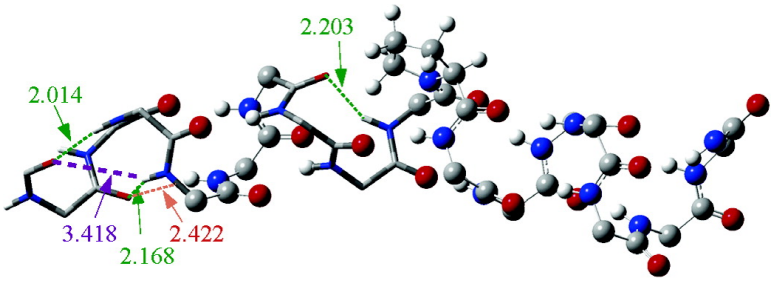
Article

The Energetic and Structural Effects of Single Amino Acid Substitutions upon Capped -Helical Peptides Containing 17 Amino Acid Residues. An ONIOM DFT/AM1 Study

Robert Wieczorek, and J. J. Dannenberg

J. Am. Chem. Soc., **2005**, 127 (49), 17216-17223• DOI: 10.1021/ja052689j • Publication Date (Web): 16 November 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 5 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/16/2005

The Energetic and Structural Effects of Single Amino Acid Substitutions upon Capped α-Helical Peptides Containing 17 Amino Acid Residues. An ONIOM DFT/AM1 Study

Robert Wieczorek† and J. J. Dannenberg*

Contribution from the Department of Chemistry, City University of New York, Hunter College and the Graduate School, 695 Park Avenue, New York, New York 10021

Received April 26, 2005; E-mail: jdannenberg@gc.cuny.edu.

Abstract: We evaluate the effect of the amino acid mutations of glycine, leucine, valine, phenylalanine, serine, and proline for the 10th alanine in the capped peptide, acetly$(Ala)_{17}NH_2$, upon the energies of the α-helices and β-strands using ONIOM DFT/AM1 molecular orbital calculations. The relative stabilities of the α-helix (to the β-strand) derive from the differences between the effects upon not only the helix but the strand as well. Thus, Ala → Pro significantly destabilizes both but destabilizes the α-helix more, while Ala → Gly stabilizes both but stabilizes the β-strand more. The theoretical results are discussed in the context of the known experimental reports. We suggest that the solvation of the unfolded state drives the helix/coil equilibrium in solution.

Protein folding remains one of the most intriguing subjects of biochemical research. The fact that all peptides do not form distinctly folded structures has long been established. Only certain sequences of amino acids (primary structures) have this capability. The folded structures contain secondary structural motifs (such as helices, sheets, etc.) which are further conformationally arranged (tertiary structure) into the completely folded protein. Simple changes of one or more amino acids for one or more others at specific positions in the primary sequence often do not lead to significant changes in the structure of a folded protein. However, each such change must, in principle, alter the relative energies of the folded protein relative to some arbitrarily chosen standard state (such as the appropriate collection of monomeric amino acids). When the aggregate energetic perturbations of enough changes in the primary sequence become significant, the perturbed peptide will no longer fold into the original structure, because other structures will become energetically competitive with the original. Modifications of the primary sequence can provoke relative energetic differences in the secondary as well as the tertiary structures. Clearly, a change in secondary structure would be likely to cause a major perturbation of the tertiary structure as well.

In this paper, we shall present ONIOM (DFT/AM1) calculations on the relative energetic effects of substituting various amino acids for the 10th alanine (Ala) in the α-helical form of the capped 17-amino acid sequence acetyl$(Ala)_{17}NH_2$ to form acetyl$(Ala)_9X(Ala)_7NH_2$, where "X" represents the new amino acid. We chose this position near the middle of the sequence as a model for evaluating the energetic and structural perturbations on an α-helical region that should be relatively free of end effects. The capped 17-amino acid model is one long enough to be robustly α-helical. If the α-helical structure unraveled upon substitution, we could not evaluate the energetic perturbation quantitatively. Earlier studies (both experimental[1−3] and theoretical[4]) have shown that short peptide sequences do not form α-helices. We needed to choose a model that was sufficiently stable to accommodate individual potentially destabilizing interactions.

When determining the stability of a conformation of a molecule (such as an α-helix), one must carefully define some reference structure. As we shall see, the apparent stabilities of the present α-helices depend on the particular reference chosen. There have been numerous experimental reports of the effects of changing individual amino acids on the stability of α-helices in solution.[5−14] These studies, while extremely useful, cannot determine the extent to which the observed stabilities are due to the effect upon the unfolded structures, solvation (of both helical and unfolded structures), or the change in the intrinsic energy of the α-helix itself. Clearly, we need to understand how

† On leave from Faculty of Chemistry, University of Wroclaw, F. Joliot-Curie 14, Wroclaw, Poland.

(1) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(2) Wallimann, P.; Kennedy, R. J.; Miller, J. S.; Shalongo, W.; Kemp, D. S. *J. Am. Chem. Soc.* **2003**, *125*, 1203−1220.
(3) Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 945−962.
(4) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2004**, *126*, 14198−205.
(5) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648−1652.
(6) Park, S.-H.; Shalongo, W.; Stellwagen, E. *Biochemistry* **1993**, *32*, 7048−7053.
(7) Petukhov, M.; Muñoz, V.; Yumoto, N.; Yoshikawa, S.; Serrano, L. *J. Mol. Biol.* **1998**, *278*, 279−289.
(8) Petukhov, M.; Uegaki, K.; Yumoto, N.; Serrano, L. *Protein Sci.* **2002**, *11*, 766−777.
(9) Petukhov, M.; Uegaki, K.; Yumoto, N.; Yoshikawa, S.; Serrano, L. *Protein Sci.* **1999**, *8*, 2144−2150.
(10) Rohl, C. A.; Chakrabartty, A.; Baldwin, R. L. *Protein Sci.* **1996**, *5*, 2623−2637.
(11) Cochran, D. A. E.; Penel, S.; Doig, A. J. *Protein Sci.* **2001**, *10*, 463−470.
(12) Sun, J. K.; Penel, S.; Doig, A. J. *Protein Sci.* **2000**, *9*, 750−754.
(13) Cochran, D. A. E.; Doig, A. J. *Protein Sci.* **2001**, *10*, 1305−1311.
(14) Iqbalsyah, T. M.; Doig, A. J. *Protein Sci.* **2004**, *13*, 32−9.

10.1021/ja052689j CCC: $30.25

peptides and proteins behave in their natural environment, which generally resembles aqueous solution more than the gas phase. However, an accurate understanding of how the environment affects the stabilities of α-helices requires an understanding of the isolated (unsolvated) system. The environment of a particular α-helix in a protein often is quite different from aqueous solution because of the particular manner in which proteins fold. For example, helices often form bundles with a significant part of each shielded from the bulk solvent. Furthermore, the in vivo environment contains many other ingredients besides water.

One must also consider the effect of the amino acid change upon the competing conformations. For example, a modification that stabilizes an α-helix may stabilize another conformation even more, causing the α-helical structure to no longer be experimentally observable. Thus, changing an amino acid could favor the helix by stabilizing it, destabilizing the alternative structures or a combination of both.

We have chosen two different reference points for comparison to the α-helices: (1) the component amino acids and (2) extended β-strands. Both of these reference points are quite arbitrary, but each can be well-defined. The energetic reference to the component amino acids can be obtained from an imaginary condensation polymerization involving the requisite amino acids that yields the polypeptide plus one water molecule for each bond formed. The appropriate groups needed for capping the peptides can easily be included in the energetic calculation. Such a reaction for the formation of acetyl-$(Ala)_{17}NH_2$ would be

$$17\text{Ala} + \text{NH}_3 + \text{CH}_3\text{COOH} \rightarrow \text{acetyl(Ala)}_{17}\text{NH}_2 + 17\text{H}_2\text{O} \quad (1)$$

The relative energy of the peptide with respect to the components would then be

$$E_{\text{rel}} = E_{\text{peptide}} - 17E_{\text{Ala}} - E_{\text{ammonia}} - E_{\text{acetic acid}} - 17E_{\text{water}} \quad (2)$$

While well-defined, this definition of the peptide energy lacks a direct connection to the kinds of experimental equilibria between the helical and nonhelical peptide conformations for which data are available. The common reference of a so-called "random coil" is rather poorly defined as it would be a complex mixture of numerous conformations. Furthermore, this mixture would be different for each peptide with a different amino acid residue, X. Several groups have recently criticized the original notion of the random coil.[1,15,16] For these reasons, we found it convenient to define a single reference conformation that might be a representative conformation of the random coil. For this, we chose the fully extended β-strand. This choice is completely arbitrary. The real population of conformations that contribute to the random coil will vary with the composition of the peptide, the temperature, and the solvent, while the structure of the α-helix remains relatively impervious to these influences for a reasonably wide range of conditions. Our choice of the arbitrary extended β-strand as a reference avoids the virtually intractable variation in the nature of the random coil.

The present study considers the following amino acids as X: Gly, Leu, Val, Phe, Ser, and Pro. We chose these amino acids for the following reasons. Gly, Leu, and Val provide examples of amino acids with alkyl side chains. We have previously suggested that C−H···O hydrogen bonds from β-CH may play an energetic role.[17,18] Of these, only Val and Leu have such a β-CH. In addition, Leu provides a conformational variation that Val does not. Phe provides another different opportunity for C−H···O interactions. The ortho H's of Phe might have stabilizing CH···O interactions that are similar to those of the alkyl side chains. Ser provides an example of an amino acid whose OH might form a H-bond. Finally, Pro is unique in that it lacks an amido hydrogen. This much studied amino acid[19−23] has a profound influence upon peptide structures.[10] The side chains of the amino acids are in their fully extended conformations where relevant.

There have been several previous reports of theoretical calculations of α-helical peptides.[4,17,22,24−32] We have previously reported the effects of mutating one or more natural amino acids at a time for alanine in acetyl$(Ala)_{17}NH_2$.[17] The mutation of lactic acid for Ala in similar smaller α-helices has also been reported.[32]

Calculational Details

We used the ONIOM[33,34] method as programmed in the Gaussian G03[35] and 98[36] suites of computer programs. ONIOM divides the system into up to three segments which can be treated at different levels of calculational complexity. Thus, one can treat the essential part of the system at the high level, while the less critical parts of the system might be calculated at the medium or low level. For this study, we only used two levels (high and medium). We treated the cores of the α-helices (equivalent to a corresponding peptide containing only glycines) at the high level, with only the side chain groups that distinguish the residues from glycine at the medium level. The high level used hybrid DFT methods at the B3LYP/D95(d,p) level. This combines Becke's three-parameter functional,[37] with the nonlocal correlation provided by the correlation functional of Lee et al.[38] In the ONIOM method, there are unsatisfied valences in the high level at the interface between it and the medium level. These valences were satisfied by using the default method of capping them with a hydrogen atom in the direction of the connecting atom in the medium level with a C−H distance of 0.723886

(15) Pappu, R. V.; Srinivasan, R.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12565−12570.
(16) Baldwin, R. L.; Zimm, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12391−12392.
(17) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, 8124−29.
(18) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, 14065−71.
(19) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 698−705.
(20) Dempsey, C. E. *Biochemistry* **1992**, *31*, 4705−12.
(21) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341−350.
(22) Bour, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45−59.
(23) Guo, H.; Beahm, R. F.; Guo, H. *J. Phys. Chem. B* **2004**, *108*, 18065−18072.
(24) Ireta, J.; Neugebauer, J.; Scheffler, M.; Rojo, A.; Galvan, M. *J. Phys. Chem. B* **2003**, *107*, 1432−1437, 9616.
(25) Mohle, K.; Hofmann, H.-J. *J. Pept. Res.* **1998**, *51*, 19−28.
(26) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2000**, *256*, 15−27.
(27) Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T.-H.; Perczel, A.; Rashin, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **2001**, *123*, 6054−60.
(28) Wu, X.; Wang, S. *J. Phys. Chem. B* **2001**, *105*, 2227−2235.
(29) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313−5319.
(30) Kubelka, J.; Gangani, R. A.; Silva, D.; Keiderling, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 5325−5332.
(31) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2004**, *126*, 12278−9.
(32) Cieplak, A. S.; Surmeli, N. B. *J. Org. Chem.* **2004**, *69*, 3250−3261.
(33) Morokuma, K. *Bull. Korean Chem. Soc.* **2003**, *24*, 797−801.
(34) Vreven, T.; Morokuma, K. *J. Chem. Phys.* **2000**, *113*, 2969−2975.
(35) Frisch, M. J.; et al. *GAUSSIAN 03*, revision C01; Gaussian, Inc.: Pittsburgh, PA, 2003.
(36) Frisch, M. J.; et al. *Gaussian 98*, revision A.11, revision A.10; Gaussian, Inc.: Pittsburg, PA, 1998−2001.
(37) Becke, A. D. *J. Chem Phys.* **1993**, *98*, 5648.
(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

Table 1. Energies (kcal/mol) for Formation from Amino Acids (eq 1) for Each Strand and Helix and Comparisons for Each Mutation with All Alanine

	ΔE from AA's		ΔE from $(Ala)_{17}$		
X	β-strand	-α-helix	β-strand	-α-helix	$\Delta\Delta E$ helix
Ala	−3.77	−36.35			
Gly	−5.35	−37.53	−1.58	−1.18	0.40
Leu	−2.96	−36.69	0.81	−0.34	−1.15
Val	−2.78	−35.83	0.98	0.52	−0.46
Phe	−2.59	−34.21	1.18	2.14	0.96
Ser	−3.10	−36.65	0.67	−0.30	−0.97
Pro	3.70	−25.16	7.47	11.19	3.73

times the C−C distance. We used the AM1[39] semiempirical molecular orbital method for the ONIOM medium level.

We have previously used this procedure in other studies of helical peptides.[4,17,18,31] All geometries were completely optimized in all internal degrees of freedom.

We tested all reasonable conformations and chose the one with the lowest energy where more than one reasonable conformation might exist for the side chain in the substituted peptides.

Results

One can conceive of the energy differences upon changing an Ala to another amino acid in a particular peptide in many different ways. We shall consider only two different ways in this paper. One is to consider the energy of this change in primary structure for various different secondary structural motifs such as extended β-strand or alpha helix. This can be accomplished by evaluating the energy of the hypothetical reaction

$$\text{acetyl(Ala)}_{17}\text{NH}_2 + \text{X} \rightarrow \text{acetyl(Ala)}_9\text{X(Ala)}_7\text{NH}_2 + \text{Ala} \quad (3)$$

where X is another amino acid for both the helical and an idealized extended β-strand of the same peptide. In this way, we can compute the energetic effect of the amino acid mutation for each structure. The difference between them gives the energetic preference for the helix or the strand compared to our standard $\text{acetyl(Ala)}_{17}\text{NH}_2$. Another method considers the hypothetical condensation reactions (similar to that mentioned above):

$$16\text{Ala} + \text{X} + \text{NH}_3 + \text{CH}_3\text{COOH} \rightarrow \text{acetyl(Ala)}_9\text{X(Ala)}_7\text{NH}_2 + 17\text{H}_2\text{O} \quad (4)$$

for the formations of both the helix and the idealized strand. Both allow us to determine the extent to which the substitution affects the idealized strand as well as the helix. In this manner, we can determine how much of the energy difference between the helix and the strand is due to the effect of amino acid substitution upon each structure. The results are tabulated in Table 1 and are illustrated in Figure 1 for the different amino acids substituted for the 10th Ala (counting from the acetyl, or N, end) of $\text{acetyl(Ala)}_{17}\text{NH}_2$. The data clearly indicate that the preference for the helix or strand form of a particular sequence over the polyalanine depends on the effects of the amino acid substitution upon both helical and extended strand forms. Substitution by proline provides the most dramatic example of this. While Ala → Pro mutation destabilizes the helix by 11.2 kcal/mol, the same substitution destabilizes the strand by 7.5. Thus, the strand is favored over the α-helix by only 3.7 kcal/mol, because the strand is less destabilized.

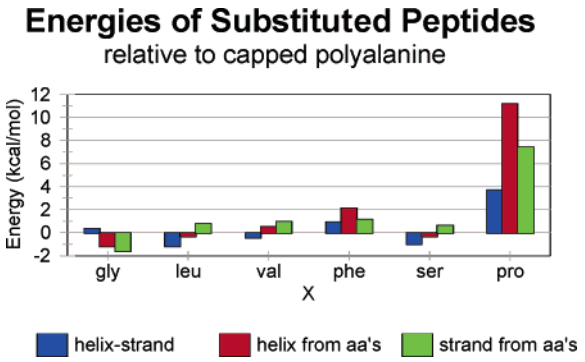


Figure 1. Comparison of the energetic effects of the mutation Ala → X for the 10th residue in $\text{acetyle(Ala)}_{17}\text{NH}_2$.

***β*-Strands.** The relative stabilities of the β-strands appear to depend on several different factors: (1) Steric distortions of the strand at the position of substitution; (2) the effect of the new amino acid upon the cooperative extended H-bonding chain involving C_5-interactions; and (3) H-bonding (including C−H···O) interactions. All three of these interactions will surely be mutually coupled to varying extents. We discuss each of the three in turn.

Steric Distortions. Steric distortions clearly will play an important role in determining the relative stabilities of the strands. The data in Figure 1 and Table 1 show that when X is any of the amino acid residues except Gly, the corresponding extended β-strand is destabilized relative to polyalanine. That only X = Gly provides a more stable β-strand seems clearly attributable to the reduced strain at the Gly position relative to any other X. Gly is the only amino acid whose chains are calculated to be planar rather than puckered.[40] The polyalanine strand is quite regular in geometry and quasi-planar with a slight pucker or "pleat" at each α-carbon. The angle, theta, between the first, eleventh, and penultimate (17th) nitrogens provides a convenient measure of this regularity. We chose the angle between these particular nitrogens since it should be 180° as all three are odd numbered on amino acids with the C═O's (nominally) pointing in the same direction (the position of the N alternates slightly for every second position). We define this angle as positive if the strand is bent toward the substituted side chain and negative if it is bent away. The 11th nitrogen is that closest to the position of substitution. In the optimized polyalanine, this angle is quite close to the theoretical value (179.8°). The (destabilized) Pro substituted strand (−137.2°) has the largest deviation from linearity (Figure 2) for this angle because of the rigidity of the ring, which separates the Ala's on either side into substrands that hardly interact with each other. In fact, the planes of their N's are twisted by about 30° from each other. However, each of these substrands is rather regular within itself (with angles analogous to that for the N's within the strand of about 177° in each substrand).

The Gly substituted strand (theta = 172.6°) also has some steric distortion from the polyalanine model, but the distortion

(39) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−9.
(40) Viswanathan, R.; Asensio, A.; Dannenberg, J. J. *J. Phys. Chem. A* **2004**, *108*, 9205−12.

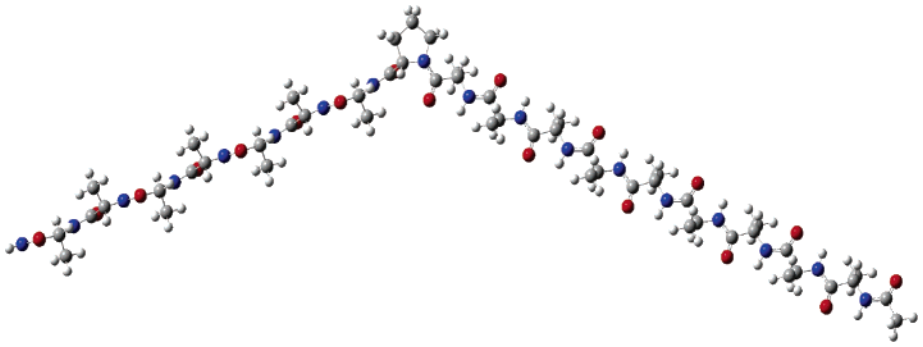


Figure 2. The extended β-strand when X = Pro.

is not caused by strain but rather from its relief. The geometric change is primarily due to the planarity of the strand at the Gly position compared to the slight pucker at the Ala's where the NCCN dihedrals are about 170°. As noted above, the Gly for Ala substitution relieves whatever strain gives rise to the pucker at the substituted position. β-sheets of glycine peptides can form planar (rather than pleated sheet) structures.

All the other amino acids considered as X have bulkier side chains than Ala. Thus, they impart more strain to the β-strand.

C_5 Hydrogen Bonds. In previous reports, we,[40] as well as others,[41] have illustrated the importance of C_5 H-bonding interactions in β-strands and β-sheets. These interactions modulate the cooperative interactions between the strands of a β-sheet, while the O···H distances in these interactions tend to shorten near the center of the strands. A small amount of energetic cooperativity accompanies this phenomenon, although it quickly reaches an asymptotic limit at about six amino acids (illustrated in Figure 10 of the reference).[4] Similar effects have been reported in long polyglycine strands.[41] The C_5 H-bonding O···H distances in acetyl$(Ala)_{17}NH_2$ are largest for that nearest the acetyl end (2.1365 Å), decrease near the center of the strand (2.0673−2.0717 Å for the 5th through 14th interactions), and increase to 2.1023 Å for that nearest the NH_2 end. The shortest distance is for the ninth of the 17 interactions. While the changes in this distance are rather small, the trend is quite evident. A comparison with the most destabilized strand, that with Pro in place of Ala, provides a very different picture. This strand contains one fewer C_5 interaction (because of the lack of an NH in proline). As a result, the C_5 H-bonding cooperativity of the polyalanine chains fragments into two shorter interactions (one of nine, the other of six C_5's). Each of these shorter chains characteristically has longer O···H distances at its ends. Thus, the ninth C_5 O···H (which was the shortest in the polyalanine) becomes 2.1605 Å for the Pro strand, as it terminates the chain of nine C_5's.

Hydrogen Bonds. We chose the (arbitrary) fully extended β-strand as a reference structure as it will not contain any of the H-bonds that stabilize the common secondary structural motifs such as sheets, helices, turns, and so forth. However, all H-bonding interactions cannot be completely eliminated. Even the polyalanine strand has the C_5 H-bonding interactions discussed above. Substitution of certain amino acids for alanine in the β-strand leads to other H-bonding interactions that cannot be avoided. In addition, H-bonds that might be present in the individual amino acids might become weakened or broken in

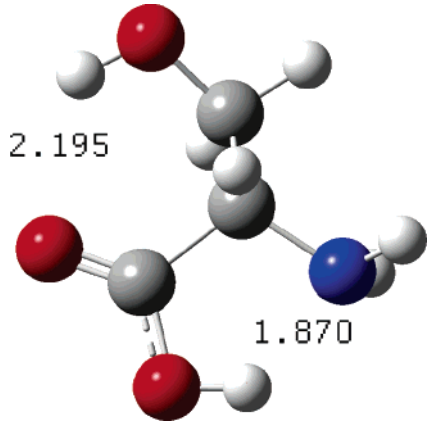


Figure 3. Optimized structure of serine with the O···H distances noted (in angstroms).

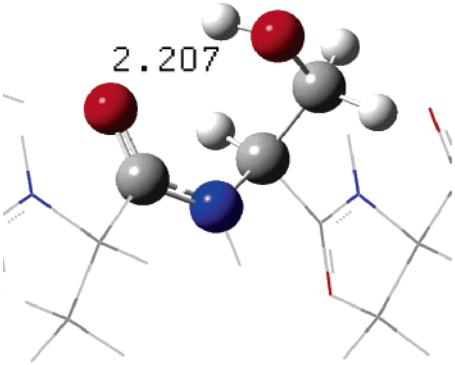


Figure 4. Detail of β-strand when X = Ser with O···H distance noted (in angstroms).

the peptide. Serine provides a good example of this phenomenon. The optimized structure of serine (Figure 3) contains an H-bond from the $-CH_2OH$ side chain to the C═O of the carboxyl group, as well as the more common H-bond between the OH of the carboxyl group to the NH_2. In the β-strand, this interaction is replaced by a different H-bond between the $-CH_2OH$ and the C═O of the adjacent alanine (Figure 4). As a consequence, the substitution of Ser for Ala in the β-strand is unfavorable by 0.67 kcal/mol. The strand distorts slightly from the quasi planarity of the polyalanine, as the 1−11−17 angle becomes −175.5°. Another slightly less stable (by 1.38 kcal/mol) conformation of the serine substituted β-strand has a different H-bond in which the $-CH_2OH$ side chain accepts an H-bond from the N−H on the other adjacent Ala.

Phe substituted β-strands contain C−H···O H-bonds from each of the ortho H's on the phenyl ring to the C═O's of each

(41) Horvath, V.; Varga, Z.; Kovacs, A. *J. Phys. Chem. A* **2004**, *108*, 6869−6873.

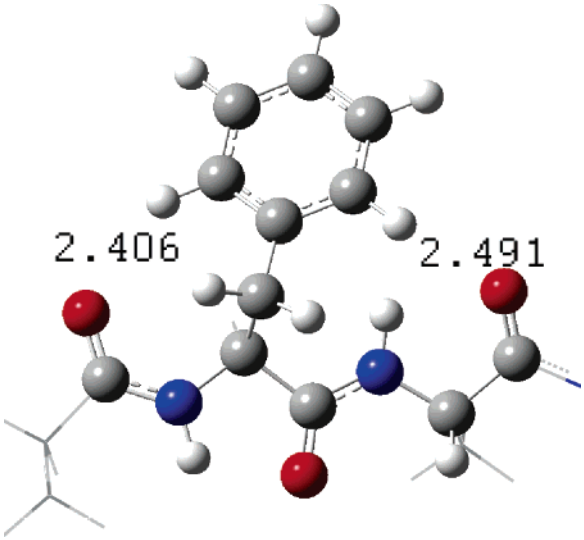


Figure 5. Detail of β-strand when X = Phe with O···H distance noted (in angstroms).

of the adjacent amino acid residues (see Figure 5). While these interactions must stabilize the structure, that stabilization is insufficient to overcome the added steric strain upon the Ala → Phe mutation.

α-Helices. Like the β-strands, the α-helices can be either stabilized or destabilized upon substitution of X for Ala. The ones stabilized are X = Gly, Leu, and Ser (the last two stabilizations are quite small), while those destabilized are Val, Met, Phe, and Pro. The stabilization for Gly can be attributed to relief of steric strain (as in the β-strand). The stabilization due to Leu can be attributed to C−H···O interactions, particularly between the β-C−H of the side chain to the C═O on the same residue. The O···H distances of the proximate helical H-bonds on either side of the side chain are slightly shorter, therefore presumably stronger, than for the corresponding H-bonds in the polyalanine helix. On the other hand, the stabilization due to Ser derives from a CH_2OH···O═C interaction from the Ser side chain to the C═O of the amino acid residue three positions distant (in the direction of the acetyl-capped N-terminus). Two conformations with other H-bonding interactions proved to be less stable. Interestingly, the two peptide H-bonds of the helix on either side of the Ser side chain are both longer than those of polyalanine.

The α-helices substituted with Val, Met, Phe, and Pro are all destabilized with respect to polyalanine. The destabilizations when X = Val and Met are quite small. The Val substitution destabilizes while the Leu substitution stabilized the helix. This small difference, which is reflected in the shorter CH···O distance for Leu (2.509 Å) than for Val (2.513 Å), is due to the lower strain in Leu, where the α-carbon of the side chain is less branched. The somewhat larger (2.14 kcal/mol) destabilization of the helix than the strand (1.18 kcal/mol) when X = Phe appears to be due to the steric strain which is compensated by only one C−H···O interaction in the helix (vs two in the strand).

Substitution by Pro provides by far the largest destabilization of the helix (11.19 kcal/mol). The helix with X = Pro contains one fewer amidic H-bond than the polyalanine, as Pro lacks the N−H necessary to a H-bond donor. The loss of the amidic H-bond is partially compensated by interactions between the

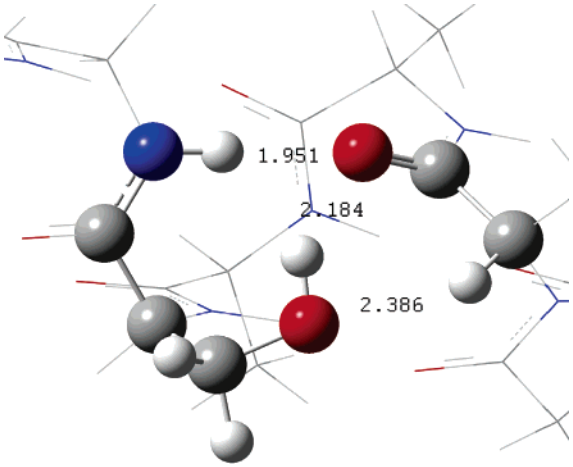


Figure 6. Detail of α-helix when X = Ser with O···H distance noted (in angstroms).

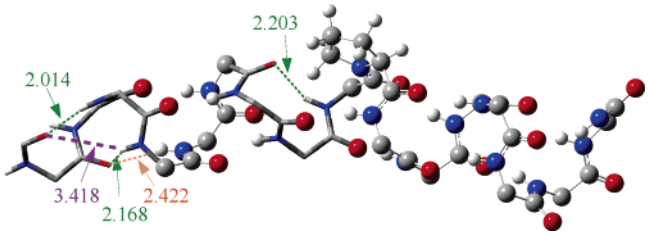


Figure 7. The Pro substituted helix with the 310-helical regions shown as tubes and the α-helical regions shown as balls and sticks. The 310 H-bonds are noted in green, the missing α-helical H-bond is noted in magenta, and the bifurcated α-helical H-bond is noted in orange. The methyl groups and Gly's attached to the α-carbons have been removed for better visualization.

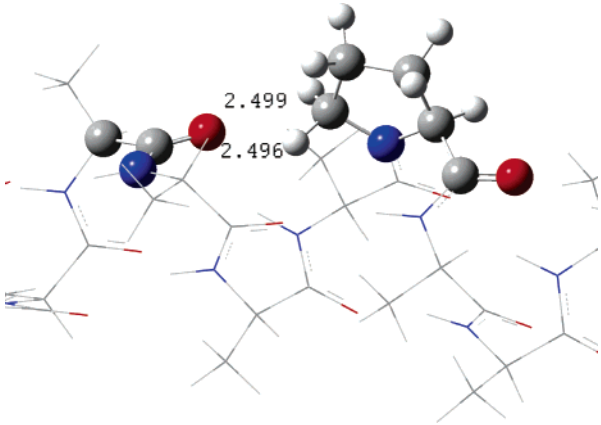


Figure 8. Detail of α-helix when X = Pro with O···H distance noted (in angstroms).

would-be H-bond acceptor and two C−H's adjacent to the N on the proline ring (see Figure 7). The missing amide H-bonding interaction occurs in a chain of five H-bonds. The disruption of the α-helix is further compensated for by formation of two 3_{10}-helical segments (see Figure 8), which are more stable for short helices. The first α-helical H-bond is broken (O···H distance = 3.418 Å) as a consequence. Increased strain in the helix, partially mitigated by the C−H···O interactions, presumably accounts for the additional destabilization. The destabilization is close to the value of 10.42 kcal/mol calculated to be the strength of the central H-bond in a chain of six formamides (containing five H-bonds).[42] This value is much larger than that

Table 2. O...H Distances (Angstroms) and Differences from the Corresponding Values for Polyalanine for Each Substituted α-Helix[a]

	Ala	Gly	ΔAla	Leu	ΔAla	Val	ΔAla	Phe	ΔAla	Ser	ΔAla	Pro	ΔAla
3_{10}	2.125	2.123	***−0.002***	2.143	0.018	2.138	0.012	2.141	0.016	2.140	0.014	2.014	***−0.111***
						Chain1							
	2.379	2.371	−0.007	2.337	−0.041	2.316	−0.063	2.322	−0.057	2.312	−0.067	2.168	−0.211
	1.995	1.984	−0.011	1.978	−0.017	1.984	−0.011	1.984	−0.011	1.980	−0.015	1.989	−0.006
	1.960	1.921	−0.039	1.946	−0.014	1.944	−0.016	1.929	−0.031	1.967	0.007		
	1.950	1.966	0.016	1.956	0.005	1.977	0.027	1.965	0.014	1.966	0.015	2.038	0.087
	2.037	2.061	0.024	2.047	0.010	2.042	0.005	2.045	0.008	2.060	0.023	2.063	0.026
average	2.064	2.061	−0.003	2.053	−0.012	2.052	−0.012	2.049	−0.015	2.057	−0.007	2.026	0.036
						Chain 2							
	2.010	2.020	0.010	2.014	0.004	2.004	−0.006	2.008	−0.003	2.007	−0.004	2.422	0.412
	1.983	1.983	0.000	1.979	−0.004	1.984	0.001	1.974	−0.009	1.981	−0.002	2.053	0.070
	1.968	1.952	−0.015	1.952	−0.015	1.994	0.026	1.962	−0.006	1.951	−0.017	2.069	0.101
	1.980	1.952	−0.028	1.978	−0.001	1.971	−0.008	1.987	0.007	1.989	0.010	1.961	−0.019
	2.031	2.046	0.015	2.021	−0.010	2.047	0.016	2.042	0.011	2.023	−0.008	2.068	0.037
average	1.994	1.991	−0.004	1.989	−0.005	2.000	0.006	1.994	0.000	1.990	−0.004	2.115	0.120
						Chain 3							
	2.024	2.029	0.005	2.032	0.007	2.035	0.010	2.034	0.010	2.033	0.008	2.009	−0.015
	1.955	1.957	0.002	1.956	0.001	1.947	−0.008	1.941	−0.014	1.994	0.039	2.203	0.248
	1.957	1.971	0.013	1.973	0.016	1.954	−0.003	1.971	0.014	1.961	0.004	1.926	−0.031
	1.973	1.970	−0.003	1.979	0.006	1.971	−0.001	1.966	−0.006	1.978	0.005	2.043	0.070
	2.194	2.191	−0.003	2.159	−0.034	2.191	−0.002	2.193	0.000	2.159	−0.035	2.194	0.001
average	2.021	2.023	0.003	2.020	−0.001	2.020	−0.001	2.021	0.001	2.025	0.004	2.075	0.054
						Overall Average							
	2.033	2.031	−0.001	2.028	−0.004	2.031	−0.001	2.029	−0.004	2.031	−0.001	2.081	0.044

[a] Values in bold italics are for 3_{10} H-bonds.

normally expected for an α-helical H-bond as it does not include the helical strain which must be taken into account when comparing helices with open structures.[4] In this case, we compare two helices (X = Pro with X = Ala). Here, only the differential strain, which should be quite small, matters, so the full energy of the H-bond should prevail. Guo reported that C−H···O interactions involving Pro can be either stabilizing or not stabilizing depending upon the environment.[23]

Table 2 presents the O···H distances for the various α-helices grouped by H-bonding amide chains. The substituted amino acid acts as a H-bond donor in chain 1and as a H-bond acceptor in chain 2. The data show that chains 1 and 2 are more affected than chain 3 and that the middle H-bonds of the first two chains are most affected. With the exception of Ala → Pro, the fact that the average O···H distance for all of the substituted α-helices are shorter than for the polyalanine suggests that the H-bonds are strengthened (or at least not weakened) by these substitutions. This observation reinforces the suggestion that whatever destabilization of the helix occurs must be due to additional strain upon substitution of X for Ala. The Ala → Gly substitution, which stabilizes the α-helix (although it stabilizes the *β*-strand even more), has virtually the same average O···H distance as polyalanine, suggesting its stabilizing influence derives from reducing strain. Comparison of the Ala → Leu with Ala → Val mutations suggests that Leu stabilizes while Val destabilizes the α-helix because of the additional steric strain induced by the additional methyl group on the first carbon of the side chain.

Comparison of α-Helices with *β*-Strands. When considering the effect of an amino acid mutation on the stability of a peptide secondary structural motif, such as an α-helix, one must always consider at least one other structure for comparison. The present data clearly show that when a mutation occurs that destabilizes (or stabilizes) the helix, the possibility that it better destabilizes another structure remains quite real. When comparing calculations with experimental results on helical propensities or other measures of helix stability, one should consider the effect of a primary structural modification not only upon the helix but also upon whatever nonhelical forms might occur. We have shown that the Ala → Pro mutations significantly destabilizes the extended *β*-strand. Clearly, the alternative structure to the α-helix in solution will not be 100% *β*-strand. The mixture of structures is usually referred to as random coil. However, the composition of this mixture for the mutated peptide must certainly be significantly different from that for the analogous polyalanine. The introduction of a Pro will certainly restrict the number of accessible conformations, probably excluding some that might allow some stabilizing H-bonds. As it has one fewer H-bond donor, it cannot form as many of these interactions, either internally or with solvent as can the polyalanine.

Comparison with Experimental Reports. Two significant and related problems impede the comparison of the gas-phase calculations with experimental reports: (1) The populations and energies of the nonhelical peptides are not known and (2) the relative effects of solvation cannot be estimated reliably experimentally and have not yet been treated theoretically. We consider each of these problems in turn.

Most experimental determinations of the effect of an amino acid mutation upon α-helix stability do so with respect to a random coil. Generally, the extent of helix formation is measured by monitoring either the circular dichroism (CD) or the NMR spectra and applying a statistical helix-coil theory, such as that formulated by Lifson and Roig.[43] Theories such as these do not account for the effects of the amino acid side chains upon the accessible conformations of either the helix or the coil.[44] Since the conformation of the helix is established, this ap-

(42) Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. A* **2003**, *107*, 10389−95.

(43) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974.
(44) Qian, H.; Schellman, J. A. *J. Phys. Chem.* **1992**, *96*, 3987−3994.

proximation poses a potential problem only for the coil. By combining helix-coil theory with experimental measurements, one can obtain thermodynamic parameters for the helix relative to the random coil. When this is done as a function of amino acid mutation, one can obtain the difference in these parameters as a function of the mutation. Thus, experiments of this type show the substitution Ala → Pro (for example) to destabilize the α-helix relative to the coil. However, these experiments cannot determine if the change in the helix/coil energy is due solely to a perturbation of the energy of the helix, the coil, or both. Nevertheless, one usually attributes the energy difference upon mutation to effects upon the helical structure. This seems natural, as the helix−coil theories (which do not take the amino acid side chains into account) neglect the effect of the amino acid mutation upon the structures that contribute to the random coil. To evaluate the effect of an amino acid mutation upon the energies of both the random coil (as experimentally determined) and the helix, one needs at least one other reference state. In this work, we have defined two such states: the component amino acids and the extended β-strand. The first is easily definable, as the heats of formation and free energies of these species in their standard states are readily available. In principle, the same thermodynamic parameters of the peptides could be obtained by measuring their heats of combustion. However, such experimental measurements pose clear practical problems. The second, the extended β-strand, is completely artificial. Nevertheless, it might make a significant contribution to the ensemble of structures that comprise random coil (as experimentally determined). To the extent that the amino acid mutation under consideration does not affect the number of hydrogen bonds (both intramolecular and with solvent) that can be formed, the β-strand probably could be used as a semiquantitative model for the determination of the effects of amino acid mutation upon the energy of the random coil in the absence of solvent. Thus, the energetic effect of the Pro → Ala mutation upon the random coil could reasonably be estimated by the unfavorable effect of 7.47 kcal/mol for the β-strand.

Natural proteins do not generally exist in random coils or in artificially conceived structures such as extended β-strands. They often contain secondary structural domains such as α-helices, β-sheets, and so forth. Thus, the relevant comparisons for the energetic effects of amino acid mutations might be between α-helices and β-sheets or aggregates such as those implicated in several (amyloid) diseases. These, of course, can be derived from comparisons of each relevant secondary structure to the same arbitrary, well-defined standard (e.g., the extended β-strand or the component amino acids).

As mentioned above, the randomness of the random coil has been questioned.[15,16] Recent reports suggest that the open structure might resemble the polyproline II structure in aqueous solution.[1] This structure contains no C_5 or other intramolecular H-bonding interactions. The C═O's and N−H's are completely exposed for H-bonding to solvent. Since this polyproline II structure has none of the stabilizing intramolecular H-bonds, assumes a somewhat strained backbone conformation, and has its H-bonding donors and acceptors exposed to solvent, one might reasonably conclude that its conformation is dictated by solvation. Thus, it would not be a minimum in the gas phase. One might suspect that the Ala → Pro mutation might not destabilize this conformation so much as the extended or β-strand model. However, the lack of a H-bond donor N−H on the proline would lower the hydration energy for the mutation.

Solvation clearly has an important effect upon the helix/coil equilibrium. For example, α-helices are relatively stabilized compared to the coil in TFE (generally 40%) compared to aqueous solution. The solvent effect upon the equilibrium can be primarily due to solvation differences of the helix, the random coil, or both. As in the effect of amino acid mutations, the effect appears to be most often attributed to solvation of the helix. This is natural as the helix has a known structure, so its solvation seems more easily interpretable, while the coil was assumed to be a random mixture. However, since the coil contains a mixture of conformations, it is the logical "structure" to be most affected by solvent. For example, a particular solvent might most favorably interact with one conformation much more advantageously than the others, which would significantly lower its energy. The apparent effect on the experimental equilibrium would be to destabilize the helix. One might expect that water stabilizes some of the nonhelical structures rather than destabilize the helix. The recent reports that short alanine-based peptides form the polyproline II structure in aqueous solution (which has all its H-bond donors and acceptors exposed to solvent)[1] reinforce this suggestion.

Most available theoretical solvation models were designed for small spherical ions or molecules that are ill-suited to large rodlike structures such as α-helices. The calculation of the solvation energies of the nonhelical state becomes daunting upon consideration of the quantity of conformations that likely have different solvation energies.

With the qualifications noted above, our results agree with some but not all of those of Baldwin, who reported that virtually all amino acid mutations that replace Ala in a poly-Ala peptide tend to disfavor helices. He found Pro and Gly to destabilize helices in both water and 40% TFE.[10] The interactions that cause the other mutations (X = Leu, Val, Ser) that we report to favor the helix over the strand all involve intramolecular H-bonds which could be replaced by H-bonds to solvent under Baldwin's experimental conditions. We suggest that the observed solvent effects probably derive from selective solvation of one or more (i.e., polyproline II) conformations of the unfolded state rather that of the helix.

Conclusions

Mutations of another amino acid (Gly, Leu, Val, Phe, Ser, Pro) for alanine near the center of a capped polyalanine can either stabilize or destabilize the α-helix; it almost always (except when X = Gly) destabilizes the β-strand. Differences in the effect of the same mutation upon the two structures give rise to the relative energetic differences between the α-helical and open structures which are somewhat related to the effects measured in solution. These relative energetic effects can be due to a difference in the stabilizations of both (Gly), a difference in the destabilizations of both (Val, Phe, Pro), or a stabilization of one and destabilization of the other (Leu, Ser). Thus, the effects of these mutations on both the helices and the group of conformations that comprise the random coil must both be considered for a proper understanding of their effects on peptide secondary structures in the absences of solvent. In light

of these results, the effects of solvation must be considered upon both the helical and population of random coil structures. Because the former is a relatively defined and unique structure while the latter is a complex mixture of flexible structures, the effect of the environment (solvation, temperature, etc.) should be greater upon the latter. The experimental observations should be interpreted in this context.

Pro has by far the largest effect upon both the β-strand and α-helix. The relative destabilization of the α-helix derives from the difference in these two large individual destabilizations.

Acknowledgment. This work was supported in part by grants from the donors of the Petroleum Research Foundation administered by the American Chemical Society and from PSC-CUNY.

Supporting Information Available: Cartesian coordinates of the relevant structures and complete refs 35 and 36. This material is available free of charge via the Internet at http://pubs.acs.org.

JA052689J