and C-capping preferences for all 20 amino acids in α-helical peptides. *Protein Sci. 4*, 1325−1336.

(36) Marqusee, S., Baldwin, R. L. (1990) α-Helix formation by short peptides in water, in *Protein Folding* (Gierasch, L. M., King, J., Eds.) pp 85−94, American Association for the Advancement of Science, Washington D.C.

(37) Huyghues-Despointes, B. M. P., Scholtz, J. M., and Baldwin, R. L. (1993) Helical peptides with 3 pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci. 2*, 80−85.

(38) Smith, J. S., and Scholtz, J. M. (1998) Energetics of polar side-chain interactions in helical peptides: Salt effects on ion pairs and hydrogen bonds. *Biochemistry 37*, 33−40.

(39) Atherton, E., Fox, H., Harkiss, D., Logan, C. J., Sheppard, R. C., and Williams, B. J. (1978) Mild procedure for solid-phase peptide-synthesis - use of fluorenylmethoxycarbonylamino-acids. *J. Chem. Soc. Chem. Commun.*, 537−539.

(40) Fields, G. B., and Noble, R. L. (1990) Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbonyl amino-acids. *Int. J. Pept. Protein Res. 35*, 161−214.

(41) Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*, 1948−1954.

(42) Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption-coefficient of a protein. *Protein Sci. 4*, 2411−2423.

(43) Lifson, S., and Roig, A. (1961) The theory of helix-coil transition in polypeptides. *J. Chem. Phys. 34*, 1963−1974.

(44) Doig, A. J., Chakrabartty, A., Klingler, T. M., and Baldwin, R. L. (1994) Determination of free-energies of N-capping in α-helices by modification of the Lifson-Roig helix-coil theory to include N-capping and C-capping. *Biochemistry 33*, 3396−3403.

(45) Chiu, H. P., Suzuki, Y., Gullickson, D., Ahmad, R., Kokona, B., Fairman, R., and Cheng, R. P. (2006) Helix propensity of highly fluorinated amino acids. *J. Am. Chem. Soc. 128*, 15556−15557.
(46) Robert, C. H. (1990) A hierarchical nesting approach to describe the stability of α-helices with side-chain interactions. *Biopolymers 30*, 335−347.
(47) Hobohm, U., and Sander, C. (1994) Enlarged representative set of protein structures. *Protein Sci. 3*, 522−524.
(48) Griep, S., and Hobohm, U. (2010) PDBselect 1992−2009 and PDBfilter-select. *Nucleic Acids Res. 38*, D318−D319.
(49) Gunasekaran, K., Nagarajaram, H. A., Ramakrishnan, C., and Balaram, P. (1998) Stereochemical punctuation marks in protein structures: Glycine and proline containing helix stop signals. *J. Mol. Biol. 275*, 917−932.
(50) Engel, D. E., and DeGrado, W. F. (2004) Amino acid propensities are position-dependent throughout the length of α-helices. *J. Mol. Biol. 337*, 1195−1205.
(51) Cheng, R. P., Girinath, P., Suzuki, Y., Kuo, H.-T., Hsu, H.-C., Wang, W.-R., Yang, P.-A., Gullickson, D., Wu, C.-H., Koyack, M. J., Chiu, H.-P., Weng, Y.-J., Hart, P., Kokona, B., Fairman, R., Lin, T.-E., and Barrett, O. (2010) Positional effects on helical Ala-based peptides. *Biochemistry 49*, 9372−9384.
(52) Cheng, R. P., Weng, Y.-J., Wang, W.-R., Koyack, M. J., Suzuki, Y., Wu, C.-H., Yang, P.-A., Hsu, H.-C., Kuo, H.-T., Girinath, P., and Fang, C.-J. (2012) Helix formation and capping energetics of arginine analogs with varying side chain length. *Amino Acids 43*, 195−206.
(53) Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577−2637.
(54) Efron, B., and Gong, G. (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation. *Amer. Stat. 37*, 36−48.
(55) Kuebler, R. R., Smith, H. (1976) Statistics-A beginning, in *Statistics-A Beginning*, p 302 John Wiley & Sons, Inc, New York.
(56) Klugh, H. E. (1970) Statistics-The essentials for research, in *Statistics-The Essentials for Research*, p 350, John Wiley & Sons, Inc, New York.
(57) Marqusee, S., and Baldwin, R. L. (1987) Helix stabilization by Glu- ... Lys+ salt bridges in short peptides of denovo design. *Proc. Natl. Acad. Sci. U. S. A. 84*, 8898−8902.
(58) Chakrabartty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) Aromatic side-chain contribution to far-ultraviolet circular-dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry 32*, 5560−5565.
(59) Chang, C. T., Wu, C. S. C., and Yang, J. T. (1978) Circular dichroic analysis of protein conformation - inclusion of the β-turns. *Anal. Biochem. 91*, 13−31.
(60) Sali, D., Bycroft, M., and Fersht, A. R. (1988) Stabilization of protein structure by interaction of α-helix dipole with a charged side chain. *Nature 355*, 740−743.
(61) Johnsson, K., Allemann, R. K., Widmer, H., and Benner, S. A. (1993) Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. *Nature 365*, 530−532.
(62) Munoz, V., and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters 0.2. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol. 245*, 275−296.
(63) Jeffrey, G. A., Saenger, W. (1994) in *Hydrogen Bonding in Biological Structures*, pp 365−371, Springer-Verlag, Heidelberg.
(64) Jeffrey, G. A., Saenger, W. (1994) Chapter 7 Metrical aspects of two-center hydrogen bonds, in *Hydrogen Bonding in Biological Structures*, pp 111−135, Springer-Verlag, Heldelberg.
(65) Andrew, C. D., Penel, S., Jones, G. R., and Doig, A. J. (2001) Stabilizing nonpolar/polar side-chain interactions in the a-helix. *Proteins: Struct. Funct. Genet. 45*, 449−455.
(66) Groebke, K., Renold, P., Tsang, K. Y., Allen, T. J., McClure, K. F., and Kemp, D. S. (1996) Template-nucleated alanine-lysine helices are stabilized by position-dependent interactions between the lysine side chain and the helix barrel. *Proc. Natl. Acad. Sci. U. S. A. 93*, 4025−4029.
(67) Luo, P., and Baldwin, R. L. (1999) Interaction between water and polar groups of the helix backbone: An important determinant of helix propensities. *Proc. Natl. Acad. Sci. U. S. A. 96*, 4930−4935.
(68) Garcia, A. E., and Sanbonmatsu, K. Y. (2002) α-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc. Natl. Acad. Sci. U. S. A. 99*, 2782−2787.
(69) McGregor, M. J., Islam, S. A., and Sternberg, M. J. E. (1987) Analysis of the relationship between side-chain conformation and secondary structure in globular-proteins. *J. Mol. Biol. 198*, 295−310.
(70) Dunbrack, R. L., and Karplus, M. (1993) Backbone-dependent rotamer library for proteins - application to side-chain prediction. *J. Mol. Biol. 230*, 543−574.
(71) Maxfield, F. R., and Scheraga, H. A. (1975) Effect of neighboring charges on helix forming ability of charged amino-acids in proteins. *Macromolecules 8*, 491−493.
(72) Barlow, D. J., and Thornton, J. M. (1983) Ion-pairs in proteins. *J. Mol. Biol. 168*, 867−885.
(73) Singh, J., Thornton, J. M., Snarey, M., and Campbell, S. F. (1987) The geometries of interacting arginine-carboxyls in proteins. *FEBS Lett. 224*, 161−171.
(74) Chakrabarti, P. (1994) Conformations of arginine and lysine side chains in association with anions. *Int. J. Peptide Protein Res. 43*, 284−291.
(75) Ghadiri, M. R., and Fernholz, A. K. (1990) Peptide architecture. Design of stable α-helical metallopeptides via a novel exchange-inert Ru^{III} complex. *J. Am. Chem. Soc. 112*, 9633−9635.
(76) Jackson, D. Y., King, D. S., Chmielewski, J., Singh, S., and Schultz, P. G. (1991) General approach to the synthesis of short α-helical peptides. *J. Am. Chem. Soc. 113*, 9391−9392.
(77) Mills, N. L., Daugherty, M. D., Frankel, A. D., and Guy, R. K. (2006) An α-helical peptidomimetic inhibitor of the HIV-1 Rev-RRE interaction. *J. Am. Chem. Soc. 128*, 3496−3497.
(78) Phelan, J. C., Skelton, N. J., Braisted, A. C., and McDowell, R. S. (1997) A general method for constraining short peptides to an α-helical conformation. *J. Am. Chem. Soc. 119*, 455−460.
(79) Judice, J. K., Tom, J. Y. K., Huang, W., Wrin, T., Vennari, J., Petropoulos, C. J., and McDowell, R. S. (1997) Inhibition of HIV type 1 infectivity by constrained α-helical peptides: Implications for the viral fusion mechanism. *Proc. Natl. Acad. Sci. U. S. A. 94*, 13426−13430.
(80) Kajino, M., Fujimoto, K., and Inouye, M. (2011) Side-chain cross-linked short α-helices that behave like original proteins in biomacromolecular interactions. *J. Am. Chem. Soc. 133*, 656−659.
(81) Schafmeister, C. E., Po, J., and Verdine, G. L. (2000) An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc. 122*, 5891−5892.
(82) Kim, Y.-W., Kutchukian, P. S., and Verdine, G. L. (2010) Introduction of all-hydrocarbon *i*,*i*+3 staples into α-helices via ring-closing olefin metathesis. *Org. Lett. 12*, 3046−3049.
(83) Kim, Y.-W., Grossmann, T. N., and Verdine, G. L. (2011) Synthesis of all-hydrocarbon stapled α-helical peptides by ring-closing olefin metathesis. *Nat. Protocols 6*, 761−771.
(84) Kumita, J. R., Flint, D. G., Smart, O. S., and Woolley, G. A. (2002) Photo-control of peptide helix content by an azobenzene cross-linker: Steric interactions with underlying residues are not critical. *Protein Eng. 15*, 561−569.
(85) Pepe-Mooney, B. J., and Fairman, R. (2009) Peptides as materials. *Curr. Opin. Struct. Biol. 19*, 483−494.
(86) Lu, H., Wang, J., Bai, Y., Lang, J. W., Liu, S., Lin, Y., and Cheng, J. (2011) Ionic polypeptides with unusual helical stability. *Nat. Commun. 2*, 206.
(87) Altman, M., Lee, P., Rich, A., and Zhang, S. G. (2000) Conformational behavior of ionic self-complementary peptides. *Protein Sci. 9*, 1095−1105.
(88) Hong, Y. S., Legge, R. L., Zhang, S., and Chen, P. (2003) Effect of amino acid sequence and pH on nanofiber formation of self-assembling peptides EAK16-II and EAK16-IV. *Biomacromolecules 4*, 1433−1442.

(89) Kuehne, J., and Murphy, R. M. (2001) Synthesis and characterization of membrane-active GALA-OKT9 conjugates. *Bioconjugate Chem. 12*, 742−749.
(90) Haas, D. H., and Murphy, R. M. (2004) Design of a pH-sensitive pore-forming peptide with improved performance. *J. Pept. Res. 63*, 9−16.
(91) Haas, D. H., and Murphy, R. M. (2004) Templated assembly of the pH-sensitive membrane-lytic peptide GALA. *J. Pept. Res. 63*, 451−459.
(92) Turk, M. J., Reddy, J. A., Chmielewski, J. A., and Low, P. S. (2002) Characterization of a novel pH-sensitive peptide that enhances drug release from folate-targeted liposomes at endosomal pHs. *Biochim. Biophys. Acta Biomembr. 1559*, 56−68.
(93) Zimenkov, Y., Dublin, S. N., Ni, R., Tu, R. S., Breedveld, V., Apkarian, R. P., and Conticello, V. P. (2006) Rational design of a reversible pH-responsive switch for peptide self-assembly. *J. Am. Chem. Soc. 128*, 6770−6771.
(94) Sheparovych, R., Roiter, Y., Yang, J., Kopecek, J., and Minko, S. (2009) Stimuli-responsive properties of peptide-based copolymers studied via directional growth of self-assembled patterns on solid substrate. *Biomacromolecules 10*, 1955−1961.
(95) Hirosue, S., and Weber, T. (2006) pH-Dependent lytic peptides discovered by phage display. *Biochemistry 45*, 6476−6487.
(96) Pagel, K., and Koksch, B. (2008) Following polypeptide folding and assembly with conformational switches. *Curr. Opin. Chem. Biol. 12*, 730−739.

Supporting Information

Effect of Glutamate Side Chain Length on Intrahelical Glutamate-Lysine Ion Pairing Interactions

Richard P. Cheng,* Wei-Ren Wang, Prashant Girinath, Po-An Yang, Raheel Ahmad, Jhe-Hao Li, Pier Hart, Bashkim Kokona, Robert Fairman, Casey Kilpatrick, Annmarie Argiros

Tables

Figures

Table S1. Mean Residue Ellipticity at 22 nm for the ZbbXaaN Peptides at pH 7 and 2

	$[\theta]_{222\ nm}$ (deg cm^2 $dmol^{-1}$)	
Peptide	pH 7	pH 2
AspLys5	-18200 ± 300	-10800 ± 300
AspOrn5	-9700 ± 500	-5300 ± 500
AspDab5	-7400 ± 500	-3900 ± 500
AspDap5	-4200 ± 400	-2700 ± 500
AspLys4	-26800 ± 1000	-8600 ± 400
AspOrn4	-25800 ± 900	-8400 ± 300
AspDab4	-13300 ± 300	-5400 ± 200
AspDap4	-6400 ± 200	-4100 ± 300
AspLys3	-9200 ± 200	-6000 ± 100
AspOrn3	-4300 ± 600	-3900 ± 200
AspDab3	-3600 ± 100	-2500 ± 100
AspDap3	-1800 ± 700	-1800 ± 100
GluLys5	-24600 ± 400	-23200 ± 300
GluOrn5	-17100 ± 300	-17200 ± 400
GluDab5	-11100 ± 500	-12100 ± 700
GluDap5	-2900 ± 200	-2700 ± 200
GluLys4	-24400 ± 300	-22800 ± 300
GluOrn4	-27000 ± 200	-15900 ± 200
GluDab4	-24300 ± 600	-14400 ± 400
GluDap4	-12500 ± 300	-5100 ± 200
GluLys3	-16400 ± 100	-13400 ± 200
GluOrn3	-6100 ± 100	-7100 ± 300
GluDab3	-4300 ± 600	-3900 ± 200
GluDap3	-1600 ± 100	-1500 ± 300
AadLys5	-25500 ± 300	-21100 ± 200
AadOrn5	-20000 ± 300	-16700 ± 200
AadDab5	-11700 ± 200	-7900 ± 200
AadDap5	-6000 ± 200	-4000 ± 200
AadLys4	-25800 ± 500	-22100 ± 400
AadOrn4	-23200 ± 600	-17500 ± 500
AadDab4	-20600 ± 200	-13800 ± 200
AadDap4	-12500 ± 300	-8100 ± 300
AadLys3	-20200 ± 200	-18800 ± 100
AadOrn3	-9500 ± 200	-11800 ± 200
AadDab3	-6900 ± 200	-7200 ± 300
AadDap3	-3200 ± 100	-3300 ± 500

Table S2. Energetics[a] of Zbb-Xaa (*i*, *i*+4) Intrahelical Interactions at pH 7 and 2

Residue *i*	Residue *i*+*n*	*n*	$\Delta G_{pH\ 7\ interaction}$ (cal·mol^{-1})	$\Delta G_{pH\ 2\ interaction}$ (cal·mol^{-1})	$\Delta\Delta G$[d] (cal·mol^{-1})
Asp	Lys	5	0[b]	ND[c]	
Asp	Orn	5	ND[c]	ND[c]	
Asp	Dab	5	ND[c]	ND[c]	
Asp	Dap	5	ND[c]	ND[c]	
Glu	Lys	5	-84.7 ± 51.4[e]	-21.2 ± 15.6[e]	-63.7 ± 53.1[e]
Glu	Orn	5	0[b,e]	0[b,e]	
Glu	Dab	5	ND[c,e]	ND[c,e]	
Glu	Dap	5	ND[c,e]	ND[c,e]	
Aad	Lys	5	-67.2 ± 59.8	-42.3 ± 37.1	-24.9 ± 70.4
Aad	Orn	5	-16.4 ± 67.7	0	-16.4 ± 67.7
Aad	Dab	5	ND[c]	ND[c]	
Aad	Dap	5	ND[c]	ND[c]	
Asp	Lys	4	-736 ± 24	0[b]	-736 ± 24
Asp	Orn	4	-1090 ± 40	-217 ± 93	-873 ± 101
Asp	Dab	4	-713 ± 31	-133 ± 43	-580 ± 53
Asp	Dap	4	-547 ± 34	0[b]	-547 ± 34
Glu	Lys	4	-459 ± 34[e]	-194 ± 31[e]	-265 ± 46[e]
Glu	Orn	4	-742 ± 43[e]	-125 ± 34[e]	-617 ± 53[e]
Glu	Dab	4	-837 ± 36[e]	-281 ± 38[e]	-556 ± 52[e]
Glu	Dap	4	-781 ± 43[e]	0[b,e]	-781 ± 43[e]
Aad	Lys	4	-130 ± 24	0[b]	-130 ± 24
Aad	Orn	4	-243 ± 75	-33.1 ± 38.0	-209 ± 84
Aad	Dab	4	-531 ± 78	-143 ± 34	-388 ± 85
Aad	Dap	4	-473 ± 48	-96.2 ± 43.0	-377 ± 64
Asp	Lys	3	-146 ± 36	0[b]	-146 ± 36
Asp	Orn	3	ND[c]	ND[c]	
Asp	Dab	3	ND[c]	ND[c]	
Asp	Dap	3	ND[c]	ND[c]	
Glu	Lys	3	-256 ± 36[e]	-143 ± 32[e]	-113 ± 48[e]
Glu	Orn	3	0[b,e]	ND[c,e]	
Glu	Dab	3	0[b,e]	ND[c,e]	
Glu	Dap	3	0[b,e]	ND[c,e]	
Aad	Lys	3	-46.9 ± 30.3	-70.1 ± 25.2	23.2 ± 39.4
Aad	Orn	3	ND[c]	ND[c]	
Aad	Dab	3	ND[c]	ND[c]	
Aad	Dap	3	ND[c]	ND[c]	

[a]The energetic values $\Delta G_{pH\ 7\ interaction}$ and $\Delta G_{pH\ 2\ interaction}$ were derived from the experimental CD data based on the nesting block method (*1-3*). [b]The experimental results are within error of theoretically calculated values without any side chain-side chain interaction. [c]Not determined because the experimental results exhibit lower helical content compared to theoretically calculated values without any side chain-side chain interaction (see main text for reason). [d]$\Delta\Delta G = \Delta G_{pH\ 7\ interaction} - \Delta G_{pH\ 2\ interaction}$. The difference in side chain-side chain interaction energy at pH 7 and 2 reflects the contribution of electrostatics in the Glu-Xaa interaction. [e]Values as previously reported (*3*).

Table S3. Summary of Low Energy Conformations from Conformational Analysis of MSZbbXaaN Peptides by Molecular Mechanics Calculations

Peptide	Lowest Energy Conformer	Conformations within 4 kcal of the Lowest Energy Conformer			
	Energy (kcal)	No.[a]	Weighed Average Energy (kcal)	Salt Bridge[b] (%)	Major conformations Number(Zbb χ_1, Xaa χ_1)[c]
MSAspLys3	-52.0	11	-51.6	100%	11(*g*+, *g*+)
MSAspLys4	-56.8	3	-56.6	100%	3(*g*-, *g*+)
MSAspLys5	-44.1	44	-44.5	0%	18(*g*-, *t*), 17(*g*-, *g*+), 7(*g*+, *g*+)
MSAspOrn3	-45.3	18	-44.5	0%	10(*g*-, *g*+), 5(*g*+, *g*+), 3(*g*-, *t*)
MSAspOrn4	-49.6	5	-49.4	60%	3(*g*-, *g*+), 2(*t*, *g*+)
MSAspOrn5	-41.8	23	-41.4	0%	9(*g*-, *t*), 8(*g*-, *g*+), 6(*g*+, *g*+)
MSAspDab3	-40.8	26	-40.5	0%	7(*g*+, *t*), 7(*g*+, *g*+), 6(*g*-, *g*+), 6(*g*-, *t*)
MSAspDab4	-43.5	12	-43.3	17%	5(*g*-, *g*+), 4(*g*-, *t*)
MSAspDab5	-39.3	12	-39.0	0%	3(*g*-, *g*+), 3(*g*+, *t*), 3(*g*+, *g*+), 3(*g*-, *t*)
MSAspDap3	-35.6	14	-35.3	0%	3(*g*+, *t*), 3(*g*+, *g*-), 2(*g*-, *g*+), 2(*g*+, *g*+), 2(*g*-, *t*), 2(*g*-, *g*-)
MSAspDap4	-37.7	10	-37.4	0%	2(*g*-, *g*+), 2(*g*+, *g*-), 2(*g*-, *t*), 2(*g*-, *g*-)
MSAspDap5	-34.6	10	-34.5	0%	3(*g*+, *g*-), 3(*g*+, *g*+)
MSGluLys3	-50.1	11	-49.6	100%	9(*g*+, *g*+)
MSGluLys4	-50.1	24	-49.6	67%	18(*g*-, *g*+), 4(*t*, *g*+)
MSGluLys5	-43.5	41	-42.9	0%	17(*g*-, *g*+), 17(*g*-, *t*), 17(*g*+, *g*+)
MSGluOrn3	-43.4	26	-42.9	19%	11(*g*-, *g*+), 7(*g*+, *g*+), 5(*g*-, *t*)
MSGluOrn4	-45.8	18	-45.3	56%	13(*g*-, *g*+), 5(*t*, *g*+)
MSGluOrn5	-40.0	17	-39.7	0%	9(*g*-, *t*), 8(*g*-, *g*+)
MSGluDab3	-39.5	19	-39.2	0%	6(*g*-, *g*+), 6(*g*-, *t*), 4(*g*+, *g*+), 3(*g*+, *t*)
MSGluDab4	-43.1	10	-43.0	50%	5(*g*-, *g*+), 4(*t*, *g*+)
MSGluDab5	-37.5	10	-37.3	0%	3(*g*-, *g*+), 3(*g*+, *g*+), 3(*g*-, *t*)
MSGluDap3	-34.3	9	-34.0	0%	2(*g*-, *g*+), 2(*g*-, *t*), 2(*g*-, *g*-)
MSGluDap4	-35.8	10	-35.7	0%	2(*g*-, *g*+), 2(*g*-, *t*), 2(*g*-, *g*-)
MSGluDap5	-32.9	6	-32.8	0%	
MSAadLys3	-55.8	21	-55.6	100%	21(*g*+, *g*+)
MSAadLys4	-54.5	60	-54.0	100%	24(*t*, *g*+), 22(*g*-, *g*+), 14(*g*+, *g*+)
MSAadLys5	-47.9	89	-47.4	0%	49(*g*-, *g*+), 34 (*g*-, *t*), 5(*g*+, *g*+)
MSAadOrn3	-50.7	14	-50.5	100%	14 (*g*+, *g*+)
MSAadOrn4	-49.6	33	-49.1	100%	21(*t*, *g*+), 9(*g*-, *g*+), 3(*g*+, *g*+)
MSAadOrn5	-44.3	54	-43.9	0%	26(*g*-, *g*+), 23 (*g*-, *t*), 5(*g*+, *g*+)
MSAadDab3	-45.3	16	-45.2	50%	7(*g*+, *g*+), 3(*t*, *t*), 2(*g*-, *g*+), 2(*g*+, *t*)
MSAadDab4	-45.2	55	-44.7	71%	32(*t*, *g*+), 10(*g*-, *g*+), 5(*g*+, *g*+), 4(*g*-, *t*)
MSAadDab5	-41.8	25	-41.5	0%	12(*g*-, *t*), 10(*g*-, *g*+)
MSAadDap3	-38.0	43	-37.2	12%	9(*g*+, *g*-), 8(*g*+, *g*+), 6(*g*-, *g*+), 6(*g*+, *t*), 6(*g*-, *t*), 6(*g*-, *g*-)
MSAadDap4	-39.5	41	-38.8	20%	7(*t*, *g*+), 7(*g*+, *g*-), 7(*g*+, *g*+), 6(*g*-, *t*), 5(*g*-, *g*+), 5(*g*-, *g*-)
MSAadDap5	-37.2	17	-37.0	0%	6(*g*-, *g*+), 5(*g*-, *g*-)

[a]The number of conformations within 5 kcal of the lowest energy conformer for each peptide. [b]The percentage of conformations within 5 kcal of the lowest energy conformer with a Zbb-Xaa salt bridge, which is a hydrogen bonded ion pair (*4*, *5*).

Table S4. Survey Results of Oppositely Charged Residues with Various Spacings in Natural Protein Helices[a]

Position *i*	Position *i+n*	*n*	Occurrence[a]	Helix Pair Propensity[b]	Z value[c]	P value[d]	Salt Bridge Occurrence[e]
Glu	Lys	3	1940	1.95 ± 0.06	30.0	9.81×10^{-198}	226
Asp	Lys	3	1054	1.37 ± 0.05	10.3	1.18×10^{-24}	49
Glu	Lys	4	1887	2.12 ± 0.07	33.4	3.74×10^{-244}	124
Asp	Lys	4	918	1.46 ± 0.06	11.6	2.90×10^{-31}	95
Glu	Lys	5	809	1.36 ± 0.06	8.69	3.62×10^{-18}	0
Asp	Lys	5	460	0.88 ± 0.04	-2.69	7.15×10^{-3}	2

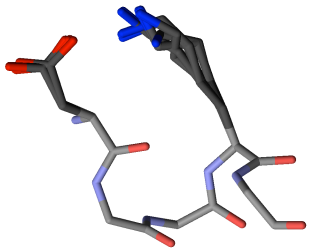
[a]The occurrence of oppositely charged residues placed three, four, and five residues apart in helices six residues or longer in the non-redundant protein structure database PDBselect (April 2009, 25% threshold) (*6, 7*). The helix conformation is defined by the backbone dihedral angles (ϕ, ψ) according to the values provided by Balaram and coworkers (*8, 9*). [b]The occurrence for the various sequence patterns divided by the corresponding expected value. The expected value was obtained by bootstrapping the complete PDBselect database, thereby including the bias for the occurrence of each residue for the helix conformation. Since bootstrapping was performed 100,000 times, this enabled the calculation of a standard deviation for the expected value and thus helix pair propensity. [c]The difference between the occurrence and the expected value divided by the standard deviation for the expected value. In other words, the number of standard deviations that separate the occurrence and the expexted value; a positive number means the occurrence is larger than the expected value, whereas a negative number means the opposite. [d]The probability that the occurrence and expected occurrence are the same based on the standard deviation obtained from bootstrapping assuming a Gaussian distribution. [e]The number of the occurrences for each sequence pattern that actually form salt bridges between the oppositely charged side chain functionalities. The presence of a salt bridge is defined by an N–O distance less than or equal to 3 Å.

Table S5. Survey Results of Oppositely Charged Residues with Various Spacings in Natural Protein Helices[a]

Position *i*	Position *i+n*	*n*	Occurrence[a]	Pair Propensity[b]	Z value[c]	P value[d]	Salt Bridge Occurrence[e]
Glu	Lys	3	1940	1.31 ± 0.03	11.3	1.31×10^{-29}	226
Asp	Lys	3	1054	1.44 ± 0.05	10.5	1.07×10^{-25}	49
Glu	Lys	4	1887	1.38 ± 0.04	14.5	1.62×10^{-47}	124
Asp	Lys	4	918	1.35 ± 0.05	8.71	3.04×10^{-18}	95
Glu	Lys	5	809	0.66 ± 0.02	-12.6	3.09×10^{-36}	0
Asp	Lys	5	460	0.76 ± 0.03	-5.87	4.36×10^{-9}	2

[a]The occurrence of oppositely charged residues placed three, four, and five residues apart in helices six residues or longer in the non-redundant protein structure database PDBselect (April 2009, 25% threshold) (*6, 7*). The helix conformation is defined by the backbone dihedral angles (ϕ, ψ) according to the values provided by Balaram and coworkers (*8, 9*). [b]The occurrence for the various sequence patterns divided by the corresponding expected value. The expected value was obtained by bootstrapping the helices in the PDBselect database, thereby removing the bias for the occurrence of each residue for the helix conformation. Since bootstrapping was performed 100,000 times, this enabled the calculation of a standard deviation for the expected value and thus helix pair propensity. [c]The difference between the occurrence and the expected value divided by the standard deviation for the expected value. In other words, the number of standard deviations that separate the occurrence and the expexted value; a positive number means the occurrence is larger than the expected value, whereas a negative number means the opposite. [d]The probability that the occurrence and expected occurrence are the same based on the standard deviation obtained from bootstrapping assuming a Gaussian distribution. [e]The number of the occurrences for each sequence pattern that actually form salt bridges between the oppositely charged side chain functionalities. The presence of a salt bridge is defined by an N–O distance less than or equal to 3 Å.

MSAspLys3



(*g*+, *g*+)

Figure S1. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSAspLys3 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Asp)– χ_1(Lys) combination is shown below the overlaid structures.

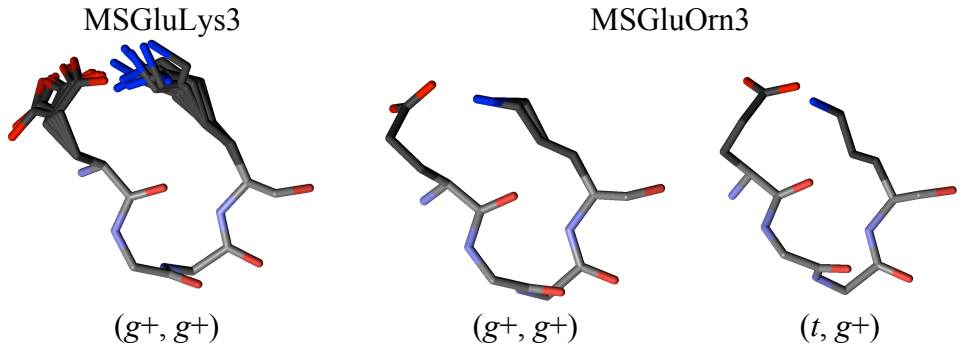


Figure S2. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSGluLys3 and MSGluOrn3 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Glu)– χ_1(Lys/Orn) combination is shown below the overlaid structures.

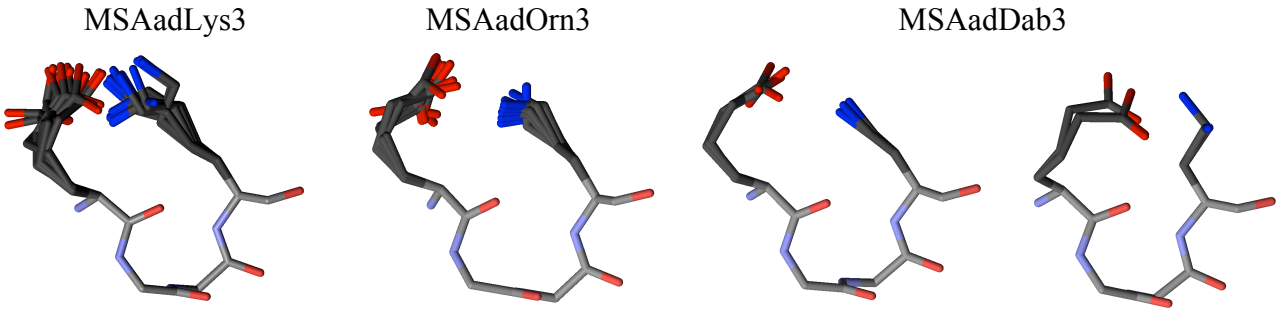


Figure S3. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSAadLys3, MSAadOrn3, and AadDab3 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Aad)– χ_1(Lys/Orn/Dab) combination is shown below the overlaid structures.

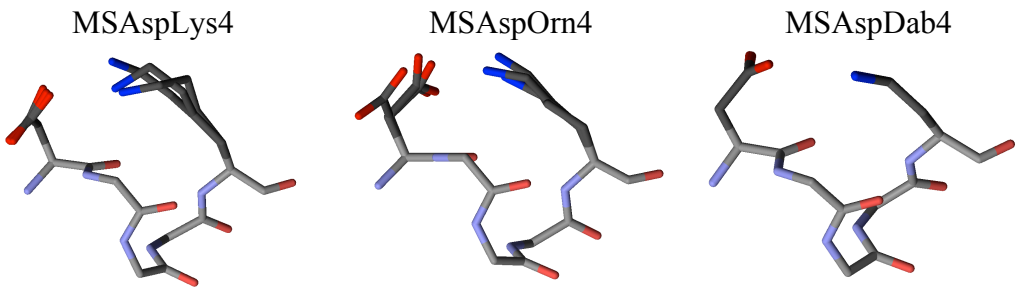


Figure S4. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSAspLys4, MSAspOrn4, and MSAspDab4 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Asp)– χ_1(Lys/Orn/Dab) combination is shown below the overlaid structures.

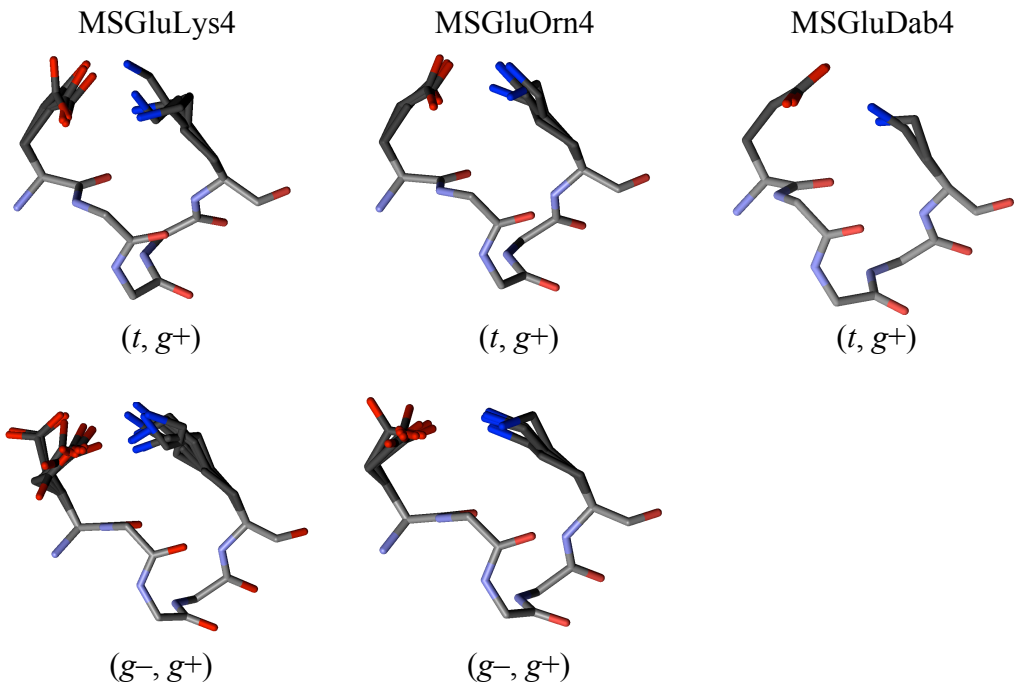


Figure S5. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSGluLys4, MSGluOrn4, and MSGluDab4 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Glu)– χ_1(Lys/Orn/Dab) combination is shown below the overlaid structures.

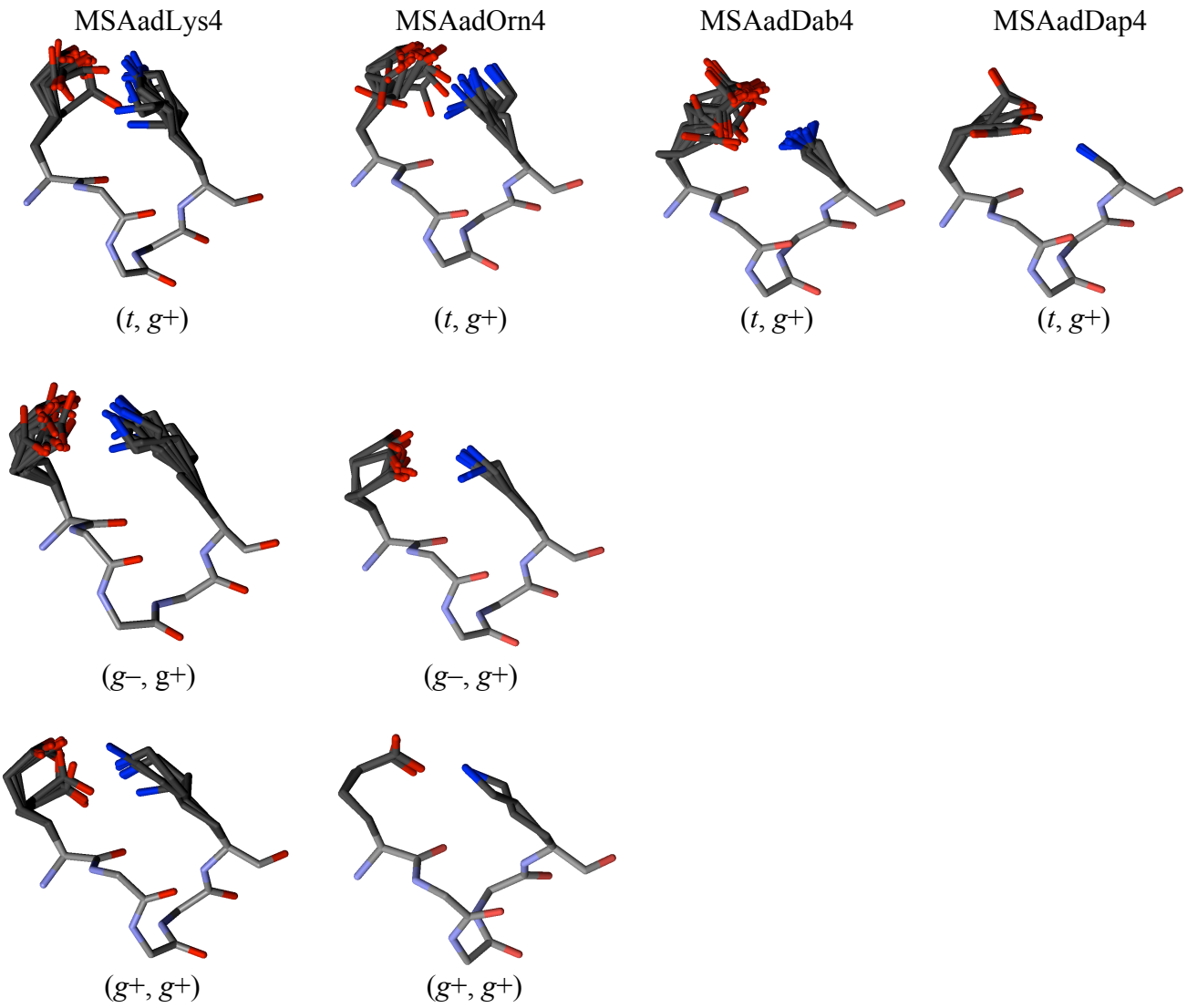


Figure S6. Overlay of the three-dimensional structures of the major low energy conformers from the conformational analysis for MSAadLys4, MSAadOrn4, MSAadDab4, and MSAadDap4 by molecular mechanics calculations. The oxygen atoms are shown in red, the nitrogens are shown in blue, and the carbons are shown in gray. The χ_1(Aad)– χ_1(Lys/Orn/Dab/Dap) combination is shown below the overlaid structures.

Material and Methods

General Section

All of the chemical reagents except those indicated otherwise were purchased from Aldrich. Organic and high performance liquid chromatography (HPLC) solvents were from EMD Science or Merck Taiwan. Fmoc amino acids, 1-hydroxybenzotrazole (HOBt), and O-1H-benzotriazol-1-yl-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were from Novabiochem, and the Fmoc-PAL-PEG-PS resin was from Applied Biosystems. Reagents and solvents were used without further purification. Analytical reverse phase (RP)-HPLC was performed on an Agilent 1100 series chromatography system using a Vydac C_{18} column (4.6 mm diameter, 250 mm length). Preparative RP-HPLC was performed on a Waters Breeze chromatography system using a Vydac C_4 and C_{18} column (22 mm diameter, 250 mm length). Mass spectrometry of the peptides was performed on a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) spectrometer (Bruker Daltonics Biflex IV) using α-cyano-4-hydroxycinnamic acid as the matrix. Determination of peptide concentration was performed on a UV-Vis spectrophotometer (Agilent 8453). Circular dichroism (CD) spectra were collected on a Jasco J715 spectrometer using 1 mm pathlength cell. Each reported CD value was the mean of three wavelength scans or the mean of 61 readings at 222 nm. Data was normalized in terms of per residue molar ellipticity ($deg \cdot cm^2 \cdot dmol^{-1}$).

Peptide Synthesis

Peptides were synthesized by solid phase peptide synthesis using Fmoc (N^{α}-fluorenylmethyloxycarbonyl)-based chemistry (*10*). For a typical peptide synthesis, Fmoc-PAL-PEG-PS (0.05 mmol) was swollen in *N, N*-dimethylformamide (DMF, 5 mL) for 30 minutes. The resin was then deprotected by 20% piperidine/DMF (5 mL, 3x8 min) and rinsed with DMF (5 mL, 5x1 min). A mixture of 3 equivalents of the appropriately protected Fmoc amino acid, HOBt, and HBTU was dissolved in DMF (1 mL). Diisopropylethylamine (DIEA, 8 equivalents) was then added to the solution and mixed thoroughly. The solution was then applied to the resin. The vial that contained the solution was rinsed with DMF (2x1 mL) and added to the reaction. The first coupling was carried out for 8 hours. The 8th to 14th residues were coupled for 1.5 hours. Other residues were coupled for 45 minutes. After each coupling, the resin was washed with DMF (5 mL, 5x1 min). For synthesis on the synthesizer (aapptec, Apex 396), procedures

were the same except different equivalents of reagents were used and different coupling times were applied. 25% piperidine/DMF, 5 equivalents of protected Fmoc amino acid, 9 equivalents of DIEA and 0.5 M of HBTU were used. Coupling time was doubled (i.e. 45 minutes to 90 minutes, 1.5 hours to 3 hours). For capping with acetic anhydride (Ac_2O), a solution of Ac_2O (20 equivalents), DIEA (20 equivalents), DMF (3 mL) was added to the resin. The reaction was shaken for 2 hours. The resin was subsequently washed with DMF (5 mL, 5x1 min) and methanol (5 mL), and was lyophilized overnight.

Peptides were deprotected and cleaved off the resin by treating the resin with 10 mL 95:5 trifluoroacetic acid (TFA)/triisopropylsilane and shaken for 2 hours. The solution was then filtered through glass wool and the resin was washed with TFA (3x3 mL). The combined filtrate was evaporated by a gentle stream of N_2. The resulting material was washed with hexanes (3x3 mL), dissolved in water, and lyophilized. The peptide (1 mg mL^{-1} aqueous solution) was analyzed using analytical RP-HPLC on a 25 cm C_{18} column (dia 4.6 mm) with flow rate 1 mL min^{-1}, temperature 25°C, linear 1 % min^{-1} gradient from 100% A to 0% A (solvent A: 99.9% water,0.1% TFA; solvent B: 90% acetonitrile, 10% water, 0.1% TFA). Appropriate linear solvent A/solvent B gradient was used for purification on preparative RP-HPLC on C_4 and C_{18} columns. The identity of the peptide was confirmed by MALDI-TOF.

AspLys5 (Ac-Tyr Gly Gly Asp Ala Ala Ala Ala Lys Asp Ala Ala Ala Ala Lys Asp Ala Ala Ala Ala Lys- NH_2). The peptide was synthesized using 0.1452 g (0.029 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1996 g of resin (85.0% yield). The cleavage yielded 64.2 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG13_28) column to 98.5% purity. Retention time on analytical RP-HPLC was 23.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{81}H_{131}N_{25}O_{29}$ [1918.96]; Observed [1918.09]. Apparent molecular weight from sedimentation equilibrium analysis: 1910 ± 130.

AspLys4 (Ac-Tyr Gly Gly Ala Asp Ala Ala Ala Lys Asp Ala Ala Ala Lys Asp Ala Ala Ala Lys Ala- NH_2). The peptide was synthesized using 0.1008 g (0.02 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1296 g of resin (67.0% yield). The cleavage yielded 37.8 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG04_17) column to 98.9% purity. Retention time on analytical RP-HPLC was 26.6 min. The identity of

the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{126}N_{24}O_{28}$ [1847.92]; Observed [1847.98]. Apparent molecular weight from sedimentation equilibrium analysis: 2000 ± 130.

AspLys3 (Ac-Tyr Gly Gly Ala Asp Ala Ala Lys Ala Asp Ala Ala Lys Ala Asp Ala Ala Lys Ala-NH_2). The peptide was synthesized using 0.1123 g (0.022 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1577 g of resin (98.0% yield). The cleavage yielded 40.8 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG04_21) column to 98.5% purity. Retention time on analytical RP-HPLC was 24.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{121}N_{23}O_{27}$ [1776.89]; Observed [1776.94]. Apparent molecular weight from sedimentation equilibrium analysis: 1780 ± 140.

AspOrn5 (Ac-Tyr Gly Gly Asp Ala Ala Ala Ala Orn Asp Ala Ala Ala Ala Orn Asp Ala Ala Ala Ala Orn- NH_2). The peptide was synthesized using 0.1054 g (0.021 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1362 g of resin (67.6% yield). The cleavage yielded 25.4 mg of crude peptide (98% yield). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 22.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{78}H_{125}N_{25}O_{29}$ [1876.91]; Observed [1877.90]. Apparent molecular weight from sedimentation equilibrium analysis: 1760 ± 140.

AspOrn4 (Ac-Tyr Gly Gly Ala Asp Ala Ala Ala Orn Asp Ala Ala Ala Orn Asp Ala Ala Ala Orn Ala- NH_2). The peptide was synthesized using 0.2796 g (0.0503 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3817 g of resin (97.1% yield). The cleavage yielded 77.4 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG04_18) column to 98.9% purity. Retention time on analytical RP-HPLC was 24 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{75}H_{120}N_{24}O_{28}$ [1805.88]; Observed [1805.46]. Apparent molecular weight from sedimentation equilibrium analysis: 1880 ± 32.

AspOrn3 (Ac-Tyr Gly Gly Ala Asp Ala Ala Orn Ala Asp Ala Ala Orn Ala Asp Ala Ala Orn Ala- NH_2). The peptide was synthesized using 0.2708 g (0.048 mmol) Fmoc-PAL-PEG-PS. The

synthesis gave 0.3578 g of resin. The cleavage yielded 74.2 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG00_13) column to 99.3% purity. Retention time on analytical RP-HPLC was 19.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{72}H_{115}N_{23}O_{27}$ [1734.84]; Observed [1734.76]. Apparent molecular weight from sedimentation equilibrium analysis: 1690 ± 120.

AspDab5 (Ac-Tyr Gly Gly Asp Ala Ala Ala Ala Dab Asp Ala Ala Ala Ala Dab Asp Ala Ala Ala Ala Dab- NH_2). The peptide was synthesized using 0.1185 g (0.024 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1550 g of resin (72.6% yield). The cleavage yielded 41.7 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C4 (PLG13_25) column to 98.1% purity. Retention time on analytical RP-HPLC was 21.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{75}H_{119}N_{25}O_{29}$ [1834.87]; Observed [1834.24]. Apparent molecular weight from sedimentation equilibrium analysis: 1920 ± 130.

AspDab4 (Ac-Tyr Gly Gly Ala Asp Ala Ala Ala Dab Asp Ala Ala Ala Dab Asp Ala Ala Ala Dab Ala-NH_2). The peptide was synthesized using 0.2855 g (0.0514 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.5467 of resin . The cleavage yielded 66.6 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG04_21) column to 98.1% purity. Retention time on analytical RP-HPLC was 23.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{72}H_{114}N_{24}O_{28}$ [1763.83]; Observed [1763.73]. Apparent molecular weight from sedimentation equilibrium analysis: 1610 ± 130.

AspDab3 (Ac-Tyr Gly Gly Ala Asp Ala Ala Dab Ala Asp Ala Ala Dab Ala Asp Ala Ala Dab Ala- NH_2). The peptide was synthesized using 0.2960 g (0.0532 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3414 g of resin. The cleavage yielded 70.3 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.1% purity. Retention time on analytical RP-HPLC was 16.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{69}H_{109}N_{23}O_{27}$ [1692.79]; Observed [1693.24]. Apparent molecular weight from sedimentation equilibrium analysis: 1970 ± 210.

AspDap5 (Ac-Tyr Gly Gly Asp Ala Ala Ala Ala Dap Asp Ala Ala Ala Ala Dap Asp Ala Ala Ala Ala Dap- NH_2). The peptide was synthesized using 0.1248 g (0.025 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1658 g of resin (79.0% yield). The cleavage yielded 43.9 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C4 (PLG13_25) column to 98.5% purity. Retention time on analytical RP-HPLC was 20.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{72}H_{113}N_{25}O_{29}$ [1792.82]; Observed [1792.10]. Apparent molecular weight from sedimentation equilibrium analysis: 1820 ± 84.

AspDap4 (Ac-Tyr Gly Gly Ala Asp Ala Ala Ala Dap Asp Ala Ala Ala Dap Asp Ala Ala Ala Dap Ala- NH_2). The peptide was synthesized using 0.2792 g (0.0503 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3178 of resin. The cleavage yielded 62.9 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG04_12) column to 99% purity. Retention time on analytical RP-HPLC was 20.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{69}H_{108}N_{24}O_{28}$ [1721.78]; Observed [1721.05]. Apparent molecular weight from sedimentation equilibrium analysis: 1990 ± 150.

AspDap3 (Ac-Tyr Gly Gly Ala Asp Ala Ala Dap Ala Asp Ala Ala Dap Ala Asp Ala Ala Dap Ala- NH_2). The peptide was synthesized using 0.2784 g (0.0501 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3188 g of resin. The cleavage yielded 38.1 mg of crude peptide (>99% yield). The peptide was purified by preparative RP-HPLC using a C_4 (PLG01_14) column to 99.2% purity. Retention time on analytical RP-HPLC was 18.7 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{66}H_{103}N_{23}O_{27}$ [1650.75]; Observed [1649.96]. Apparent molecular weight from sedimentation equilibrium analysis: 1730 ± 140.

GluLys5 (Ac-Tyr Gly Gly Glu Ala Ala Ala Ala Lys Glu Ala Ala Ala Ala Lys Glu Ala Ala Ala Ala Lys- NH_2). The peptide was synthesized using 0.2804 g (0.0505 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3328 g of resin (59.8% yield). The cleavage yielded 75.3 mg of crude peptide (>99% yield, 39.8% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.3% purity. Retention time on analytical RP-HPLC was 26.8 min. The identity

of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{84}H_{137}N_{25}O_{29}$ [1961.01]; Observed [1960.94].

GluLys4 (Ac-Tyr Gly Gly Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala- NH_2). The peptide was synthesized using 0.2010 g (0.0504 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2605 g of resin (60.7% yield). The cleavage yielded 72.4 mg of crude peptide (>99% yield, 35.7% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 26.6 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{81}H_{132}N_{24}O_{28}$ [1889.95]; Observed [1889.79]. Apparent molecular weight from sedimentation equilibrium analysis: 1700 ± 170.

GluLys3 (Ac-Tyr Gly Gly Ala Glu Ala Ala Lys Ala Glu Ala Ala Lys Ala Gly Ala Ala Lys Ala-NH_2). The peptide was synthesized using 0.2002 g (0.0501 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2602 g of resin (62.3% yield). The cleavage yielded 62.8 mg of crude peptide (>99% yield, 46.1% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column. To 98.5% purity. Retention time on analytical RP-HPLC was 24.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{127}N_{23}O_{27}$ [1818.88]; Observed [1817.90]. Apparent molecular weight from sedimentation equilibrium analysis: 1770 ± 200.

GluOrn5 (Ac-Tyr Gly Gly Glu Ala Ala Ala Ala Orn Glu Ala Ala Ala Ala Orn Glu Ala Ala Ala Ala Orn-NH_2). The peptide was synthesized using 0.2801 g (0.0504 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3203 g of resin (59.1% yield). The cleavage yielded 50.2 mg of crude peptide (>99% yield, 34.9% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.1% purity. Retention time on analytical RP-HPLC was 24.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{81}H_{131}N_{25}O_{29}$ [1918.94]; Observed [1918.96]. Apparent molecular weight from sedimentation equilibrium analysis: 2160 ± 200.

GluOrn4 (Ac-Tyr Gly Gly Ala Glu Ala Ala Ala Orn Glu Ala Ala Ala Orn Glu Ala Ala Ala Orn Ala- NH_2). The peptide was synthesized using 0.2170 g (0.0544 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2752 g of resin (74.1% yield). The cleavage yielded 77.4 mg of crude peptide

(>99% yield, 48.3% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 25.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{126}N_{24}O_{28}$ [1847.95]; Observed [1847.63]. Apparent molecular weight from sedimentation equilibrium analysis: 1800 ± 300.

GluOrn3 (Ac-Tyr Gly Gly Ala Glu Ala Ala Orn Ala Glu Ala Ala Orn Ala Glu Ala Ala Orn Ala-NH_2). The peptide was synthesized using 0.2100 g (0.0525 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2732 g of resin (78.2% yield). The cleavage yielded 74.2 mg of crude peptide (>99% yield, 38.1% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.5% purity. Retention time on analytical RP-HPLC was 23.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{121}N_{23}O_{27}$ [1776.88]; Observed [1776.23]. Apparent molecular weight from sedimentation equilibrium analysis: 950 ± 200.

GluDab5 (Ac-Tyr Gly Gly Glu Ala Ala Ala Ala Dab Glu Ala Ala Ala Ala Dab Glu Ala Ala Ala Ala Dab- NH_2). The peptide was synthesized using 0.2782 g (0.0501 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3216 g of resin (67.2% yield). The cleavage yielded 62.3 mg of crude peptide (>99% yield, 42.4% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.8% purity. Retention time on analytical RP-HPLC was 24.6 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{125}N_{25}O_{29}$ [1876.47]; Observed [1876.64]. Apparent molecular weight from sedimentation equilibrium analysis: 1750 ± 200.

GluDab4 (Ac-Tyr Gly Gly Ala Glu Ala Ala Ala Dab Glu Ala Ala Ala Dab Glu Ala Ala Ala Dab Ala- NH_2). The peptide was synthesized using 0.2190 g (0.0549 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2730 g of resin (74% yield). The cleavage yielded 66.6 mg of crude peptide (>99% yield, 26.3% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.5% purity. Retention time on analytical RP-HPLC was 24.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{120}N_{24}O_{28}$ [1805.95]; Observed [1806.18]. Apparent molecular weight from sedimentation equilibrium analysis: 700 ± 100.

GluDab3 (Ac-Tyr Gly Gly Ala Glu Ala Ala Dab Ala Glu Ala Ala Dab Ala Glu Ala Ala Dab Ala-NH_2). The peptide was synthesized using 0.2170 g (0.0542 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2582 g of resin (81% yield). The cleavage yielded 70.3 mg of crude peptide (>99% yield, 35.2% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.1% purity. Retention time on analytical RP-HPLC was 21.3 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{72}H_{115}N_{23}O_{27}$ [1734.88]; Observed [1734.89]. Apparent molecular weight from sedimentation equilibrium analysis: 1300 ± 200.

GluDap5 (Ac-Tyr Gly Gly Glu Ala Ala Ala Ala Dap Glu Ala Ala Ala Ala Dap Glu Ala Ala Ala Ala Dap- NH_2). The peptide was synthesized using 0.2780 g (0.0500 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3321 g of resin (73.4% yield). The cleavage yielded 61.3 mg of crude peptide (>99% yield, 41.6% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 23.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{119}N_{25}O_{29}$ [1834.94]; Observed [1834.47].

GluDap4 (Ac-Tyr Gly Gly Ala Glu Ala Ala Ala Dap Glu Ala Ala Ala Dap Glu Ala Ala Ala Dap Ala- NH_2). The peptide was synthesized using 0.2072 g (0.0507 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2400 g of resin (71.3% yield). The cleavage yielded 62.9 mg of crude peptide (>99% yield, 34.9% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 23.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{72}H_{114}N_{24}O_{28}$ [1764.95]; Observed [1763.44]. Apparent molecular weight from sedimentation equilibrium analysis: 2600 ± 300. The sedimentation equilibrium data was best fit with mostly monomer with a hint of large order aggregation.

GluDap3 (Ac-Tyr Gly Gly Ala Glu Ala Ala Dap Ala Glu Ala Ala Dap Ala Glu Ala Ala Dap Ala-NH_2). The peptide was synthesized using 0.2030 g (0.0517 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1430 g of resin (45% yield). The cleavage yielded 38.1 mg of crude peptide (>99% yield, 43.5% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.7% purity. Retention time on analytical RP-HPLC was 20.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+]

$C_{69}H_{109}N_{23}O_{27}$ [1693.88]; Observed [1692.16]. Apparent molecular weight from sedimentation equilibrium analysis: 1600 ± 170.

AadLys5 (Ac-Tyr Gly Gly Aad Ala Ala Ala Ala Lys Aad Ala Ala Ala Ala Lys Aad Ala Ala Ala Ala Lys -NH_2). The peptide was synthesized using 0.2569 g (0.046 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3221 g of resin(54.8% yield). The cleavage yielded 94.5 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C4 (PLG13_28) column to purity 98.5%. Retention time on analytical RP-HPLC was 26.3 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{87}H_{143}N_{25}O_{29}$ [2003.06]; Observed [2003.18]. Apparent molecular weight from sedimentation equilibrium analysis: 1860 ± 120.

AadLys4 (Ac-Tyr Gly Gly Ala Aad Ala Ala Ala Lys Aad Ala Ala Ala Lys Aad Ala Ala Ala Lys Ala-NH_2). The peptide was synthesized using 0.3000 g (0.054 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3979 g of resin(73.6% yield). The cleavage yielded 63.8 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 26.6 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{84}H_{138}N_{24}O_{28}$ [1932.02]; Observed [1931.04]. Apparent molecular weight from sedimentation equilibrium analysis: 1670 ± 110.

AadLys3 (Ac-Tyr Gly Gly Ala Aad Ala Ala Lys Ala Aad Ala Ala Lys Ala Aad Ala Ala Lys Ala-NH_2). The peptide was synthesized using 0.2962 g (0.0533 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3402 g of resin (62.3% yield). The cleavage yielded 65.1 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.5% purity. Retention time on analytical RP-HPLC was 24.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{81}H_{133}N_{23}O_{27}$ [1860.98]; Observed [1860.25]. Apparent molecular weight from sedimentation equilibrium analysis: 1840 ± 180.

AadOrn5 (Ac-Tyr Gly Gly Aad Ala Ala Ala Ala Orn Aad Ala Ala Ala Ala Orn Aad Ala Ala Ala Ala Orn-NH_2). The peptide was synthesized using 0.2701 g (0.048 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3339 g of resin (51.6% yield). The cleavage yielded 104.9 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_4 (PLG13_28) column to purity 98.5%. Retention time on analytical RP-HPLC was 25.8 min. The identity of the peptide

was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{84}H_{137}N_{25}O_{29}$ [1961.01]; Observed [1960.02]. Apparent molecular weight from sedimentation equilibrium analysis: 1840 ± 180.

AadOrn4 (Ac-Tyr Gly Gly Ala Aad Ala Ala Ala Orn Aad Ala Ala Ala Orn Aad Ala Ala Ala Orn Ala-NH_2). The peptide was synthesized using 0.2796 g (0.0503 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.2952 g of resin (74.1% yield). The cleavage yielded 72.9 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.9% purity. Retention time on analytical RP-HPLC was 25.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{81}H_{132}N_{24}O_{28}$ [1889.97]; Observed [1889.33]. Apparent molecular weight from sedimentation equilibrium analysis: 1960 ± 150.

AadOrn3 (Ac-Tyr Gly Gly Ala Aad Ala Ala Orn Ala Aad Ala Ala Orn Ala Aad Ala Ala Orn Ala- NH_2). The peptide was synthesized using 0.2778 g (0.0500 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3032 g of resin (78.2% yield). The cleavage yielded 70.4 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_4 (PLG09_22) column to purity 98.5%.. Retention time on analytical RP-HPLC was 21.8 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{127}N_{23}O_{27}$ [1818.94]; Observed [1818.22]. Apparent molecular weight from sedimentation equilibrium analysis: 1600 ± 110.

AadDab5 (Ac-Tyr Gly Gly Aad Ala Ala Ala Ala Dab Aad Ala Ala Ala Ala Dab Aad Ala Ala Ala Ala Dab- NH_2). The peptide was synthesized using 0.1221 g (0.024 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1668 g of resin. The cleavage yielded 50.1 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_{18} column to 98.5% purity. Retention time on analytical RP-HPLC was 25.4 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH+] $C_{81}H_{131}N_{25}O_{29}$ [1918.96]; Observed [1918.10]. Apparent molecular weight from sedimentation equilibrium analysis: 1610 ± 130.

AadDab4 (Ac-Tyr Gly Gly Ala Aad Ala Ala Ala Dab Aad Ala Ala Ala Dab Aad Ala Ala Ala Dab Ala- NH_2). The peptide was synthesized using 0.1283 g (0.026 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.1811 g of resin. The cleavage yielded 51.8 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_4 (PLG12_25) column to 98.5% purity.

Retention time on analytical RP-HPLC was 24.5 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{78}H_{126}N_{24}O_{28}$ [1847.93]; Observed [1847.09]. Apparent molecular weight from sedimentation equilibrium analysis: 1780 ± 87.

AadDab3 (Ac-Tyr Gly Gly Ala Aad Ala Ala Dab Ala Aad Ala Ala Dab Ala Aad Ala Ala Dab Ala- NH_2). The peptide was synthesized using 0.2804 g (0.0505 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3182 g of resin (81% yield). The cleavage yielded 70.3 mg of crude peptide (>99% yield, 35.2% purity). The peptide was purified by preparative RP-HPLC using a C_{18} column to 99.1% purity. Retention time on analytical RP-HPLC was 21.3 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{121}N_{23}O_{27}$ [1776.89]; Observed [1776.11]. Apparent molecular weight from sedimentation equilibrium analysis: 1900 ± 140.

AadDap5 (Ac-Tyr Gly Gly Aad Ala Ala Ala Ala Dap Aad Ala Ala Ala Ala Dap Aad Ala Ala Ala Ala Dap- NH_2). The peptide was synthesized using 0.2643 g (0.048 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3149 g g of resin (42.6% yield). The cleavage yielded 86.5 mg of crude peptide. The peptide was purified by preparative RP-HPLC using a C_4 (PLG13_25) column to 98.5% purity. Retention time on analytical RP-HPLC was 25.0 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MNa+] $C_{78}H_{125}N_{25}O_{29}$ [1898.90]; Observed [1898.20]. Apparent molecular weight from sedimentation equilibrium analysis: 1610 ± 130.

AadDap4 (Ac-Tyr Gly Gly Ala Aad Ala Ala Ala Dap Aad Ala Ala Ala Dap Aad Ala Ala Ala Dap Ala- NH_2). The peptide was synthesized using 0.2950 g (0.0531 mmol) Fmoc-PAL-PEG-PS. The synthesis gave 0.3327 g of resin (71.3% yield). The cleavage yielded 62.9 mg of crude peptide (>99% yield, 34.9% purity). The peptide was purified by preparative RP-HPLC using a C_4 (PLG07_20) column to 98.9% purity. Retention time on analytical RP-HPLC was 23.9 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{75}H_{120}N_{24}O_{28}$ [1805.88]; Observed [1805.5]. Apparent molecular weight from sedimentation equilibrium analysis: 1870 ± 110.

AadDap3 (Ac-Tyr Gly Gly Ala Aad Ala Ala Dap Ala Aad Ala Ala Dap Ala Aad Ala Ala Dap Ala- NH_2). The peptide was synthesized using 0.2958 g (0.0532 mmol) Fmoc-PAL-PEG-PS. The

synthesis gave 0.3217 g of resin (45% yield). The cleavage yielded 38.1 mg of crude peptide (>99% yield, 43.5% purity). The peptide was purified by preparative RP-HPLC using a C_4 (PLG05_19) column to 99.7% purity. Retention time on analytical RP-HPLC was 20.1 min. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for [MH^+] $C_{72}H_{115}N_{23}O_{27}$ [1734.84]; Observed [1733.47]. Apparent molecular weight from sedimentation equilibrium analysis: 1800 ± 140.

Sedimentation Equilibrium Experiments by Analytical Ultracentrifugation

All proteins were dissolved in 120 μL buffer of 10 mM MOPS ((3-*N*-morpholino)-propanesulfonic acid) at pH 7.0 to give an absorbance between 0.2 and 0.4 at 275 nm. The Glu-containing peptides were spun at 30, 40, and 50k rpm using an equilibration delay time of 8 hours at 25°C. The Asp- and Aad-containing peptides (in buffer that also included 150 mM NaCl) were spun at 50k rpm using an equilibration delay time of 8 hours at 4°C. The data were truncated using the program WinReedit (version 0.999, 1998) and analyzed using WinNonLin (version 1.035, 1997) (*11*). For the analysis, the solution density (1.00620 kg/m^3) and the peptide partial specific volume were calculated by using Sednterp as described by Laue *et al* (*12*). The partial specific volume for the non-standard amino acids was calculated by using a procedure described by Durschschlag *et al* (*13*). Single species analysis for all proteins suggested a molecular weight consistent with a monomer.

Circular Dichroism Spectroscopy

CD data was obtained using a 1 mm pathlength cell. The concentration of peptide stock solutions were determined by the tyrosine absorbance in 6 M guanidinium chloride (ε_{282}=1220, ε_{280}=1285, ε_{278}=1395, ε_{276}=1455) (*14, 15*). CD measurements were performed at peptide concentrations near 50 μM in 1 mM phosphate, 1 mM citrate, 1 mM borate buffer (pH 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12) at 0°C. The data was analyzed using Kaleidagraph 3.52 (Synergy Software, CA). Each reported CD value was the mean of at least 3 determinations. Data are expressed in terms of mean residue molar ellipticity (deg·cm^2·$dmol^{-1}$) normalized to the number of backbone amide bonds. The mean residue molar ellipticity of the peptides was independent of peptide concentration (50-100 μM). The helical content of two peptides were considered to be significantly different if the CD signal of the less helical peptide was less than 60% of the CD

signal for the more helical peptide, and the P-value for comparing the CD signal of the two peptides was less than 10^{-5}. The fraction helix of each peptide (f_{helix}) was derived from the mean residue molar ellipticity at 222 nm and the number of backbone amides (N) using equation 1 (*16*).

$$f_{helix} = \frac{[\theta]_{222}}{40000(1 - \frac{2.5}{N})} \qquad \text{equation 1}$$

N, total number of residues

Deriving f_{helix} and Interaction Energy Using Modified Lifson-Roig Theory

These calculations were performed following previously published procedures (*3*). The modified Lifson-Roig theory (*17*) that included end-capping effects (*18*), and interactions between charged residues and the helix macrodipole (*2*) was used to derived the fraction helix (f_{helix}) of each peptide in the absence of side chain-side chain interactions (*18*). These calculations were performed using compiled computer code written in C++. Literature values for the helix propagation parameter w (*16, 18-20*) and the N-terminal capping parameter n (*16, 18, 19*) were used for the calculations, and the helix initiation parameter v and the C-terminal capping parameter c for all residues were set to 0.048 (*16*) and 1 (*16*), respectively. The contribution of interactions between charged residues and the helix macro-dipole was derived based on the method as described by Scholtz (*2*). Modifications to the method were made to account for helical states with multiple helical segments (*3*).

To obtain the side chain-side chain interaction energetics, the statistical weight for each specific intra-helical side chain-side chain interaction (p) was derived from the experimental f_{helix} based on the nesting block method (*1-3*). Modifications to the nesting block method were made to explicitly include helix propagation parameters for each amino acid (*3, 16, 18, 19*), interactions between charged residues and helix macrodipole (*2, 3*), and also states with multiple helical segments (*3*). For every possible state of each peptide, the f_{helix} and probability were calculated and combined. The free energy of each specific side chain-side chain interaction was calculated by $\Delta G = -RT \cdot \ln(p)$ (*2, 3*).

Conformational Analysis by Molecular Mechanics

The conformational analysis was performed using the program Discovery Studio 2.1 (Accelrys, CA) on an IBM x3550M2 workstation (CPU: Dual Xeon E5530 2.4 GHz with Quad cores; RAM: 48 G) running the operating system CentOS 5.3. The models were initially created with ideal backbone dihedral angles for α-helix with various combinations of potential low energy side chain dihedrals. Each conformation was then minimized using the CFF forcefield. The nonbond radius of 99 Å, nonbond higher cutoff distance of 98 Å, and nonbond lower cutoff distance of 97 Å were employed to effectively model no cutoffs. Distance dependent dielectric constant of 2 was used as the implicit solvent model. Minimization was performed by steepest descent and conjugate gradient protocols until convergence (converging slope was set to 0.1 kcal/(mol x Å). After minimization, each conformation was re-examined to remove duplicating conformations, because minimization with different starting conformations occasionally resulted in the same final conformation. When the same conformation was represented more than once, only the lowest energy conformation was considered in further analyses.

Survey of Glu-Lys (i, i+3) and (i, i+4) Salt Bridges in Protein α-Helices

The survey was performed on the PDBselect (April 2009, 25% threshold) (*6, 7*), a database of non-redundant protein chains. The α-helical conformation was defined based on backbone dihedral angles (*8, 9, 21, 22*). The helical residues were selected using in-house code written in ActivePerl 5.8.8.819. The backbone dihedral angles were compiled using DSSP (*23*). The side chain dihedrals were compiled using Discovery Studio 2.1 (Accelrys, CA, USA). Segments of six or more α-helical residues were considered to avoid end effects (*3, 21, 22*). The occurrence was compiled for Glu/Asp-Lys (*i*, *i+3*), (*i*, *i+4*), and (*i*, *i+5*) sequence patterns. These occurrences were compiled using in-house code written in ActivePerl 5.8.8.819. The helix pair propensity of each residue pattern was calculated by dividing the occurrence of the sequence pattern in α-helices by the expected occurrence for the sequence pattern based on the all structures in the database. The pair propensity of each residue pattern was calculated by dividing the occurrence of the sequence pattern in α-helices by the expected occurrence for the sequence pattern based on the α-helices in the database. The expected occurrence and the corresponding standard deviation was obtained by bootstrapping (*24*) the sequence pattern against the

appropriate context across the PDBselect database. The bootstrapping was performed using in-house code written in C++. Dividing the difference between the occurrence and the expected occurrence by the standard deviation gave the z value, which was used to obtain the P value based on a normal distribution (*25, 26*). A salt bridge was considered to be present if any of the oxygens of the side chain carboxylate of Asp or Glu was within 3Å of the nitrogen of the side chain amine of Lys. The occurrence of intrahelical salt bridge was compiled using the program Discovery Studio 2.1 (Accelrys, CA, USA) on a personal computer running the operating system Windows XP (Microsoft, WA, USA).

References

1. Robert, C. H. (1990) A hierarchical nesting approach to describe the stability of α-helices with side-chain interactions, *Biopolymers 30*, 335-347.
2. Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide, *Biochemistry 32*, 9668-9676.
3. Cheng, R. P., Girinath, P., and Ahmad, R. (2007) Effect of lysine side chain length on intra-helical glutamate-lysine ion pairing interactions, *Biochemistry 46*, 10528-10537.
4. Marqusee, S., and Baldwin, R. L. (1987) Helix stabilization by Glu- ... Lys+ salt bridges in short peptides of denovo design, *Proc. Natl. Acad. Sci. U. S. A. 84*, 8898-8902.
5. Jeffrey, G. A., and Saenger, W. (1994) in *Hydrogen Bonding in Biological Structures*, pp 365-371, Springer-Verlag, Heidelberg.
6. Hobohm, U., and Sander, C. (1994) Enlarged representative set of protein structures, *Protein Sci. 3*, 522-524.
7. Griep, S., and Hobohm, U. (2010) PDBselect 1992-2009 and PDBfilter-select, *Nucleic Acids Res. 38*, D318-D319.
8. Gunasekaran, K., Nagarajaram, H. A., Ramakrishnan, C., and Balaram, P. (1998) Stereochemical Punctuation Marks in Protein Structures: Glycine and Proline Containing Helix Stop Signals, *J. Mol. Biol. 275*, 917-932.
9. Engel, D. E., and DeGrado, W. F. (2004) Amino Acid Propensities are Position-dependent Throughout the Length of α-Helices, *J. Mol. Biol. 337*, 1195-1205.
10. Fields, G. B., and Noble, R. L. (1990) Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbonyl amino-acids, *Int. J. Pept. Protein Res. 35*, 161-214.
11. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques, *Biophys. J. 36*, 575-588.
12. Laue, T. M., Shah, B., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E. R., A. J.; Horton, J. C., Ed.), pp 90-125, Royal Society of Chemistry, Cambridge.

13. Durchschlag, H., and Zipper, P. (1994) Calculations of the partial volume of organic compounds and polymers, *Prog. Colloid Polym. Sci. 94*, 20-39.
14. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins, *Biochemistry 6*, 1948-1954.
15. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption-coefficient of a protein, *Protein Sci. 4*, 2411-2423.
16. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix propensities of the amino-acids measured in alanine-based peptides without helix-stabilizing side-chain interactions, *Protein Sci. 3*, 843-852.
17. Lifson, S., and Roig, A. (1961) The theory of helix-coil transition in polypeptides., *J. Chem. Phys. 34*, 1963-1974.
18. Doig, A. J., Chakrabartty, A., Klingler, T. M., and Baldwin, R. L. (1994) Determination of free-energies of N-capping in α-helices by modification of the Lifson-Roig helix-coil theory to include N-capping and C-capping, *Biochemistry 33*, 3396-3403.
19. Padmanabhan, S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1996) Helix propensities of basic amino acids increase with the length of the side-chain, *J. Mol. Biol. 257*, 726-734.
20. Chiu, H. P., Suzuki, Y., Gullickson, D., Ahmad, R., Kokona, B., Fairman, R., and Cheng, R. P. (2006) Helix propensity of highly fluorinated amino acids, *J. Am. Chem. Soc. 128*, 15556-15557.
21. Cheng, R. P., Girinath, P., Suzuki, Y., Kuo, H.-T., Hsu, H.-C., Wang, W.-R., Yang, P.-A., Gullickson, D., Wu, C.-H., Koyack, M. J., Chiu, H.-P., Weng, Y.-J., Hart, P., Kokona, B., Fairman, R., Lin, T.-E., and Barrett, O. (2010) Positional effects on helical Ala-based peptides, *Biochemistry 49*, 9372-9384.
22. Cheng, R. P., Weng, Y.-J., Wang, W.-R., Koyack, M. J., Suzuki, Y., Wu, C.-H., Yang, P.-A., Hsu, H.-C., Kuo, H.-T., Girinath, P., and Fang, C.-J. Helix formation and capping energetics of arginine analogs with varying side chain length, *Amino Acids*, in press.
23. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577-2637.
24. Efron, B., and Gong, G. (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation, *Amer. Stat. 37*, 36-48.
25. Kuebler, R. R., and Smith, H. (1976) Statistics-A beginning, in *Statistics-A beginning*, John Wiley & Sons, Inc, New York. pp 302.
26. Klugh, H. E. (1970) Statistics-The essentials for research, in *Statistics-The essentials for research*, John Wiley & Sons, Inc, New York. pp 350.