A Synthetic Retrotransition (Backward Reading) Sequence of the Right-Handed Three-Helix Bundle Domain (10-53) of Protein A Shows Similarity in Confomation as Predicted by Computation

Krista Witte, Jeff Skolnick, and Chi-Huey Wong*

Contribution from the Department of Chemistry and The Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received June 24, 1998

Abstract: A recent computational analysis of retro-proteins (backward reading of native proteins) suggests that the retro-protein has a tendency to adopt a structure similar to that of the natural one, as demonstrated by a case study using a truncated version of the B-domain sequence (10-53) of protein A which forms a right-handed, three-helix bundle (Olszewski, K. A.; Kolinski, A.; Skolnick, J. *Protein Eng.* **1996**, *9*, 5−14). To test this hypothesis, both the natural 44 amino acid peptide and its retro-sequence have been synthesized by solid phase and purified to homogeneity. Circular dichroism studies indicate that both peptides adopt right-handed α-helical structures in the presence of trifluoroethanol. Though it is not clear if this tendency is general, this work does provide useful information for the study of protein folding.

Introduction

It is widely believed that the three-dimensional structure of a protein is largely determined by its amino acid sequence and that this tertiary structure, in turn, often determines the protein's biological function. It has therefore been a long-standing goal to be able to predict the three-dimensional structure of a protein from the amino acid sequence. If these rules of folding could be understood, a protein's structure, and possibly its function, could be determined by the sequence of its corresponding mRNA. Despite intensive research in this area, these rules still remain largely a mystery. [2−9] Prediction of the structure through computational modeling has become one method by which to tackle this problem. [10−13]

Recently Olszewski et al. published a computational treatment of a unique peptide called retro-protein A, **1** (Figure 1).[1] This peptide is the "retrotransition" (backward reading) sequence of a fragment from a naturally occurring domain of protein A, a cell wall component of *Staphylococcus aureus*. The B-domain from which the fragment originated (**2**) is one of five homologous domains in protein A and is responsible for the binding of the Fc portion of human polyclonal immunoglobulin G.[14] The full length of the B-domain consists of a 59 amino acid sequence which forms a three-helix bundle as determined by both NMR[15] and X-ray crystallography.[16] Of the three helices, helix I (Gln10 to His19) is tilted with respect to helices II (Glu25 to Asp37) and III (Ser42 to Ala55) which are antiparallel. All three helices are right-handed. Having such an extensive knowledge of the natural protein but no structural information of the retro-protein provided a unique opportunity for computational analysis followed by experiment validation.

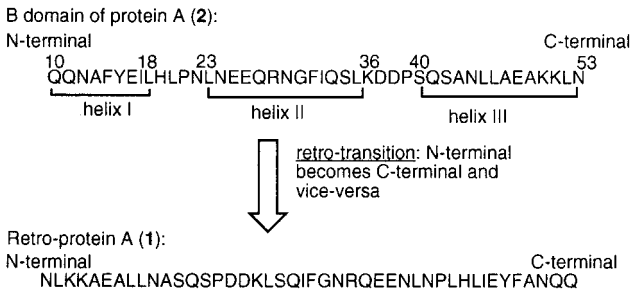

Figure 1. Sequences of B-domain and retro-protein A.

The aforementioned computational study was performed using a fragment of the B-domain (residues 10-53) as a base set for the retro-protein (**2**). A truncated form was used, as a survey of the literature showed both the solution[15] and the crystal structure[16] demonstrated a high degree of structural heterogeneity in the N- and C-termini of the B-domain. In the crystal structure, residues 1-5 and 49-59 were unable to be resolved. In the solution structure, residues 1-9 and 56-59 showed little determinable structure. Using the lattice model of proteins,

(1) Olszewski, K. A.; Kolinski, A.; Skolnick, J. *Protein Eng.* **1996**, *9*, 5−14.
(2) Vasquez, M.: Nemethy, G.; Scherage, H. A. *Chem. Rev.* **1994**, *94*, 2183.
(3) Pain, R. H. *Mechanisms of Protein Folding*; Pain, R. H., Ed.; IRL Press: New York, 1994; p 265.
(4) Creighton, T. E. *Protein Folding*; Creighton, T. E., Ed.; W. H. Freeman and Co.: New York, 1992; p 547.
(5) Langone, J. J. *Molecular Design and Modeling: Concepts and Applications*; Academic Press: San Diego, CA, 1991; Vols. 202−203.
(6) Schneider, J. P.; Lombardi, A.; DeGrado, W. F. *Fold. Des.* **1998**, *3*, R29−R40.
(7) King, J.; Haase-Pettingell, C.; Robinson, A. S.; Speed, M.; Mitraki, A. *FASEB J.* **1996**, *10*, 57−66.
(8) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935−941.
(9) Mitraki, A.; King, J. *FEBS Lett.* **1992**, *307*, 20−25.
(10) Wright, P. E.; Dyson, H. J.; Lerner, R. A. *Biochemistry* **1988**, *27*, 7167−7175.
(11) Dandekar, T.; Konig, R. *Biochim. Biophys. Acta* **1997**, *14*, 1−15.
(12) Kolinski, A.; Skolnick, J. *Acta Biochim. Pol.* **1997**, *44*, 389−422.
(13) Jones, D. T. *Curr. Opin. Struct. Biol.* **1997**, *7*, 377−387.
(14) Friesner, R. A.; Gunn, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 315−342.
(15) Langone, J. J., Dixon, F. J., Kunkel, H. G., Eds.; Academic Press: New York, 1982; Vol. 32, pp 156−252.
(16) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arate, Y.; Schimada, I. *Biochemistry* **1992**, *31*, 9665−9672.

10.1021/ja982203h CCC: $15.00
Published on Web 12/03/1998

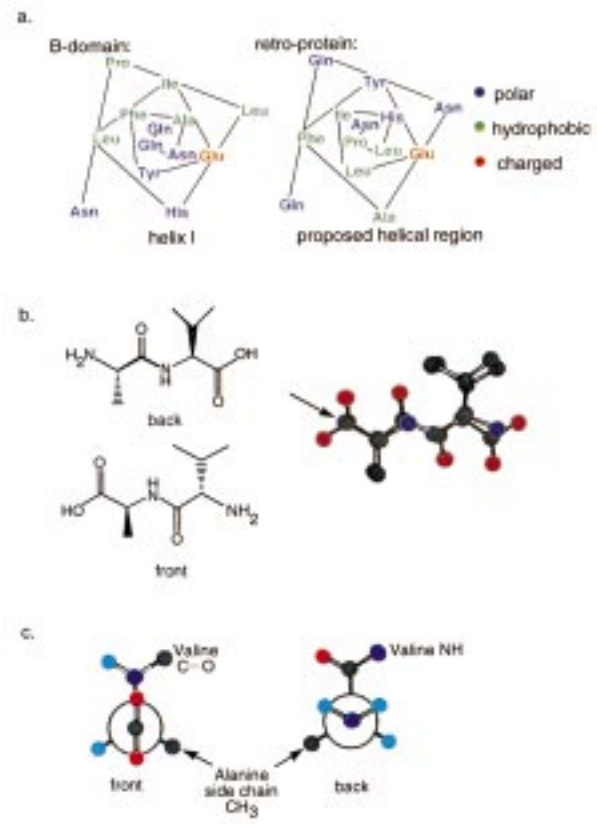

Figure 2. (a) Helical wheel analysis of the hydrophobicity of residues. For b and c, atoms are colored by type as follows: carbons, gray; oxygens, red; nitrogens, blue; hydrogens, cyan. (b) Antiparallel overlay of forward and retro sequences. (c) Side chain direction of the antiparallel peptides as viewed down the bond indicated by the arrow in b.

Olszewski et al. modeled the structure of retro-protein A. It was predicted to have a tendency to adopt a structure similar to that of the natural B domain, i.e., a right-handed, three-helix bundle, but is predicted to be less stable then the parent structure.[1] This is in contradiction to an earlier publication of Guptasarma which hypothesized that a retro-protein might adopt the mirror image structure of the native protein, in this case a left-handed α-helix.[17] Serrano and co-workers[18] have synthesized the retro-sequence of protein A and find that it is unfolded. Here we explore if it can be made more stable by suitable addition of trifluoroethanol, which is know to stabilize α-helical structures.[19]

Results and Discussion

The study of the retro-protein provides a unique opportunity to study the importance of context and hydrophobic effects as driving forces in secondary structure formation. When the sequences of the retro-protein and of the B-domain are analyzed by alignment to a helical wheel, it can be seen the pattern of hydrophobic and hydrophilic residues are similar (Figure 2a). If the primary driving force for the formation of secondary structure is to group and protect the hydrophobic residues from solvent, one could expect a similar structure from the two peptides. Each residue is also bordered by the same two residues in both the retro-protein and the B-domain, albeit the positions of the neighboring residues have switched. An important

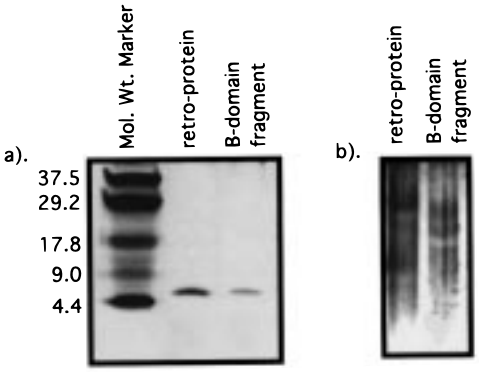

Figure 3. (a) Denaturing SDS tricine−glycerol acrylamide gel of the retro-protein and the corresponding B-domain fragement. (b) Native tricine−glycerol acrylamide gel of the retroprotein and the B-domain fragment.

difference between the retro-protein and the B-domain can be seen if one overlays the residues of the two peptides in an antiparallel fashion. When overlaid, identical residues are matched from both peptides (Figure 2b); however, a vital difference is immediately apparent. The direction that the side chains protrude from the backbone is opposite (Figure 2c). It is this important difference and its affect on the secondary structure of the peptide which the previously published computational study and this experimental work set out to address.

Both the 44 amino acid retro-protein and the 44 amino acid peptide were synthesized by solid-phase peptide synthesis. Both were characterized by mass spectrometry and purified by HPLC to greater than 98% purity. A denaturing polyacrylamide gel shows identical mobility of the two peptides (Figure 3a). However a native gel shows only a smear of bands for both the B-domain as well as the retro-protein indicating a lack of defined secondary structure under the running conditions (Figure 3b).

Circular Dichroism Studies. Circular dichroism (CD) is a common method used to determine the secondary structural elements of a protein.[20] Random coil, β-sheet, and α-helix all have very distinct CD spectra. Random coil is distinguished by an intense negative band at approximately 200 nm. The α-helix conformation has an intense negative band at 222 nm and another at 208 nm. Therefore it is fairly easy to distinguish each of these conformations. The standard unit used to report CD data is the mean residue ellipticity (θ_{MRW}) which is essentially the amount of optical activity which can be attributed to each residue on average. Because it takes into account the concentration, the molecular weight, and the number of residues in the protein being studied, it provides a standard measurement of secondary structure which can be compared with literature values regardless of the experimental details.

The predicted structure for retro-protein A was a propensity for α−helical structure. However, in agreement with the work of Serrano, initial studies in a 10 mM borate buffer indicated random coil. These studies were conducted at various temperatures and at pH 7 (Figure 4a). There was no increase in α-helical content as monitored by the intensity of the band at 222 nm and the shape of the trace between 180 and 260 nm upon either cooling to 5 °C or at room temperature (21 °C). It was thought that perhaps upon purification, retro-protein A was precipitated in a random coil state and could not renature to the α-helical conformation. To try to overcome this difficulty, the protein was denatured using 7 M urea and refolded. Two different methods of refolding were tested (Figure 4b). A fast

(17) Guptasarma, P. *FEBS Lett.* **1992**, *310*, 205−210.
(18) LaCroix, E.; Viguera, A. R.; Serrano, L. *Fold Des.* **1998**, *3*, 79−81.
(19) Liudens, M. K.; Figge, J.; Breese, K.; Vajda, S. *Biopolymers* **1996**, *39*.
(20) For a review of the technique, see: Schmid, F. X. *Protein Structure; a practical approach*; IRL Press: Oxford, UK, 1990.

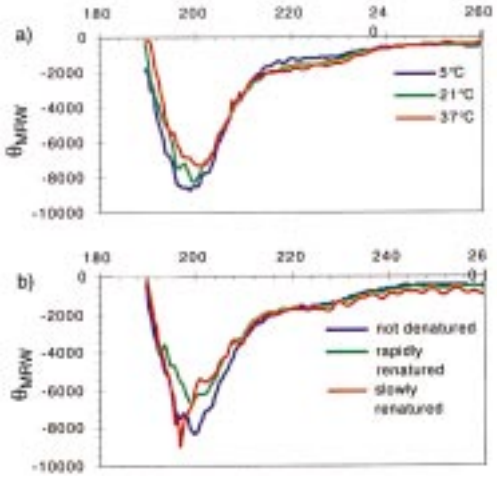

Figure 4. (a) Effect of temperature on conformation of the retro-protein A. (b) Effect of unfolding and refolding on the conformation of retro-protein A at 21 °C. All CD studies were done at pH 7.0.

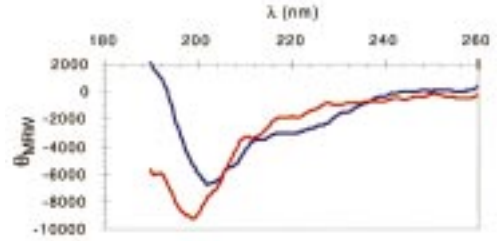

Figure 5. Comparison of the CD spectra of retro-protein A (red) and the B-domain fragment (blue) with 0% TFE.

method in which the urea was removed by a desalting column provided a CD spectrum which appeared no different than above. A slower method was then used in hopes that slow refolding would provide more helicity. The denatured protein was dialyzed against buffer to slowly remove the urea. However, this also did not increase helicity over what was seen before denaturing.

The B-domain fragment **2** was used as a control in these studies. CD data were taken at identical conditions as those described in Figure 4. Surprisingly, the results were similar to those found with the retro-protein (Figure 5). Very little α-helix was evident, and the conformation appeared to be primarily random coil. As mentioned earlier, the structure of the complete B-domain of protein A has been extensively studied and has been determined to be predominantly α-helical. However peptide **2** is not the complete B-domain but a fragment lacking the first 10 and last 6 residues. It is possible that the residues which were left out of **2** are important for the initiation or stabilization of the helical structure.

CD Studies in the Presence of Trifluoroethanol. Because both the B-domain fragment as well as retro-protein A were predicted to have a tendency toward α-helicity, we reasoned that if the α-helical state could be stabilized, even slightly, the protein might fold into the predicted conformation. Trifluoroethanol (TFE) is known to stabilize α-helical structures by stabilizing the essential internal hydrogen-bond structure.[19]

A TFE titration of the retro-protein was performed at both pH 5 and 7. At pH 5 and the addition of 10% TFE, the band at 200 nm attributed to random coil was no longer visible. An increase in the intensity of the band at 222 nm continued with increased TFE concentration. At pH 7 and 10% TFE, the band at 200 nm was still fairly visible; however, at pH 7 and 20%

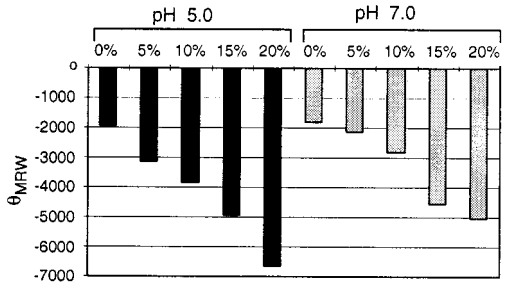

Figure 6. Comparison of the crucial qMRW at 222 nm at pH 5 and 7 along a TFE titration. TFE concentrations are as indicated.

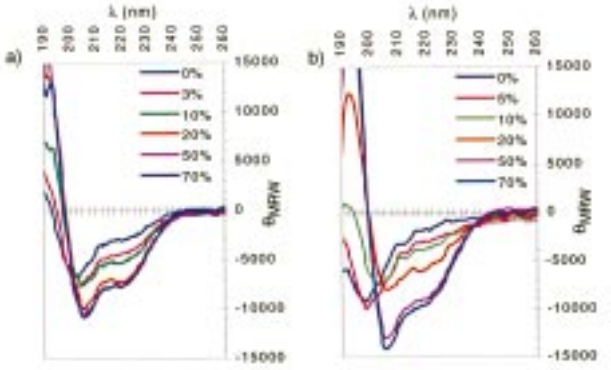

Figure 7. TFE titration of the (a) B-domain fragment and (b) retro-protein A at pH 5. The concentrations of the TFE cosolvent are as indicated in the legends.

TFE, α-helix dominates the spectrum. A comparison of $[\Theta]_{MRW}$ at 222 nm at pH 5 and 7 shows a higher degree of helicity at the more acidic pH at any given concentration of TFE (Figure 6). This indicates the helix was more stable at this lower pH. In addition, it should be noted that the helicity is right-handed as indicated by the negative value for the band at 222 nm, also in agreement with the predicted structure.

An extensive TFE titration study was performed using TFE concentrations ranging from 0 to 70% on both the retro-protein **1** and the B-domain fragment **2** (data shown in Figure 7). In both cases a marked increase in α-helicity is seen as the TFE concentration increases, as evidenced by the appearance of the critical negative peak at 222 nm. As can be seen from Figure 7a the B-domain fragment shows very little α-helical character without TFE as a cosolvent. From 3 to 10% TFE, the characteristic peak shape at 222 nm began to appear, indicating an α-helical conformation was emerging. When the TFE concentration reached the 20% range, a pronounced negative peak at 222 nm became apparent. A further increase of TFE from 20 to 50% and 70% did little to increase the α-helical content of peptide **2** with a maximum θ_{MRW} at 222 nm of 7592.

Similarly, the retro-protein **1** showed very little α-helical character at low concentrations of TFE (Figure 7b). However, in contrast to peptide **2** the retro-protein **1** showed a dramatic increase in θ_{MRW} at 222 nm when the TFE concentration was increased from 20 to 50%. A further increase of TFE to 70% shows very little affect, giving a maximum θ_{MRW} of 9292. A comparison of the θ_{MRW} value at 222 nm indicates that the retro-protein has a slightly greater α-helical content than the control B-domain fragment **2** at TFE concentrations greater than 50% and slightly lower α-helical content at TFE concentration lower than 20% (Figure 8). Aside from these small differences the behavior of these two peptides consisting of identical residues in opposite order was remarkably similar.

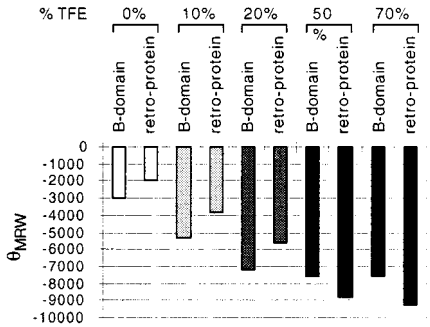

Figure 8. Comparison of the mean residue ellipticity (qMRW) at 222 nm for the B-domain fragment (control) and retro-protein A.

Conclusion

Although retro-protein A (**1**) did not initially form the predicted α-helical structure, neither did the forward version of the peptide, the B-domain fragment (**2**). Using trifluoroethanol, an α-helix stabilizer, as a cosolvent both the retro-peptide and the control peptide showed α-helical character. A TFE titration of both peptides provided a TFE concentration at which the mean residue ellipticity (θ_{MRW}) no longer increased with increasing TFE concentration. For the B-domain fragment this concentration of TFE was 20% whereas for the retro-peptide it was 50%. However at concentrations of TFE greater than or equal to 50%, the retro-peptide demonstrated more α-helical character than the control peptide. In addition, like the natural B-domain, the helical structure of both peptides was right-handed as predicted. This was contradictory to earlier suggestions. The helicity of retro-protein A was more pronounced at a slightly acidic pH than at neutral pH, and a similar trend was reported with the native B-domain. It is noted, however, that CD is a rather crude criterion (compared to NMR) to determine conformation and that the helical structure was not observed until TFE was added to the forward- or the retro-sequence. Work is in progress to study other sequences to test the generality of this hypothesis in the absence of TFE.

Experimental Section

General. The peptides corresponding to the sequence of retro-protein A (**1**) and the B-domain fragment (**2**) were synthesized by Quality Controlled Biochemicals. All other reagents and buffers were purchased from Aldrich and were of the highest quality available. Circular dichroism studies were performed on a AVIV Model 62DS CD spectrometer. SDS-PAGE and native PAGE were run on Biorad Protean II system using a 20% tricine gel with 8% glycerol added and stained using nonammonical silver staining.

Characterization of the Peptides. Mass spectrum was determined by electrospray mass spectroscopy. Retro-protein A (**1**): expected 5040, obsd 5040. B-domain fragment (**2**): expected 5040, obsd: 5040. HPLC was performed on a Hitachi L6200A equipped with a UV detector and with a C18 reverse phase analytical column. Gradient of H_2O and CH_3CN was used, both containing 0.1% trifluoroacetic acid.

Circular Dichroism. A typical procedure is as follows: A 50 μM stock of the peptide to be studied was made in the desired buffer (0.2 mg in 1 mL). A further dilution of 100 μL of the above stock in 2.4 mL of buffer was needed to make the 3.1 μM final solution. If trifluoroethanol was used it was added to the peptide stock and the buffer pH readjusted, and then the 100 μL of stock solution was added. CD spectra were taken at the temperatures stated. At least 30 min of equilibration time was needed to flush the chamber with nitrogen and reach the desired temperature. A minimum of 10 sweeps was taken at the step size of 0.5 nm. Data were then filtered two times through a standard second-order filter (for smoothing), and results were calculated as described in the text.

Denaturing and Refolding of Retro-protein A. Denaturing the retro-protein A was accomplished by adding 2 mg of the protein to 100 μL of 7 M urea. This solution was allowed to sit at room temperature for 0.5−2 h. For the slow refolding experiments, a 50 μL aliquot was dialyzed against 10 mM borate buffer (4 L for 4 h repeated 3 times) to remove the urea. The resulting solution was then diluted 1:10 with buffer, and 100 μL of this solution was added to 2.4 mL of buffer. For the quick refolding experiments, the samples were desalted using a small column of G-250 equilibrated with 10 mM borate buffer. The resulting protein solution was then treated as before.

JA982203H