REVIEW ARTICLES

Multiple conformations of proteins in native state*

SHENG Xiangrong** and PAN Xianming

(National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, 100101, China)

Received May 11, 2001; revised July 5, 2001

Abstract Examples of protein sequences that can adopt multiple native states are recently accumulated. Characterization of the protein multiple conformations will have important implications for our understanding of the relationship between structure and function, and their folding kinetics. In present review, the experimental evidence for the existence of multiple conformations in the native state of proteins, the molecular basis and the biological significance of multiple conformations of proteins are focused.

Keywords: protein folding, multiple native conformations, molecular basis, biological significance.

It has been accepted that amino acid sequence of a protein contains the information necessary to define its fold and that a globular protein in its native state adopts a single, well-defined conformation[1]. However, the polypeptide chain may fluctuate between preferred conformations with amplitudes and angles up to 50 Å and 20°, respectively. This observation suggests that the adage "one sequence, one conformation" is not strictly true. Recently, examples of sequences that can adopt multiple native states are emerging[2~10]. Characterization of the multiple conformations will have important implications for our understanding of the relationship between structure and function and folding kinetics of proteins.

1 Evidence for the existence of multiple conformations in the native state of proteins

The direct evidence for multiple conformations of proteins has come from nuclear magnetic resonance (NMR) studies in the case of staphylococcal nuclease. Dobson et al. used magnetization transfer (MT)-NMR to demonstrate the presence of at least two native forms of nuclease in equilibrium and to study their interconversion with the unfolded state under the conditions of the thermal unfolding transition[2]. The experiments revealed that two distinct native forms of the protein fold and unfold independently and that they could interconvert directly or via the unfolded state. Although the spectra of the different forms suggested that they were structurally similar, the MT experiments showed that the kinetics of folding and unfolding were quite different.

The multiple folded conformations of calbindin D_{9k} were characterized by two-dimensional (2D) ^1H NMR spectroscopy[5]. The 2D ^1H NMR approach is a unique method for studying multiple forms of proteins in solution because it provides the identity of the specific residue(s) in chemical equilibrium, as well as the information on the relative concentration, the rate of interconversion, and the extent of differences in the local and global conformations. The NMR spectra resolved major and minor resonance in a ratio of 3:1, and the exchanged process was slow on the NMR time scale with a value in the range of 10^{-1} s^{-1} ~ 10^{-2} s^{-1}.

Furthermore, the conformational heterogeneity in the native state of globular proteins has also been found in crystalline state. Two crystalline forms of adenylate kinase from pig muscle have been reported, which can be interconverted depending on the pH of the medium[11,12]. In the case of calbindin D_{9k}[5], the temperature factors of the second calcium-binding domain were higher than that of the first one, and very high values were observed in the loop that bridges the two domains (containing Pro-43). High temperature factors may be associated with conforma-

* Supported by the National Natural Science Foundation of China (Grant No. 30000029) and the Major State Basic Research Development Program of China (G199075606)

** To whom correspondence should be addressed. E-mail: xrsheng@sun5.ibp.ac.cn

tional heterogeneity.

However, the high-resolution detection of multiple forms of globular proteins by NMR and crystallography was limited by the need of proper size of protein samples and the availability of high quality crystals. There is literature reporting that the evidence for multiple folded forms of protein could be obtained from electrophoresis, chromatography and spectrometry. Russell et al.[13] reported that the higher *Mr* values of rabbit muscle adenylate kinase determined from gel filtration were obtained in the presence of dithiothreitol (DTT), and at higher pH and higher substrate concentrations. Furthermore, Zhang et al.[14~16] reported that the time course of 8-anilino-1-naphthalenesulfonic acid (ANS) binding to adenylate kinase was a biphasic process. The burst phase ended in the dead time of the stopped-flow apparatus (about 15 ms), whereas the slow phase completed in about 10 min. A kinetic approach was used to explore the mechanisms of the fluorescence building. The results clearly revealed the existence of at least two native forms of adenylate kinase in equilibrium in solution. One form (denoted by N_1) rapidly bound to ANS, whereas the other (denoted by N_2) did not. N_1 binding with ANS caused the burst phase of fluorescence. N_2 was slowly converted to N_1 and then bound to ANS, resulting in the slow phase of fluorescence. The result that the fraction of the burst phase decreased and that of the slow phase increased with increasing ANS concentration excluded the possibility of the conformational changes during adenylate kinase binding with ANS. The unfolding processes of the tertiary and secondary structures of adenylate kinase were both biphasic. Meanwhile, the protein-ANS complex showed a monophasic unfolding process. The results could not be interpreted as the formation of an unfolding intermediate because the amplitude of the burst phase remained constant during the adenylate kinase unfolding process. The data demonstrated that the ratio of two forms of adenylate kinase was 0.7 : 0.3 and the changes of the pH in the solution or the presence of substrates would shift the equilibrium to fit the new conditions.

2 Molecular basis of multiple conformations

The molecular basis of multiple conformations is not unique. In most cases, proline isomerization will lead to the conformational heterogeneity of proteins. Proline residues play specific roles in determining protein structure; they act as N-terminal caps to α-helices, as helix termination signals, and as corner residues in β-turn sequences. Moreover, *cis* and *trans* conformations of Xaa-Pro bonds have comparable energies, leading to a 10% ~ 30% population of *cis* Xaa-Pro in the unfolded assemble and an about 6% frequency of *cis* Xaa-Pro bonds in native proteins[17~23]. Some studies have demonstrated that some proteins show isomerization about an Xaa-Pro peptide bond in the native state, which gives rise to structural heterogeneity in the native state[2,5,7,9,10]. According to the NMR studies of SNase, the structural heterogeneity can be ascribed to the isomerizations of the Lys116-Pro117[2,4,24~26] and His46-Pro47 peptide bonds[6]. For native wild-type SNase, the predominant configuration (about 90%) of the Lys116-Pro117 peptide bond is assigned to *cis*, and the minor configuration (about 10%) is assigned to *trans*. For the His46-Pro47 peptide bond, the major configuration (about 80%) is assigned to *trans* whereas the minor configuration (about 20%) is assigned to *cis*. The heterogeneity in the native state was eliminated by the substitution of Gly for Pro117[27,28]. The isomerizations about these two peptide bonds of SNase occur independently and produce the four different native conformers[29].

In addition, the rate constant of the slow phase reaction of the ANS binding to rabbit muscle adenylate kinase was accelerated about 2.4-fold in the presence of 2.0 μmol/L peptidyl prolyl *cis*/*trans*-isomerase, suggesting that the domain movement during ANS binding to adenylate kinase involves proline isomerization. The activation energy of the slow phase was determined to be 74.6 kJ/mol, which is comparable to the activation energy of proline *cis*/*trans*-isomerization (about 80 kJ/mol)[30].

Furthermore, X-ray analysis of different crystal forms of proteins and peptides occasionally showed the occurrence of isoforms with different disulfide bond chirality[31]. Recently, two conformational isomers were observed in the ^{1}H NMR spectra of the basic pancreatic trypsin inhibitor (BPTI) and of a mutant protein with Gly 36 replaced by Ser, BPTI (G36S). The two conformers were different in the chirality of the disulfide bridge formed by Cys 14 and Cys 38. In BPTI, the population of the minor conformer increased from about 1.5% at 4 ℃ to 8% at 68 ℃. In BPTI (G36S), the population of the minor conformation was about 15% of the total protein[32].

In earlier work, slow local conformational equilibrium in peptides and proteins was also shown to involve 180° flips of aromatic rings about the Cβ-Cγ

bond[33,34]. It was reported that multiple conformations did exist for the 7～64 peptide fragment of the C5 domain of human α3-chain type Ⅵ collagen in solution, and that the origin of the multiple conformations in the central β-sheet was a reorientation of the Trp 21 ring. The rate of this reorientation was $(1.01 \pm 0.05)\,s^{-1}$ at 303 K, as determined by 2D-NMR, and the population of the minor conformation was found to be $(6.4 \pm 0.2)\%$ at 303 K. It seems that all Kunitz-domain proteins with a tryptophan in position 21 will exist in a major and a minor conformation, slowly exchanging on the NMR time scale[35～37].

Flexible surface-loops in protein structure are thought to play an important role in substrate binding and molecular recognition. Crystal structures have shown that, for a number of enzymes and antibodies, substrate and/or antigen binding causes a conformational change of a surface-loop either from an "open" to a "closed" form to partially cover the bound substrate[38～41], or from a "partially closed" form to "open" form, resulting in an induced fit for the antigen or ligand binding[42,43]. However, some data disagree with a long-held assumption that the loop closure is ligand gated. The results have revealed that the loop motion is not ligand gated, but a natural motion of the protein, and the population ratio between loop open and closed forms is shifted by the presence of ligand. In the case of loop-6 in triosephosphate isomerase, the population ratio changes from 10 : 1 (without ligand bound) to 1 : 10 (with ligand bound)[44]. Similarly, the open and closed forms of wild type cytochrome *c* peroxidase (CCP) are in constant equilibrium in solution even in the absence of ligand observed directly with protein crystallography and characterization with solution kinetics[8]. The loop open conformer in the native form of CCP leaves an empty binding pocket to which ligand can bind. The binding of ligand selects this open conformer and shifts the loop close-open equilibrium toward the loop open conformer. The binding of ligand does not "cause" the loop to open; but, because the ligand binds to the open conformer, it selects this conformer and shifts the equilibrium to the open conformer. This binding only appears to be induced fit because of the equilibrium shift; before ligand addition, a small amount of the protein pre-exists in the open conformer as shown by the pre-equilibration kinetics.

3 Biological significance of multiple conformations

It is of great interest to elucidate the influence of multiple conformations on protein activity.

In the previous study, the catalysis and inhibition kinetics of adenylate kinase was re-examined on the basis of multiple conformations of the enzyme[45]. The results demonstrated that one isoform of the enzyme could bind to the substrates in the forward reaction such as MgATP and AMP, and the inhibitor P^1, P^5-bis (adenosine-5'-) pentaphosphate (Ap_5A), while the other isoform could bind to the substrates in the backward reaction, such as MgADP and ADP. As a consequence, Ap_5A acted as a competitive inhibitor for the forward reaction and a mixed noncompetitive inhibitor for the backward reaction, which strongly supported the suggestion that two native forms of adenylate kinase be involved in the catalytic reactions. Further experiments indicated that both conformations distinguished by the ANS probe of adenylate kinase were active[46].

The difference in the conformation of multiple conformers of protein may lead to the difference in their functional properties. Cardiotoxins (CTXs) are involved in the depolarization of the excitable membranes and are hypothesized to act on membrane proteins involved in transmission of Na^+ and Ca^{2+} currents[47]. They also induce cell lysis and stimulate the activity of tissue lipases[48]. The lytic activity of CTXs is associated with their binding to and damaging of cell membranes[49,50]. The interaction of the P-type cardiotoxin Ⅱ from *Naja oxiana* snake venom (CT Ⅱ) with perdeuterated dodecylphoscholine (DPC) was studied using ^{1}H-NMR spectroscopy and diffusion measurements[51]. The structure of micelle-bound CT Ⅱ was found to be similar to the "major" form of CT Ⅱ in aqueous solution, with the Val7-Pro8 peptide bond in the *trans*-configuration[52]. No "minor" form having the *cis*-configuration of the bond was found in the micelle-bound toxin[51]. What is the functional role of the "minor" form is unclear now.

Some evidence showed that the multiple conformations involving loop movement play an important role in the protein activity. Many observations suggest that the conformational changes in loops are needed for enzymatic catalysis[53～55]. Triosephosphate isomerase is one of the first proteins for which large conformational changes in a loop region be demonstrated. However, the loop motion of "open" to "closed" conversion is not ligand-gated but a natural motion of the protein[44]. The relative energies of

the two stable states of the loop do depend on the presence of the ligand, and the rate for the most evident motion between the "open" and "closed" states is closely matched to the catalytic turnover rate. It is reasonable to suggest that the population ratio for the two states of the loop and the energetic barrier for opening and closing have crucial effect on the catalytic process of triosephosphate isomerase, since triosephosphate isomerase is a reversible enzyme, and the reactants and products are very similar. However, to extend the conclusion to other enzymes needs more experimental measurements.

There were also examples showing that no significant relations are present between multiple conformations and protein functions. Although the loop 190 ~195 opening and closing were determined to be the source of multiple conformers of CCP and related to the binding of cytochrome *c*, the idea that the loop opening is involved in the electron transfer process was disfavoured. In fact, the loop opening kinetics (calculated half-life $\gg 10^2$ s) appeared to be too slow for the electron transfer. There were no reported structures of the ES form, indicating that the B-factor of loop 190~195 was increased. The structure of the CCP-cytochrome *c* complex did not show the B-factor of loop 190~195 being significantly altered[8].

Similar results were found in BPTI and in human α3-chain type Ⅵ collagen C-terminal domain (C5 domain). In the former, it was observed earlier that selective reduction of the Cys14-Cys38 disulfide bond resulted in virtually unaltered binding affinity to trypsin, the strictly local changes coupled with the disulfide flips made a significant influence on the functional properties rather unlikely[32]. In the latter, the C5 domain was high homologous to the BPTI. Examination of the crystal structure of the trypsin/BPTI complex suggested that a reorientation of the Trp21 ring had little or no effect on the inhibitory activity[35].

Acknowledgments The authors express their sincere thanks to Prof. Tsou, C. L. at the Institute of Biophysics, Chinese Academy of Sciences for many helpful suggestions.

References

1 Anfinsen, C. B. Principles that govern the folding of protein chains. Science, 1973, 181: 223.
2 Fox R. O. et al. Multiple conformations of a protein demonstrated by magnetization transfer NMRspectroscopy. Nature, 1986, 320: 192.
3 Evans, P. A. et al. Proline isomerism in staphylococcal nuclease characterized by NMR and site-directed mutagenesis. Nature, 1987, 329: 266.
4 Evans, P. A. et al. A magnetization-transfer nuclear resonance study of the folding of staphylococcal nuclease. Biochemistry, 1989, 28: 362.
5 Chazin, W. J. et al. Proline isomerism leads to multiple folded conformations of calbindin D_{9k}: Direct evidence from two-dimensional 1H NMR spectroscopy. Proc. Natl. Acad. Sci. USA, 1989, 86: 2195.
6 Loh, S. N. et al. Techniques in Protein Chemistry. San Diego: Academic Press, 1991, Vol. 2, 275~282.
7 Amodeo, P. et al. Multiple conformations and proline *cis-trans* isomerization in salmon calcitonin: A combined nuclear magnetic resonance distance geometry and molecular mechanics study. Biochemistry, 1994, 33: 10754.
8 Cao, Y. et al. Protein conformer selection by ligand binding observed with crystallography. Protein Science, 1998, 7: 72.
9 Shalongo, W. et al. A kinetic study of the folding of staphylococcal nuclease using size-exclusion chromatography. Biochemistry, 1992, 31: 11390.
10 Huisman, J. G. et al. Recognition properties of V3-specific antibodies to V3 loop peptides derived from HIV-1 gp120 presented in multiple conformations. Biochemistry, 2000, 39: 10866.
11 Dreusicke, D. et al. Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution . J. Mol. Biol., 1988, 199: 359.
12 Dreusicke, D. et al. The switch between two conformations of adenylate kinase. J. Mol. Biol., 1988, 203: 1021.
13 Russell, P. J., Jr. et al. Evidence for conformers of rabbit muscle adenylate kinase. J. Biol. Chem., 1990, 265: 11804.
14 Zhang, H. J. et al. Evidence for at least two native forms of rabbit muscle adenylate kinase in equilibrium in aqueous solution. J. Biol. Chem., 1998, 273: 7448.
15 Zhang, H. J. et al. Activation of adenylate kinase by denaturants is due to the increasing conformational flexibility at its active sites. Biochem. and Biophy. Res. Commun., 1997, 238: 382.
16 Zhang, H. J. et al. Refolding of urea-denatured adenylate kinase. Biochem. J., 1998, 333: 401.
17 Forood, B. et al. Stabilization of α-helical structures in short peptides via end capping. Proc. Natl. Acad. Sci. USA, 1993, 90: 838.
18 Gunasekaran, K. et al. Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. J. Mol. Biol., 1998, 275: 917.
19 Prieto, J. et al. C-capping and helix stability: the Pro C-capping motif. J. Mol. Biol., 1997, 274: 276.
20 Rose, G. D. et al. Turns in peptides and proteins. Advan. Protein Chem., 1985, 37: 1.
21 Cheng, H. N. et al. *Cis-trans* equilibrium and kinetic studies of acetyl-L-proline and glycyl-L-proline. Biopolymers, 1977, 16: 1465.
22 Grathwohl, C. et al. NMR studies of the rates of proline *cis-trans* isomerization in oligopeptides. Biopolymers, 1981, 20: 2623.
23 Stewart, D. E. et al. Occurrence and role of *cis* peptide bonds in protein structures. J. Mol. Biol., 1990, 214: 253.
24 Alexandrescu, A. T. et al. Hydrogen-1 NMR evidence for three interconverting forms of staphylococcal nuclease: Effects of mutations and solution conditions on their distribution. Biochemistry, 1989, 28: 204.
25 Alexandrescu, A. T. et al. Coupling between local structure and global stability of a protein: Mutants of staphylococcal nuclease. Biochemistry, 1990, 29: 4516.

26 Hinck, A. P. et al. NMR strategy for determining Xaa-pro peptide bond configurations in proteins: Mutants of staphylococcal nuclease with altered configurations at proline-117. Biochemistry, 1993, 32: 11810.

27 Kuwajima, K. et al. The Pro-117 to glycine mutation of staphylococcal nuclease simplifies the unfolding-folding kinetics. FEBS Letters, 1991, 290: 135.

28 Sugawara, T. et al. Folding of staphylococcal nuclease A studied by equilibrium and kinetic circular dichroism spectra. Biochemistry, 1991, 30: 2698.

29 Ikura, T. et al. Kinetic folding and *cis/trans* prolyl isomerism of staphylococcal nuclease: A study by stopped-flow absorption, stopped-flow circular dichroism and molecular dynamics stimulations. Biochemistry, 1997, 36: 6529.

30 Sheng, X. R. et al. Domain movement in rabbit muscle adenylate kinase might involve proline isomerization. FEBS Letters, 1997, 413: 429.

31 Richardson, J. S. The anatomy and taxonomy of protein structure. Adv. Protein Chem., 1981, 34: 167.

32 Otting, G. et al. Disulfide bond isomerization in BPTI and BPTI (G36S): An NMR study of correlated mobility in proteins. Biochemistry, 1993, 32: 3571.

33 Wüthrich, K. NMR in Biological Research: Peptides and Proteins, Amsterdam: North Holland, 1976.

34 Wüthrich, K. NMR of Proteins and Nucleic Acids. New York: Wiley, 1986.

35 Sørensen, M. D. et al. Elucidation of the origin of multiple conformations of the human α3-chain type Ⅵ collagen C-terminal Kunitz domain: The reorientation of the Trp ring. Journal of Biomolecular NMR, 1996, 8: 391.

36 Zweckstetter, M. et al. Structure and multiple conformations of the Kunitz-type domain from human type Ⅵ collagen alpha 3(Ⅵ) chain in solution. Structure, 1996, 4: 195.

37 Heald, S. L. et al. Sequencial NMR resonance assignment and structure determination of the Kunitz-type inhibitor domain of the Alzheimer's β-amyloid precursor protein. Biochemistry, 1991, 30: 10467.

38 Joseph, D. et al. Anatomy of a conformational change: Hinged "lid" motion of the triosephosphate isomerase loop. Science, 1990, 249: 1425.

39 Lolis, E. et al. Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5 Å resolution: Implications for catalysis. Biochemistry, 1990, 29: 6619.

40 Derewenda, U. et al. Catalysis at the interface: The anatomy of a conformational change in a triglyceride lipase. Biochemistry, 1992, 31: 1532.

41 Jia, Z. et al. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. Science, 1995, 268: 1754.

42 Rini, J. M. et al. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. Science, 1992, 255: 959.

43 Concha, N. O. et al. Rat annexin Ⅴ crystal structure: Ca^{2+}-induced conformational changes. Science, 1993, 261: 1321.

44 Williams, J. C. et al. Dynamics of the flexible loop of triosephosphate isomerase: The loop motion is not ligand gated. Biochemistry, 1995, 34: 8309.

45 Sheng, X. R. et al. An iso-random Bi Bi mechanism for adenylate kinase. J. Biol. Chem., 1999, 274: 22238.

46 Li, X. et al. Both native conformers of rabbit muscle adenylate kinase are active. FEBS Letters, 2000, 480: 235.

47 Harvey, A. L. Cardiotoxins from cobra venoms: possible mechanisms of action. J. Toxicol. Toxin Rev., 1985, 4: 41.

48 Fletcher, J. E. et al. Similarities and differences in mechanisms of cardiotoxin, melittin and other myotoxins. Toxicon, 1996, 34: 1301.

49 Vincent, J. P. et al. Molecular mechanism of cardiotoxin action on axonal membranes. Biochemistry, 1976, 15: 3171.

50 Dufourcq, J. et al. Specific binding of a cardiotoxin from *Naja mossambica mossmbica* to charged phospholipids detected by intrinsic fluorescence. Biochemistry, 1978, 17: 1170.

51 Dubovskii, P. V. et al. Membrane binding motif of the P-type cardiotoxin. J. Mol. Biol., 2001, 305: 137.

52 Dementieva, D. V. et al. Two forms of cytotoxin Ⅱ (cardiotoxin) from *Naja naja oxiana* in aqueous solution. Spatial structures with tightly bound water molecules. Eur. J. Biochem, 1999, 263: 152.

53 Branden, C. et al. Introduction to Protein Structure. New York-London: Garland Publishing, Inc., 1991, 302.

54 Chothia, C. et al. Conformations of immunoglobulin hypervariable regions. Nature, 1989, 342: 877.

55 Kempner, E. S. Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity. FEBS Letters, 1993, 326: 4.