Amino Acids (2012) 43:195–206
DOI 10.1007/s00726-011-1064-2

ORIGINAL ARTICLE

Helix formation and capping energetics of arginine analogs with varying side chain length

Richard P. Cheng · Yi-Jen Weng · Wei-Ren Wang · Marc J. Koyack · Yuta Suzuki · Cheng-Hsun Wu · Po-An Yang · Hao-Chun Hsu · Hsiou-Ting Kuo · Prashant Girinath · Chun-Jen Fang

Received: 12 June 2011 / Accepted: 20 August 2011 / Published online: 16 September 2011

Abstract Arginine (Arg) has been used for recognizing negatively charged biological molecules, cell penetration, and oligosaccharide mass signal enhancement. The versatility of Arg has inspired the need to develop Arg analogs and to research the structural effects of incorporating Arg analogs. Accordingly, we investigated the effect of Arg side chain length on helix formation by studying 12 Ala-based peptides containing the Arg analogs (*S*)-2-amino-6-guanidino-hexanoic acid (Agh), (*S*)-2-amino-4-guanidinobutyric acid (Agb), and (*S*)-2-amino-3-guanidinopropionic acid (Agp). Solid phase guanidinylation with orthogonal protection strategies was necessary to synthesize Agb- and Agp-containing peptides using Fmoc-based chemistry. The fraction helix for the peptides was determined by circular dichroism spectroscopy, and used to derive the statistical mechanical parameters and energetics for N-capping, C-capping, and helix propagation (propensity). All four Arg analogs were unfavorable for N-capping. The C-cap parameter followed the trend Agp < Agb < Arg < Agh, showing more favorable C-cap energetics with increasing side chain length. In contrast, helix propensity followed the trend Agp < Agb < Arg > Agh, highlighting the uniqueness of the Arg side chain length in helix formation. Molecular mechanics calculations and a survey on protein structures were consistent with the experimental results. Furthermore, calculations and survey both showed that the $g-$ conformation for the χ_1 dihedral was present for the first two residues at the N-terminus of helices, but not favored in the center or C-terminus of helices due to sterics. These results should serve as the foundation for developing Arg-related bioactive compounds and technologies.

Keywords Arginine · Side chain length · Peptide · Helix

Electronic supplementary material The online version of this article (doi:10.1007/s00726-011-1064-2) contains supplementary material, which is available to authorized users.

R. P. Cheng (✉) · Y.-J. Weng · W.-R. Wang · C.-H. Wu · P.-A. Yang · H.-C. Hsu · H.-T. Kuo · C.-J. Fang
Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan
e-mail: rpcheng@ntu.edu.tw

M. J. Koyack · Y. Suzuki · P. Girinath
Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260-3000, USA

Introduction

Arginine and lysine are the two positively charged amino acids with side chain pK_a values greater than ten (Fig. 1). Incorporated into proteins, Arg is important for recognizing anionic biological entities such as phosphates (Hirsch et al. 2007), glycosaminoglycans (Blaum et al. 2010), and others (Chakrabarti 1994). Furthermore, Arg is responsible for the cell penetration activity of Arg-containing proteins/peptides (Wender et al. 2000, 2008; Mitchell et al. 2000), and is frequently observed at protein–protein interaction interfaces and hotspots (Janin et al. 1988; Argos 1988; Jones and Thornton 1995; Tsai et al. 1997; Bogan and Thorn 1998; Hu et al. 2000). For technological advantages, Arg derivatives have been covalently attached to carbohydrates to enhance the detection sensitivity of oligosaccharides in MALDI (matrix-assisted laser desorption ionization) mass spectrometry analysis (Zhao et al. 1997; Shinohara et al. 2004; Baumgart et al. 2004; Nishikaze and Takayama 2006; Northern et al. 2008). Arg is different from Lys in bearing a guanidinium group to carry the positive charge (Fig. 1). This creates a more diffuse positive charge compared to the Lys ammonium group. Furthermore, the

guanidinium group is capable of forming multiple hydrogen bonds in a multidentate fashion. This unique guanidinium bearing side chain of Arg is critical for various bioactivities such as cell penetration (Mitchell et al. 2000), and cannot be replaced with Lys. Cell penetration activity can be achieved by attaching the guanidinium group onto the side chain of oligomers with non-natural backbones including peptoids (Wender et al. 2000), oligo-carbamates (Wender et al. 2002), and β-peptides (Umezawa et al. 2001). Accordingly, it is important to synthesize guanidinium-containing Arg analogs and to investigate the structural and functional effects of these analogs.

One of the most direct methods to synthesize guanidinium-containing Arg analogs would be to guanidinylate the corresponding amine using one of the various guanidinylation reagents (Robinson and Roskamp 1997; Shey and Sun 1998; Ho and Sun 1999; Katritzky and Rogovoy 2005). Using these guanidinylation reagents, a number of Arg analogs have been synthesized such as (*S*)-2-amino-6-guanidinohexanoic acid (Agh; Fig. 1) (Poss et al. 1992; Drake et al. 1994), α-(trifluoromethyl)- and α-(difluoromethyl)-arginines (Moroni et al. 2001), and cyclopropane-containing Arg analogs (Pradhan et al. 2009). Furthermore, solid phase guanidinylation has been used to prepare peptides containing Agh (Fig. 1) (Yong et al. 1997, 1999; Feichtinger et al. 1998a; Zhang and Kennan 2001), Agb (Fig. 1) (Yong et al. 1997, 1999; Zhang and Kennan 2001; Diss and Kennan 2008), and Agp (Fig. 1) (Diss and Kennan 2008). Some of these Arg analog-containing peptides have been prepared by 9-fluorenylmethoxycarbonyl (Fmoc)-based chemistry, including Agb-containing dipeptide (Yong et al. 1997, 1999) and heptapeptide (Mitchell et al. 2000), and Agh-containing tripeptide (Yong et al. 1997, 1999) and octapeptide (Feichtinger et al. 1998a). However, peptides with more than ten residues have only been prepared by Boc-based chemistry, including Agp- and Agb-containing 30-mers (Zhang and Kennan 2001; Diss and Kennan 2008), and Agh-containing 14-mers (Zhang and Kennan 2001), whereas protocols using Fmoc-based chemistry remain to be fully demonstrated.

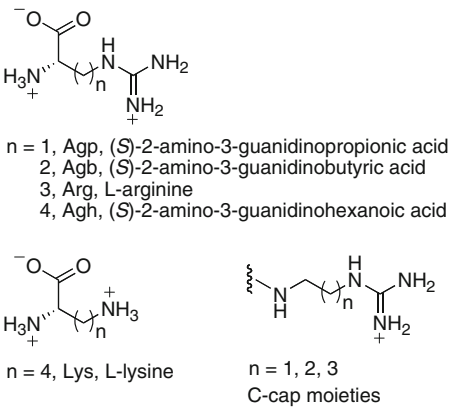


Fig. 1 Chemical structures of arginine, arginine analogs, lysine, and aminoguanidinium C-cap moieties

The ability to generate Arg analog-containing peptides appears reasonable, however there have only been limited studies on the structural effect of the Arg side chain length (Diss and Kennan 2008). Guanidinium-containing Arg analogs with various chain lengths were incorporated at the C-terminus of helical peptides (Schneider and DeGrado 1998). Optimal C-capping energetics was observed for analogs with two and four methylenes ($n = 1$ and 3, Fig. 1) (Schneider and DeGrado 1998). Interestingly, the molecule with four methylenes can be viewed as an achiral version of Arg, however analogs with even longer chain lengths were not investigated. More complex effects of the Arg side chain length on interhelical interactions were studied in a heterodimeric coiled coil system (Diss and Kennan 2008). The Arg analogs were incorporated at the helix–helix interface *a* position (Diss and Kennan 2008). Although Arg could be shortened by one methylene without affecting the stability when incorporated into pairing monomers, shortening the Arg side chain further resulted in significant decrease in stability ($\Delta T_m = -10°C$) (Diss and Kennan 2008). Also, interhelical interactions between the Arg analogs and Asp/Glu (both at the *a* position) were not significantly affected by the side chain length (Diss and Kennan 2008). However, Arg analogs with side chain lengths longer than Arg were not investigated. These pioneering studies provided some initial understanding on the structural effects of Arg analogs. Importantly, many biological activities require specific structures and conformations. Therefore, thorough studies on the structural effect of incorporating Arg analogs would serve as the foundation to facilitate the use of these analogs in various applications. Since one-third of protein residues adopt a helical conformation (Chou and Fasman 1974; Barlow and Thornton 1988; Cheng et al. 2010), herein we report the effect of Arg side chain length on helix formation and capping. We incorporated Arg analogs with varying side chain lengths into 19-residue peptides using Fmoc-based chemistry. Furthermore, we investigated the helix formation and capping parameters for the various Arg analogs by circular dichroism spectroscopy coupled with modified Lifson–Roig theory, conformational analysis by molecular mechanics calculations, and a survey of the protein structure database.

Results

Synthesis of protected Arg analogs

The appropriately protected Arg analogs suitable for Fmoc-based solid phase peptide synthesis were synthesized

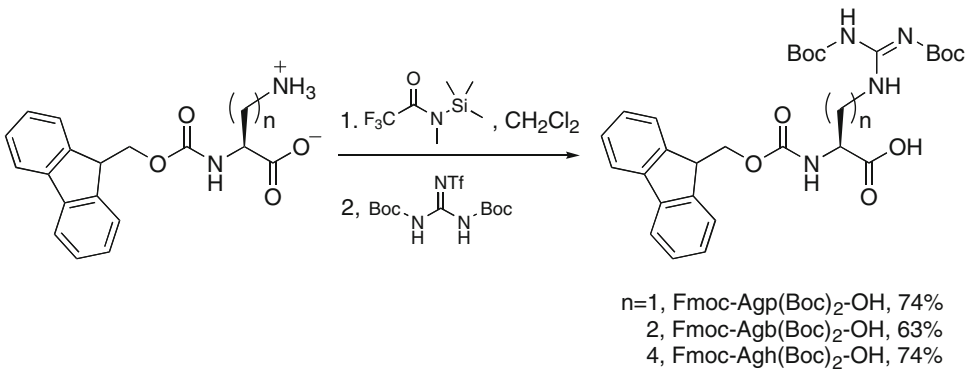
following known published procedures using *N*-triflylguanidinine derivatives (Scheme 1) (Feichtinger et al. 1998a, 1998b). In short, the N^{α}-Fmoc-protected Lys analog was silated using methyl(trimethylsilyl)-trifluoroacetanimide in refluxing dichloromethane to give a clear solution. Guanidinylation was then performed by treatment with *N*,*N*′-di-Boc-*N*′′-triflylguanidine (Feichtinger et al. 1998a, 1998b) at room temperature to give the desired products. The resulting amino acid backbone amine was already protected with Fmoc and the side chain guanidine group was protected with two Boc groups, suitable for Fmoc-based solid phase peptide synthesis (Fields and Noble 1990).

Design and synthesis of Arg analog-containing Ala-based peptides

The Ala-based peptides were designed according to analogous peptides initially studied by Baldwin (Chakrabartty et al. 1994; Doig and Baldwin 1995) and later by our group (Chiu et al. 2006; Cheng et al. 2010) (Table 1). To investigate the effect of the Arg side chain length on helix propensity (*w*), four XaaAla peptides were designed with four positions substituted simultaneously with the same Arg analog (Chakrabartty et al. 1994). The four NXaa01 peptides and the four CXaa19 peptides (each with one Arg analog at the N-terminus and C-terminus, respectively) were designed to derive the N-cap (*n*) and C-cap (*c*) parameters for the Arg analogs, respectively (Doig and Baldwin 1995). The N-terminus of all peptides was acetylated and the C-terminus was designed to be a carboxyamide so there would be no bias (from charged termini) on the statistical mechanical capping parameters derived for the Arg analogs (Cheng et al. 2010). Based on the modified Lifson–Roig theory (Lifson and Roig 1961; Doig et al. 1994; Chakrabartty et al. 1994; Doig and Baldwin 1995; Chiu et al. 2006; Cheng et al. 2010), the probability of residue Xaa for all conformational states in NXaa01 peptides would be represented by only *v* (initiation parameter), *n* (N-cap parameter), and 1 (the probability for random coil surrounded by residues in random coil conformation). This enables the exclusive derivation of the N-cap *n* parameter without needing to consider the helix propensity *w* parameter for residue Xaa. Similarly, the CXaa19 peptides enable the exclusive derivation of the C-cap *c* parameter without needing to consider the helix propensity *w* parameter for residue Xaa.

Peptides were synthesized by Fmoc-based chemistry (Fields and Noble 1990). For the Arg residues in the Arg-containing peptides, Fmoc–Arg(Pbf)–OH was double coupled for 25 min, because intramolecular cyclization of the activated ester occurs readily in this time frame, effectively terminating the coupling reaction. For peptides with Arg as the first residue coupled to the resin, quadruple coupling was performed to ensure attachment to the resin. For the Agh residues in the Agh-containing peptides, Fmoc–Agh(Boc)$_2$–OH was coupled for 45 min using standard protocols. For peptides with Agh as the first residue coupled to the resin, coupling was performed for 8 h to ensure attachment to the resin. The lack of need for double coupling is consistent with the unlikelihood of intramolecular cyclization. Intramolecular cyclization of the Agh activated ester would form a seven-membered ring, a medium size ring, which is inherently more difficult to form compared to six-membered rings.

For the Agb residues in peptide AgbAla, Fmoc–Agb(Boc)$_2$–OH could not be used to introduce all four Agb residues even with triple coupling or microwave conditions. Instead, the corresponding Dab-containing peptide was synthesized with Fmoc–Dab(ivDde)–OH (Dab, (*S*)-2,4-diaminobutyric acid; ivDde, isovaleryldimedone; Scheme 2) (Chhabra et al. 1998). The ivDde protecting group was then selectively removed using hydrazine followed by solid phase guanidinylation using *N*,*N*′-di-Boc-*N*′′-triflylguanidine in the presence of triethylamine (Scheme 2). The corresponding Dab-containing peptide synthesized using Fmoc–Dab(Mtt)–OH did not yield the intended products; coupling of this residue appeared to be difficult. For the Agp residues in peptide AgpAla, coupling



Scheme 1 Synthesis of appropriately protected Arg analogs for Fmoc-based solid phase peptide synthesis

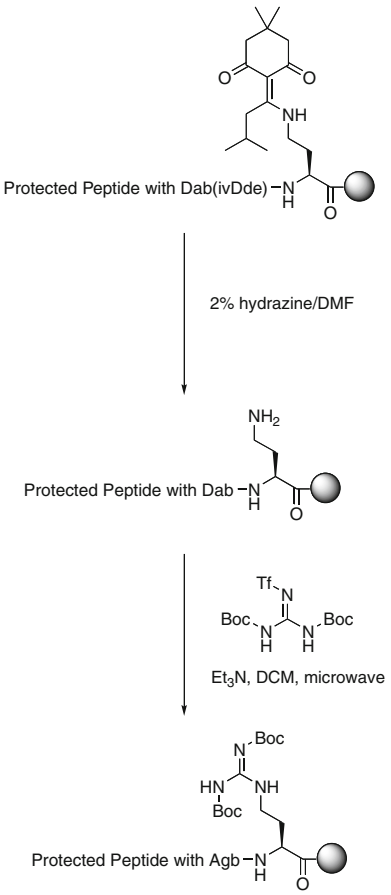
Table 1 Sequence of Ala-based peptides for determining the N-cap parameter, C-cap parameter, and helix propensity of Arg analogs with varying side chain length

Peptide	Sequence
NAgp01	Ac-**Agp** Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Gly Gly Tyr-NH_2
NAgb01	Ac-**Agb** Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Gly Gly Tyr-NH_2
NArg01	Ac-**Arg** Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Gly Gly Tyr-NH_2
NAgh01	Ac-**Agh** Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Gly Gly Tyr-NH_2
CAgp19	Ac-Tyr Gly Gly Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala **Agp**-NH_2
CAgb19	Ac-Tyr Gly Gly Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala **Agb**-NH_2
CArg19	Ac-Tyr Gly Gly Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala **Arg**-NH_2
CAgh19	Ac-Tyr Gly Gly Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala Lys Ala Ala Ala Ala **Agh**-NH_2
AgpAla	Ac-Tyr Gly Gly **Agp** Ala Ala Ala Ala **Agp** Ala Ala Ala Ala **Agp** Ala Ala Ala Ala **Agp**-NH_2
AgbAla	Ac-Tyr Gly Gly **Agb** Ala Ala Ala Ala **Agb** Ala Ala Ala Ala **Agb** Ala Ala Ala Ala **Agb**-NH_2
ArgAla	Ac-Tyr Gly Gly **Arg** Ala Ala Ala Ala **Arg** Ala Ala Ala Ala **Arg** Ala Ala Ala Ala **Arg**-NH_2
AghAla	Ac-Tyr Gly Gly **Agh** Ala Ala Ala Ala **Agh** Ala Ala Ala Ala **Agh** Ala Ala Ala Ala **Agh**-NH_2

The Arg analogs are shown in bold

using Fmoc–Agp(Boc)$_2$–OH gave very small amounts of desired peptide even with triple couple or microwave conditions. This is most likely due to steric hindrance caused by the close proximity of the Boc protecting groups to the backbone because of the short side chain length. Since insufficient material was obtained using Fmoc–Agp(Boc)$_2$–OH, a solid phase guanidinylation strategy was adopted (Scheme 3). The corresponding Dap-containing peptide was synthesized using Fmoc–Dap(Mtt)–OH (Dap, (*S*)-2,3-diaminopropionic acid; Mtt, 4-methyltrityl), then the Mtt protecting group was selectively removed using 1% trifluoroacetic acid/methylene chloride followed by solid phase guanidinylation using *N*,*N*′-di-Boc-*N*″-triflylguanidine in the presence of triethylamine under microwave conditions. We did not explore the synthetic route involving the ivDde protecting group for the Agp-containing peptides, because side chain to main chain migration may occur with Dap(ivDde)-containing peptides during solid phase synthesis.

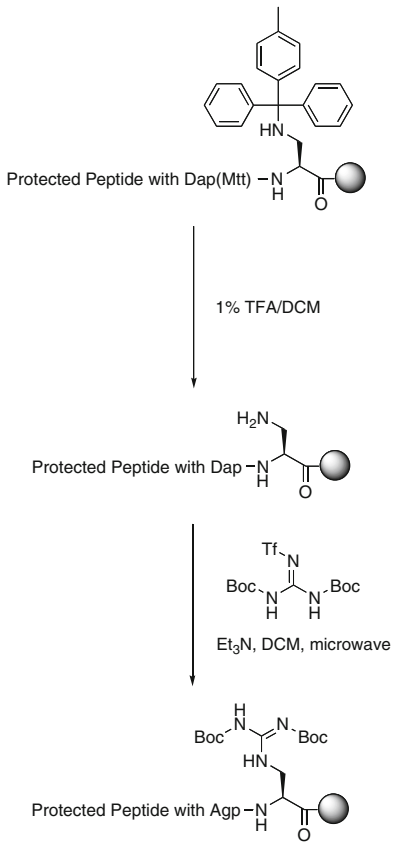
Upon cleavage with concomitant side chain deprotection, the peptides were confirmed by MALDI-TOF mass spectrometry and purified by reverse phase high performance liquid chromatography to >98% purity. The concentration of the peptides was determined by the Edelhoch method (Edelhoch 1967; Pace et al. 1995). The peptides bear multiple positive charges at neutral pH and the peptides are not amphiphilic. As such, these peptides should not associate in neutral aqueous buffers. Furthermore, analogous peptides have been shown to be monomeric in solution (Padmanabhan et al. 1990; Chakrabartty et al. 1994; Doig and Baldwin 1995; Chiu et al. 2006; Chiu and Cheng 2007; Cheng et al. 2010), and the CD spectrum of each peptide did not alter significantly between 80 and 160 μM. Therefore, the peptides are most likely monomeric in solution, and intermolecular interactions should not contribute significantly to the helical content of these peptides in solution.



Scheme 2 Synthesis of Agb-containing peptides by solid phase guanidinylation

Circular dichroism spectroscopy of the Arg analog-containing peptides

The circular dichroism (CD) spectrum for all peptides was acquired at pH 7 in the presence of 1 M NaCl to derive the fraction helix (f_{helix}) (Table 2 and Fig. 2); these were the



Scheme 3 Synthesis of Agp-containing peptides by solid phase guanidinylation

conditions for the experiments performed by Baldwin and coworkers (Chakrabartty et al. 1994). The mean residue ellipticity at 222 nm reflects the helical content of a given peptide. More specifically, the more negative the CD signal at 222 nm, the more helical the peptide. The helical content of the NXaa01 peptides followed the trend NAgh01 ~ NArg01 ~ NAgb01 > NAgp01 (Fig. 2a). This trend suggests similar N-capping abilities for the Arg analogs except Agp, which showed slightly weaker N-cap capabilities. The helical content of the CXaa19 peptides were quite similar but followed the trend CAgh19 > CArg19 > CAgb19 > CAgp19 (Fig. 2b). This trend suggests that C-capping ability for the Arg analogs increases with increasing side chain length. The helical content of the XaaAla peptides followed the trend ArgAla > AghAla > AgbAla > AgpAla (Fig. 2c).

Helix formation parameters for Arg analogs

The helix formation parameters including N-cap (n), C-cap (c), and helix propensity (w) were derived from the CD data using modified Lifson–Roig theory (Lifson and Roig 1961; Doig et al. 1994; Chakrabartty et al. 1994; Doig and Baldwin 1995; Chiu et al. 2006; Cheng et al. 2010) (Table 3). The corresponding free energies were also derived (Table 3). The N-cap n parameter was less than zero for all Arg analogs, suggesting extremely unfavorable energetics for placing these positively charged residues at the N-terminus. However, values less than zero do not bear physical meaning, therefore the n values were set to zero. The C-cap c parameter followed the trend Agh > Arg > Agb > Agp (Table 3). In other words, the c values increased with increasing side chain length. Interestingly, only Agh was clearly energetically favorable at the C-cap position, whereas Arg was indifferent; Agb was somewhat energetically unfavorable, but Agp was very unfavorable. The helix propensity w followed the trend Agp < Agb < Arg > Agh. Only the longer Arg and Agh were energetically favorable in the helical conformation, whereas the shorter Agb and Agp were unfavorable. However, Arg was more energetically favored in the helical conformation than Agh, highlighting the uniqueness of the Arg side chain length in helix formation.

Conformational analysis by molecular mechanics calculations

Conformational analysis was performed by molecular mechanics calculations to gain further insight into the helix capping and formation energetic trends of the Arg analogs with varying side chain length (Table 4). The peptides were composed of 20 alanines and one Arg analog at various positions. Each Arg analog was incorporated at the first four positions (N1–N4), the last four positions (C4–C1), and the center position of a 21-residue helical peptide. For all peptide models, the backbone was fixed to ideal helix dihedrals ($\phi = -57°$, $\psi = -47°$) (Edsall et al. 1966) to avoid drastic changes in the backbone conformation especially at the termini of the helix. For a given peptide, different side chain conformations were generated using various χ angle combinations. For each χ angle involving only sp^3 atoms such as χ_1, three possible low-energy staggered conformations were considered (Fig. 3): *gauche*− (60°, $g-$), *trans* (180°, t), and *gauche*+ (300°, $g+$) (McGregor et al. 1987; Dunbrack and Karplus 1993). For χ angles involving the planar guanidinium nitrogen (due to sp^2 hybridization of the nitrogen and carbon atoms), 12 different angles were considered: 0°, 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, and 330°. A combined total of 12,960 conformations were minimized using the CFF forcefield and analyzed.

The energy of the lowest energy conformations for peptides with the same Arg analog at different positions was then compared to gain information regarding the positional preference (Table 4). Incorporating a given Arg analog near the helix N-terminus resulted in higher energy conformers compared to incorporating the same Arg

Table 2 Mean residue ellipticity at 222 nm and fraction helix (f_{helix}) of Ala-based peptides containing Arg analogs with varying side chain length

Peptide	$[\theta]_{222}$	f_{helix}
NAgp01	−15800 ± 200	0.452 ± 0.006
NAgb01	−18200 ± 200	0.520 ± 0.006
NArg01	−18500 ± 200	0.529 ± 0.006
NAgh01	−18000 ± 200	0.514 ± 0.006
CAgp19	−17000 ± 200	0.485 ± 0.004
CAgb19	−17500 ± 200	0.500 ± 0.005
CArg19	−18100 ± 200	0.517 ± 0.005
CAgh19	−18400 ± 200	0.529 ± 0.004
AgpAla	−5600 ± 200	0.161 ± 0.007
AgbAla	−16000 ± 200	0.463 ± 0.005
ArgAla	−24400 ± 200	0.697 ± 0.006
AghAla	−21000 ± 200	0.594 ± 0.006

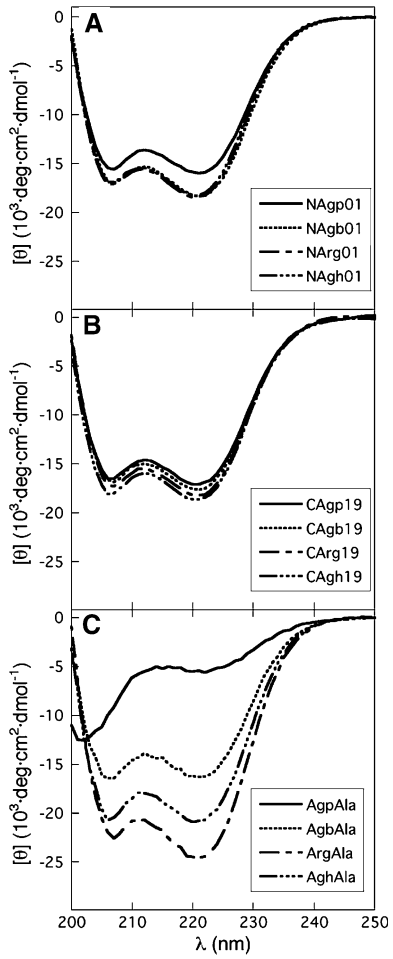


Fig. 2 Circular dichroism spectra of the peptides at pH 7 (273 K) in 1 mM phosphate, borate, and citrate buffer with 1 M NaCl. **a** NAgp01, NAgb01, NArg01, NAgh01. **b** CAgp19, CAgb19, CArg19, CAgh19. **c** AgpAla, AgbAla, ArgAla, AghAla

analog near the C-terminus or at the center of the helix. Furthermore, the energy difference between positioning a given Arg analog at N1 and C1/center was smaller for the shorter residues Agp and Agb compared to the longer residues Arg and Agh. These position-dependent energy trends are consistent with the experimental results (vide supra), showing unfavorable energetics for the N-cap parameter (n) of the Arg analogs compared to the C-cap parameter (c) and helix propensity (w).

Conformations within 4 kcal of the lowest energy conformer for each peptide were then examined in more detail (Table 4), because room temperature can provide up to 4 kcal mol^{-1} of thermal energy. The positional dependence for the weighed average energy (based on Boltzmann distribution) of the low energy ensembles followed the same trend as the energy for the lowest energy conformations. Interestingly, the number of conformations at positions N1 and N2 was consistently higher than the number of conformations at the other positions for peptides containing Agb, Arg, and Agh. This positional dependence in low energy conformers was barely present for Agp-containing peptides, perhaps because there were not many low energy conformers thereby suppressing the difference in number of conformations. Disregarding the Agp peptides, there were no low energy structures with $g-$ conformation at the χ_1 dihedral except for the N1 and N2 positions. This lack of the $g-$ conformation (at positions other than N1 and N2) is most likely because the Cα–Cβ vector inherently points toward the N-terminus of a helix, and the $g-$ conformation orients the side chain towards the upstream backbone resulting in unfavorable steric interactions with the i-3 position (Fig. 3). This unfavorable steric interaction remains present at the N3 and N4 positions, but is alleviated at the N1 and N2 positions due to the absence of the upstream i-3 residue. As for Agp, the side chain is so short that the guanidinium group is directly attached to the Cβ carbon. The planar geometry of the guanidinium group inherently directs the side chain away from the backbone, thereby avoiding this unfavorable steric clash. As such, the $g-$ conformation at the χ_1 dihedral is present in all the low energy conformers of the Agp-containing peptides. To further confirm these molecular mechanics results, we performed a survey on natural protein structures.

Survey of protein structures

The non-redundant protein structure database PDBselect (April 2009, 25% threshold) (Hobohm and Sander 1994; Griep and Hobohm 2010) was surveyed for Arg in the context of various helix-related structures (Table 5). The definition of the helical conformation was based on backbone dihedrals (Gunasekaran et al. 1998; Engel and DeGrado 2004; Cheng et al. 2010), and only helices with at

Table 3 Statistical mechanical helix formation parameters for Arg analogs with varying side chain length derived from experimentally measured fraction helix based on modified Lifson–Roig theory

Residue	n	c	w	$\Delta G^{b}_{\text{C-cap}}$	$\Delta G^{c}_{\text{helix}}$
Agp	0[a]	0.0730 ± 0.1036	0.289 ± 0.035	1.42	0.674
Agb	0[a]	0.497 ± 0.155	0.845 ± 0.039	0.380	0.0914
Arg	0[a]	1.06 ± 0.18	1.68 ± 0.08	−0.0316	−0.281
Agh	0[a]	1.48 ± 0.16	1.19 ± 0.06	−0.213	−0.0944

These statistical mechanical helix formation parameters are derived from experimental CD data of the peptides in Table 2; for sequences of the peptides see Table 1

[a] The parameter initially converged to a negative probability (which carries no physical meaning), therefore the value was set to zero

[b] $\Delta G_{\text{C-cap}} = -\text{RT ln } (c)$ reported in kcal mol^{-1}

[c] $\Delta G_{\text{helix}} = -\text{RT ln } (w)$ reported in kcal mol^{-1}

least six residues were considered to be a helix to avoid "helices" less than one turn (Cheng et al. 2007, 2010). A total of 4,418 protein chains and 666,086 residues were considered involving 17,622 helices and 236,790 helical residues. Approximately 33.7% of the residues in the database were helical (236,790 out of 666,086) (Cheng et al. 2010), similar to earlier surveys (Chou and Fasman 1974; Barlow and Thornton 1988). Interestingly, 44.4% of the Arg residues in the database were helical (14,866 out of 33,447) with a statistical propensity of 1.20 ± 0.01 (Z value 22.01; P value 2.31×10^{-107}). This result is different from an earlier survey (Chou and Fasman 1974), which found the statistical helix propensity for Arg to be 0.79, but similar to a more recent survey (Kumar and Bansal 1998), which found the statistical propensity for Arg to be 1.33. To gauge the significance of the occurrences for Arg at various positions in a helix, the statistical propensity was derived by dividing the occurrence by the corresponding expected occurrence. The expected occurrence was obtained by bootstrapping (Efron and Gong 1983) the entire database including all structures. 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 (Table 5), which was used to derive the P values (Klugh 1970; Kuebler and Smith 1976). Statistical propensities greater than unity represent higher occurrence than expected based on occurrence in all structures, whereas propensities less than unity indicate lower occurrence than the expected value. To survey Arg at various positions in a helix, we considered positions near the N-terminus, near the C-terminus, and internal positions. The residues near the helix termini were designated as follows:

N″–N′–N–N1–N2–N3–…–C3–C2–C1–C–C′–C″

The first residue of a helix with helical dihedral angles is N1, and the immediate preceding residue with non-helical dihedrals is designated N. The last residue of a helix with helical dihedral angles is C1, and the immediate downstream residue with non-helical dihedrals is designated C. This definition is similar but not identical to previous definitions for surveying helix capping (Presta and Rose 1988; Richardson and Richardson 1988; Aurora and Rose 1998; Gunasekaran et al. 1998).

The propensity for Arg to appear as the first three residues in a helix was clearly less than unity (Table 5), consistent with potentially repulsive interactions between the positively charged Arg side chain and the partial positive charge of the helix macrodipole (Blagdon and Goodman 1975; Sali et al. 1988). Furthermore, the propensity for the three positions prior to the helix was also less than unity, most likely due to unfavorable interactions with the helix macrodipole even though the positions are no longer in the helix, or perhaps inherently lower incorporation of Arg in non-helical structures. The propensity for Arg in the last three positions of a helix was higher than unity, consistent with potentially favorable interactions between the positively charged Arg side chain and the partial negative charge of the helix macrodipole (Blagdon and Goodman 1975; Sali et al. 1988). In contrast, the propensity for the three positions after the helix was less than unity, most likely due to the inherently lower incorporation of Arg in non-helical structures.

The side chain conformation of the Arg residues was then examined in detail (Table 6). Residues at positions upstream (N, N′, N″) and downstream (C, C′, C″) of the helix are connected to the helix by residues with non-helical backbone dihedrals at positions N and C, respectively. As such, Arg at these positions beyond the helix was not included in this analysis. Since the total occurrence of Arg at the different positions is different, the percentage for the three low energy conformations of the χ_1 dihedral at each position was derived to enable more meaningful comparisons (Table 6). For the χ_1 dihedral, the $g-$ conformation occurred less than t or $g+$ at all positions, consistent with our conformational analysis by molecular mechanics calculations. This distribution in conformation is most likely due to sterics (vide supra; Fig. 3). However, the non-ideal helical dihedrals in natural proteins may

Table 4 Low energy conformations from conformational analysis of Arg analog-containing peptides by molecular mechanics calculations

Residue	Position	Lowest energy conformer Energy (kcal)	Conformations within 4 kcal of lowest			
			Number of conformations[a]	χ_1[b]		
				g−	*t*	*g*+
Agp	N1	−33.7	15	4	5	6
Agp	N2	−33.8	14	4	4	6
Agp	N3	−34.3	12	3	4	5
Agp	N4	−34.7	13	3	5	5
Agp	Center	−35.3	11	3	3	5
Agp	C4	−35.5	12	3	4	5
Agp	C3	−35.6	12	3	4	5
Agp	C2	−35.6	12	3	4	5
Agp	C1	−36.0	15	4	5	6
Agb	N1	−35.0	44	14	16	14
Agb	N2	−35.2	44	14	16	14
Agb	N3	−35.7	30	0	14	16
Agb	N4	−36.3	32	0	16	16
Agb	Center	−36.5	34	0	17	17
Agb	C4	−36.7	33	0	16	17
Agb	C3	−36.7	31	0	15	16
Agb	C2	−36.8	33	0	15	18
Agb	C1	−37.4	32	0	14	18
Arg	N1	−37.9	113	26	43	44
Arg	N2	−38.1	118	29	42	47
Arg	N3	−39.3	69	0	29	40
Arg	N4	−39.7	60	0	30	30
Arg	Center	−40.2	55	0	26	29
Arg	C4	−40.4	58	0	27	31
Arg	C3	−40.4	54	0	25	29
Arg	C2	−40.4	58	0	27	31
Arg	C1	−40.8	48	0	18	30
Agh	N1	−41.0	281	66	103	112
Agh	N2	−41.3	292	77	97	118
Agh	N3	−42.4	178	0	74	104
Agh	N4	−43.0	155	0	70	85
Agh	Center	−43.7	124	0	47	77
Agh	C4	−43.9	116	0	42	74
Agh	C3	−43.8	108	0	37	71
Agh	C2	−43.9	113	0	41	72
Agh	C1	−44.1	113	0	31	82

[a] The number of conformations within 4 kcal of the lowest energy conformer for each peptide

[b] The number of conformations for the three staggered χ_1 dihedrals (Fig. 3): *g*− (*gauche*–, $\chi_1 = 60°$), *t* (*trans*, $\chi_1 = 180°$), *g*+ (*gauche*+, $\chi_1 = 300°$)

avoid this unfavorable steric interaction, leading to some representation in the survey. Low occurrence of *g*– was particularly significant at the internal, C3, and N3 positions, with slightly higher percentage of *g*– for the positions near the helix termini (C1, C2, N1, N2), most likely due to end fraying effects. Importantly, there is a higher percentage of *g*– for the positions N1 and N2 compared to C1 and C2. This can be attributed to the absence of the *i*-3 helical residue for residues at positions N1 and N2, inherently avoiding unfavorable steric interactions (Fig. 3).

Discussion

All Arg analogs, regardless of side chain length, exhibited extremely unfavorable energetics at the N-cap position of a helix (i.e. *n* parameter; Table 3). This may be due to destabilizing interactions with the partial positive charge of the helix macrodipole at the N-terminal end (Blagdon and Goodman 1975; Sali et al. 1988). Accordingly, one would a priori predict that positively charged amino acids should be favored at the C-cap position (i.e. *c* parameter)

due to stabilizing interactions with the partial negative charge of the helix macrodipole (Blagdon and Goodman 1975; Sali et al. 1988). However, only Agh was energetically favorable at the C-cap position. Arg was energetically indifferent, but Agb and Agp were energetically unfavorable. Apparently, short analogs of Arg cannot provide stabilization through interaction with the helix macrodipole at the C-terminal end of a helix, perhaps due to the need to adopt alternative non-helical conformations for the favorable interaction. Regardless, the *c* parameter followed the trend Agh > Arg > Agb > Agp (Table 3). This trend is not exactly the same as the findings by (Schneider and DeGrado 1998) on linear alkyl aminoguanidines, most likely because of the lack of the carboxyamide group and stereocenter in the DeGrado study. Interestingly, the helix propensity *w* followed the trend Agp < Agb < Arg > Agh (Table 3). This is in sharp contrast with previous studies on positively charged Lys analogs (Padmanabhan et al. 1996), showing that helix propensity *w* increases with increasing side chain length (Padmanabhan et al. 1996). Similarly, the helix propensity of linear hydrophobic residues also increases with increasing side chain length (Padmanabhan and Baldwin 1991). Accordingly, our result on helix propensity highlights the uniqueness of Arg side chain length in favoring helix formation.

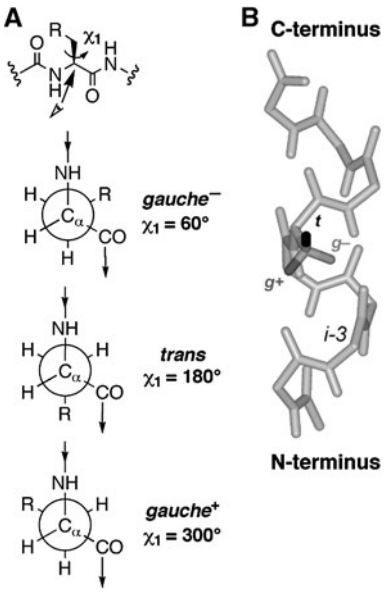


Fig. 3 **a** Newman projections of the three low energy χ_1 dihedrals *gauche–*, *trans*, and *gauche+* **b** Three-dimensional model of a generic helix backbone with the three low energy χ_1 dihedrals highlighted for the center residue

Conformational analysis by molecular mechanics calculations showed that incorporating a given Arg analog at the N-terminus of a helix is higher in energy (and thus less stable) compared to incorporating at the center or C-terminus (Table 4), consistent with our experimentally derived N-capping, C-capping, and helix propensity energetics (Table 3). Furthermore, the difference between the energy of

Table 5 Survey of arginine in helices and near the termini of helices with at least six residues in protein structures

Position	Occurrence[a]	Propensity[b]	*Z* value[c]	*P* value[d]
N″	736	0.884 ± 0.030	−3.45	5.82×10^{-4}
N′	660	0.780 ± 0.026	−6.57	5.03×10^{-11}
N	485	0.548 ± 0.018	−13.7	8.85×10^{-43}
N1	835	0.946 ± 0.031	−1.64	1.01×10^{-1}
N2	813	0.920 ± 0.030	−2.47	1.35×10^{-2}
N3	696	0.789 ± 0.026	−6.47	9.80×10^{-11}
Internal	8,167	1.247 ± 0.015	20.53	1.16×10^{-93}
C3	1,243	1.41 ± 0.05	12.5	5.80×10^{-36}
C2	1,318	1.49 ± 0.05	15.0	8.54×10^{-51}
C1	1,103	1.25 ± 0.04	7.68	1.59×10^{-14}
C	723	0.819 ± 0.027	−5.54	3.02×10^{-8}
C′	723	0.874 ± 0.029	−3.73	1.91×10^{-4}
C″	768	0.951 ± 0.032	−1.44	1.50×10^{-1}

[a] The occurrence of Arg at the designated position (in helices and near the termini of helices) in the non-redundant protein structure database PDBselect (Hobohm and Sander 1994; Griep and Hobohm 2010) (April 2009, 25% threshold). Only helices with at least six consecutive helical residues were considered to avoid counting any given residue more than once due to proximity to both N- and C-termini

[b] The occurrence divided by the corresponding expected value. The expected value was obtained by bootstrapping the entire PDBselect structure database 100,000 times

[c] The difference between the occurrence and the expected value divided by the SD for the expected value

[d] The probability that the occurrence and expected occurrence are the same based on the SD obtained from bootstrapping assuming a Gaussian distribution

Table 6 Occurrence and percent occurrence of χ_1 dihedral conformations of arginine in helices and near the termini of helices with at least six residues in protein structures

Position	χ_1		
	g−	*t*	*g*+
N1	88 (11%)	409 (51%)	308 (38%)
N2	87 (11%)	334 (43%)	358 (46%)
N3	23 (3%)	336 (49%)	321 (47%)
Internal	247 (3%)	3625 (46%)	4079 (51%)
C3	56 (5%)	496 (41%)	667 (55%)
C2	107 (8%)	473 (37%)	713 (55%)
C1	90 (8%)	147 (14%)	850 (78%)

The occurrence of Arg at the designated position (in helices and near the termini of helices) in the non-redundant protein structure database PDBselect (Hobohm and Sander 1994; Griep and Hobohm 2010) (April 2009, 25% threshold). Only helices with at least six consecutive helical residues were considered to avoid counting any given residue more than once due to proximity to both N- and C-termini. The percent occurrence is presented in parentheses. The percent occurrence for each conformation at a given position was calculated by dividing the occurrence by the sum of the occurrences for all three conformations for that given position

incorporating a given Arg analog at the N-terminus and C-terminus (or center) increases with increasing side chain length (Table 4), again consistent with our experimental results (Table 3). However, the slight energetic difference between the experimentally determined helix propensity and C-cap energetics (less than 1 kcal mol^{-1}) could not be clearly distinguished with the conformational analysis using molecular mechanics calculations as performed in this study. Although this could be due to fixing the backbone to an ideal helix conformation, results from the same calculation without fixing the backbone gave similar energetic differences. Another possibility is that the simple conformational analysis was performed using an implicit solvent model. More accurate modeling involving explicit water may be necessary to capture the subtle difference between helix propensity and C-cap parameter. However, the inherent error for these calculations may be too large to distinguish the energetic difference between the helix propensity and C-cap parameter. Regardless, the *g*− conformation for the χ_1 dihedral is not present except at the N1 or N2 position, or in the Agp-containing peptides. Importantly, these conformational results from molecular mechanics calculations are consistent with the survey results on natural protein structures.

We have focused on helix formation energetics including N-capping, C-capping and helix propensity in this study. However, protein secondary structures include helices, sheets, and turns. Although 44.4% of the Arg residues adopt a helical conformation (vide supra), the other structures cannot be ignored. Therefore, studies on the stability of non-helix structures such as sheets and turns would be necessary for a more complete understanding of the structural effect of the Arg side chain length, to enable facile use of Arg analogs in various applications.

Conclusions

We have demonstrated the synthesis of 19-residue peptides containing Arg analogs with varying side chain length at different positions using Fmoc-based chemistry. Straightforward solid phase peptide synthesis was sufficient for incorporating Agh, whereas solid phase guanidinylation using orthogonal protection strategies were necessary for Agb and Agp-containing peptides. Importantly, the effect of Arg side chain length on helix formation and capping was investigated in Ala-based peptides by CD. All four Arg analogs were energetically unfavorable for N-capping. Interestingly, Arg was the most energetically favorable at internal helix positions, whereas Agh was the most energetically favorable for C-capping. Agb and Agp were not suitable for helical structures, but may be useful for sheets or turns. Molecular mechanics calculations and a survey on a non-redundant protein structure database were consistent with our experimental findings, and provided conformational insight into the experimentally observed trends. The calculations and survey both showed that the *g*− conformation for the χ_1 dihedral is present for the first two residues at the N-terminus of helices, but not favored in the center or the C-terminus of helices. These thorough studies on the effect of Arg side chain length on helix formation and capping should serve as the foundation for the use of Arg analogs in developing Arg-related bioactive compounds and technologies.

Acknowledgments 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., CHE0809633), National Taiwan University (R.P.C.), National Science Council (R.P.C., NSC-97-2113-M-002-019-MY2, NSC-98-2119-M-002-025, NSC-99-2113-M-002-002-MY2; Y.J.W., NSC-98-2815-C-002-056-M). 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.

References

Argos P (1988) An investigation of protein subunit and domain interfaces. Protein Eng 2:101–113

Aurora R, Rose GD (1998) Helix capping. Protein Sci 7:21–38

Barlow DJ, Thornton JM (1988) Helix geometry in proteins. J Mol Biol 201(3):601–619

Baumgart S, Lindner Y, Kühne R, Oberemm A, Wenschuh H, Krause E (2004) The contributions of specific amino acid side chains to signal intensities of peptides in matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 18:863–868

Blagdon DE, Goodman M (1975) Mechanisms of protein and polypeptide helix initiation. Biopolymers 14:241–245

Blaum BS, Deakin JA, Johansson CM, Herbert AP, Barlow PN, Lyon M, Uhrin D (2010) Lysine and arginine side chains in glycosaminoglycan–protein complexes investigated by NMR, cross-linking, and mass spectrometry: a case study of the factor H-heparin interaction. J Am Chem Soc 132:6374–6381

Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9

Chakrabarti P (1994) Conformations of arginine and lysine side chains in association with anions. Int J Peptide Protein Res 43:284–291

Chakrabartty A, Kortemme T, Baldwin RL (1994) Helix propensities of the amino-acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. Protein Sci 3(5):843–852

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

Cheng RP, Girinath P, Suzuki Y, Kuo H-T, Hsu H-C, Wang W-R, Yang P-A, Gullickson D, Wu C-H, Koyack MJ, Chiu H-P, Weng Y-J, Hart P, Kokona B, Fairman R, Lin T-E, Barrett O (2010) Positional effects on helical Ala-based peptides. Biochemistry 49:9372–9384

Chhabra SR, Hothi B, Evans DJ, White PD, Bycroft BW, Chan WC (1998) An appraisal of new variants of Dde amine protecting group for solid phase peptide synthesis. Tetrahedron Lett 39:1603–1606

Chiu H-P, Cheng RP (2007) Chemoenzymatic synthesis of (*S*)-hexafluoroleucine and (*S*)-tetrafluoroleucine. Org Lett 9:5517–5520

Chiu HP, Suzuki Y, Gullickson D, Ahmad R, Kokona B, Fairman R, Cheng RP (2006) Helix propensity of highly fluorinated amino acids. J Am Chem Soc 128(49):15556–15557

Chou PY, Fasman GD (1974) Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry 13(2):211–222

Diss ML, Kennan AJ (2008) Orthogonal recognition in dimeric coiled coils via buried polar-group modulation. J Am Chem Soc 130:1321–1327

Doig AJ, Baldwin RL (1995) N- and C-capping preferences for all 20 amino acids in α-helical peptides. Protein Sci 4:1325–1336

Doig AJ, Chakrabartty A, Klingler TM, Baldwin RL (1994) Determination of free-energies of N-capping in a-helices by modification of the Lifson–Roig helix-coil theory to include N-capping and C-capping. Biochemistry 33(11):3396–3403

Drake B, Patek M, Lebl M (1994) A convenient preparation of monosubstituted *N*,*N*'-di(Boc)-protected guanidines. Synthesis: 579–582

Dunbrack RL, Karplus M (1993) Backbone-dependent rotamer library for proteins—application to side-chain prediction. J Mol Biol 230(2):543–574

Edelhoch H (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry 6(7):1948–1954

Edsall JT, Flory PJ, Kendrew JC, Liquori AM, Némethy G, Ramachandran GN, Scheraga HA (1966) A proposal of standard conventions and nomenclature for the description of polypeptide conformations. J Mol Biol 15:399–407

Efron B, Gong G (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation. Am Stat 37(1):36–48

Engel DE, De Grado WF (2004) Amino Acid propensities are position-dependent throughout the length of α-helices. J Mol Biol 337:1195–1205

Feichtinger K, Sings HL, Baker TJ, Matthews K, Goodman M (1998a) Triurethane-protected guanidines and triflyldiurethane-protected guanidines: new reagents for guanidinylation reactions. J Org Chem 63:8432–8439

Feichtinger K, Zapf C, Sings HL, Goodman M (1998b) Diprotected triflyguanidines: a new class of guanidinylation reagents. J Org Chem 63:3804–3805

Fields GB, Noble RL (1990) Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbonyl amino-acids. Int J Pept Protein Res 35(3):161–214

Griep S, Hobohm U (2010) PDBselect 1992–2009 and PDBfilter-select. Nucleic Acids Res 38:D318–D319

Gunasekaran K, Nagarajaram HA, Ramakrishnan C, Balaram P (1998) Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. J Mol Biol 275:917–932

Hirsch AKH, Fischer FR, Diederich F (2007) Phosphate recognition in structural biology. Angew Chem Int Ed 46:338–352

Ho K-C, Sun C-M (1999) Liquid phase parallel synthesis of guanidines. Bioorg Med Chem Lett 9:1517–1520

Hobohm U, Sander C (1994) Enlarged representative set of protein structures. Protein Sci 3(3):522–524

Hu Z, Ma B, Wolfson H, Nussinov R (2000) Conservation of polar residues as hot spots at protein interfaces. Proteins Struct Funct Genet 39:331–342

Janin J, Miller S, Chothia C (1988) Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol 204:155–164

Jones S, Thornton JM (1995) Protein–protein interactions: a review of protein dimer structures. Prog Biophys Mol Biol 63:31–65

Katritzky AR, Rogovoy BV (2005) Recent developments in guanylating agents. ARKIVOC iv:49–87

Klugh HE (1970) Statistics-the essentials for research. In: Statistics-The essentials for research. John Wiley & Sons Inc, New York. pp 350

Kuebler RR, Smith H (1976) Statistics-A beginning. In: Statistics-A beginning. John Wiley & Sons, Inc, New York. pp 302

Kumar S, Bansal M (1998) Dissecting α-helices: position-specific analysis of α-helices in globular proteins. Proteins Struct Funct Genet 31:460–476

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

McGregor MJ, Islam SA, Sternberg MJE (1987) Analysis of the relationship between side-chain conformation and secondary structure in globular-proteins. J Mol Biol 198(2):295–310

Mitchell DJ, Kim DT, Steinman L, Fathman CG, Rothbard JB (2000) Polyarginine enters cells more efficiently than other polycationic homopolymers. J Pept Res 56:318–325

Moroni M, Koksch B, Osipov SN, Crucianelli M, Frigerio M, Bravo P, Burger K (2001) First synthesis of totally orthogonal protected α-(trifluoromethyl)- and α-(difluoromethyl)arginines. J Org Chem 66:130–133

Nishikaze T, Takayama M (2006) Cooperative effect of factors governing molecular ion yields in desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 20:376–382

Northern TR, Lee J-C, Hoang L, Raymond J, Hwang D-R, Yannone SM, Wong CH, Siuzdak G (2008) A nanostructure-initiator mass spectrometry-based enzyme activity assay. Proc Natl Acad Sci USA 105:3678–3683

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption-coefficient of a protein. Protein Sci 4(11):2411–2423

Padmanabhan S, Baldwin RL (1991) Straight-chain non-polar amino acids are good helix-formers in water. J Mol Biol 219:135–137

Padmanabhan S, Marqusee S, Ridgeway T, Laue TM, Baldwin Robert L (1990) Relative helix-forming tendencies of nonpolar amino acids. Nature 344:268–270

Padmanabhan S, York EJ, Stewart JM, Baldwin RL (1996) Helix propensities of basic amino acids increase with the length of the side-chain. J Mol Biol 257(3):726–734

Poss MA, Iwanowicz E, Reid JA, Lin J, Gu Z (1992) A mild and efficient method for the preparation of guanidines. Tetrahedron Lett 40:5933–5936

Pradhan TK, Joosten A, Vasse J-L, Bertus P, Karoyan P, Szymoniak J (2009) A concise stereoselective synthesis of 2-substituted 1-aminocyclopropanecarboxyllic acids. Eur J Org Chem 5072–5078

Presta LG, Rose GD (1988) Helix signals in proteins. Science 240(4859):1632–1641

Richardson JS, Richardson DC (1988) Amino acid preferences for specific locations at the ends of α helices. Science 240:1648–1652

Robinson S, Roskamp EJ (1997) Solid phase synthesis of guanidines. Tetrahedron 53:6697–6705

Sali D, Bycroft M, Fersht AR (1988) Stabilization of protein structure by interaction of alpha-helix dipole with a charged side chain. Nature 355:740–743

Schneider JP, DeGrado WF (1998) The design of efficient α-helical C-capping auxiliaries. J Am Chem Soc 120:2764–2767

Shey J-Y, Sun C-M (1998) Soluble polymer-supported synthesis of *N*,*N*-di(Boc)-protected guanidines. Synlett 1423–1425

Shinohara Y, Furukawa J, Niikura K, Miura N, Nishimura S-I (2004) Direct N-glycan profiling in the presence of tryptic peptides on MALDI-TOF by controlled ion enhancement and suppression upon glycan-selective derivatization. Anal Chem 76:6989–6997

Tsai C-J, Lin SL, Wolfson HJ, Nussinov R (1997) Studies of protein–protein interfaces: a statistical analysis of the hydrophobic effect. Protein Sci 6:53–64

Umezawa N, Gelman MA, Haigis MC, Raines RT, Gellman SH (2001) Translocation of a β-peptide across cell membranes. J Am Chem Soc 124:368–369

Wender PA, Mitchell DJ, Pattabiraman K, Pelkey ET, Steinman L, Rothbard JB (2000) The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. Proc Natl Acad Sci USA 97: 13003–13008

Wender PA, Rothbard JB, Jessop TC, Kreider EL, Wylie BL (2002) Oligocarbamate molecular transporters: design, synthesis, and biological evaluation of a new class of transporters for drug delivery. J Am Chem Soc 124:13382–13383

Wender PA, Galliber WC, Goun EA, Jones JR, Pillow TH (2008) The design of guanidinium-rich transporters and their internalization mechanisms. Adv Drug Deliv Rev 60:452–472

Yong YF, Kowalski JA, Lipton MA (1997) Facile and efficient guanylation of amines using thioureas and Mukaiyama's reagent. J Org Chem 62:1540–1542

Yong YF, Kowalski JA, Thoen JC, Lipton MA (1999) A new reagent for solid and solution phase synthesis of protected guanidines from amines. Tetrahedron Lett 40:53–56

Zhang Y, Kennan AJ (2001) Efficient introduction of protected guanidines in BOC solid phase peptide synthesis. Org Lett 3:2341–2344

Zhao Y, Kent SBH, Chait BT (1997) Rapid, sensitive structure analysis of oligosaccharides. Proc Natl Acad Sci USA 94:1629–1633