FEBS Letters 509 (2001) 90–94 FEBS 25502

Crystal structure of a cyclic form of bovine pancreatic trypsin inhibitor

Istvan Botos[a], Zhibin Wu[b], Wuyuan Lu[b], Alexander Wlodawer[a,*]

[a] *Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, MD 21702, USA*
[b] *Institute of Human Virology, University of Maryland, Baltimore, MD 21201, USA*

Received 28 August 2001; revised 21 October 2001; accepted 24 October 2001

First published online 15 November 2001

Edited by Takashi Gojobori

Abstract The crystal structure of a cyclic form of a mutant of bovine pancreatic trypsin inhibitor has been solved at 1.0 Å resolution. The protein was synthesized by native chemical ligation and its structure is almost indistinguishable from the previously described recombinant form of the same mutant; however, the new loop containing the former termini became much better ordered.

Key words: Bovine pancreatic trypsin inhibitor; Cyclic protein; Atomic resolution

1. Introduction

Native proteins contain only linear polypeptide chains, although some peptidic antibiotics, defensins [1], plant protease inhibitors [2] and an antibacterial peptide [3] can be cyclic. Thus all the protein structures currently represented in the Protein Data Bank consist of only linear chains containing both amino and carboxy termini, with the exception of a few nuclear magnetic resonance (NMR) structures of proteins shorter than ~70 amino acids [2–6]. However, artificially prepared cyclic versions of several proteins have been reported [7–10]. Such artificial constructs proved to be interesting for studies on the influence of the termini on the stability and folding pathways. In particular, a cyclic form of bovine pancreatic trypsin inhibitor (cBPTI) was synthesized and its folding and stability properties were extensively analyzed [7,11,12]. In addition, the structure of cBPTI was studied by ^{1}H NMR in solution, although no atomic model resulted from this study that relied on early procedures of two-dimensional (2D) NMR [13].

Several recent developments convinced us that it might now be worthwhile to apply modern tools of structural biology to this old problem. While the original technique utilized to produce cBPTI relied on the treatment of native protein with water-soluble carbodiimide and resulted in rather low yield and purity of the product [7], we realized that a simple modification of the recently described procedure of total chemical synthesis of BPTI by native chemical ligation [14,15] could produce large amounts of extremely pure product. While crystal growth of native BPTI is not particularly easy and reproducible [16], it was recently reported that a four-point BPTI mutant (T11A, P13A, K15R, M52L) produces excellent crystals overnight [17]. The structure of that mutant, solved at ultrahigh resolution [18], is very similar to that of native BPTI, indicating that this form of the protein might be optimal for the preparation of cBPTI. Finally, current progress in collecting diffraction data at low temperature on synchrotron sources made it likely that an atomic resolution structure of cBPTI could be obtained with only a moderate effort.

2. Materials and methods

2.1. Materials

Boc-L-amino acids were purchased from Peptide Institute (Osaka, Japan). All other chemicals used in peptide syntheses and purifications, unless noted otherwise, were obtained from the same sources as previously described [14]. Analytical and preparative reversed-phase high performance liquid chromatography (HPLC) were performed on Waters 2690 and Delta Prep 600 systems, respectively, using Vydac C18 columns. Analysis of crude and purified peptides by electrospray ionization mass spectrometry was performed on Micromass ZQ-4000. Machine-assisted stepwise solid phase peptide synthesis (SPPS) was done on an Applied Biosystems 433A synthesizer using a custom-designed chemistry tailored from the in situ neutralization/HBTU activation protocol for Boc-SPPS [19].

2.2. Chemical synthesis of mutant cBPTI

The procedure for preparation of BPTI by native chemical ligation [14] has been modified to yield the cyclic version of the mutant protein. The chain was assembled starting from Gly37, continuing through the previous N terminus by addition of the C-terminal

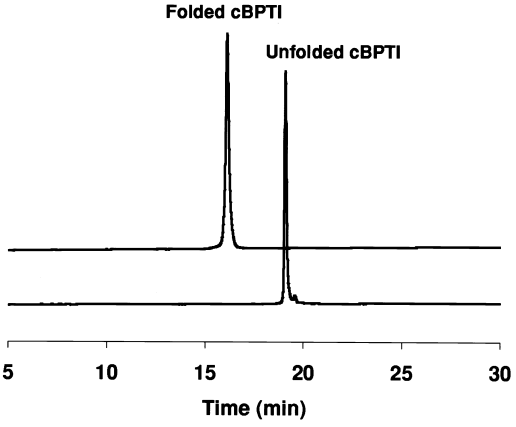

Fig. 1. Analytical HPLC traces of unfolded and folded cBPTI. Conditions were 15–40% acetonitrile containing 0.1% TFA over 30 min at a flow rate of 1 ml/min at 40°C.

*Corresponding author. Fax: (1)-301-846 6128.
E-mail address: wlodawer@ncifcrf.gov (A. Wlodawer).

PII: S0014-5793(01)03113-1

residues, and finishing the synthesis with Cys38. Synthesis of the rearranged 58-residue peptide thioester, i.e. (38–58)-(1–37)αCOSR (R = CH_2CO-Leu-OH), was accomplished on Boc-Gly-SCH_2CO-Leu-OCH_2-PAM (4-(hydroxymethyl)-phenylacetamidomethyl) resin. After chain assembly, the peptide was deprotected and cleaved by HF in the presence of 5% *p*-cresol at 0°C for 1 h, then loaded onto preparative reversed-phase HPLC for purification. To obtain cBPTI, 100 mg of purified peptide thioester were dissolved in 16 ml of 0.25 M phosphate buffer, pH 7.5, containing 6 M GuHCl and 1% thiophenol. The cyclization reaction, i.e. intra-molecular native chemical ligation, was completed in 3 h. The cBPTI polypeptide was further purified by HPLC using a gradient of 15–40% acetonitrile containing 0.1% TFA over 60 min. The observed mass of the cyclization product is 6453.2 Da, which is in good agreement with the expected value of 6453.5 Da calculated based on average isotope compositions. It is worth pointing out that inter-molecular ligation products were not detected under the conditions used. This is in general agreement with the fact that intra-molecular chemical reactions are always much faster than their inter-

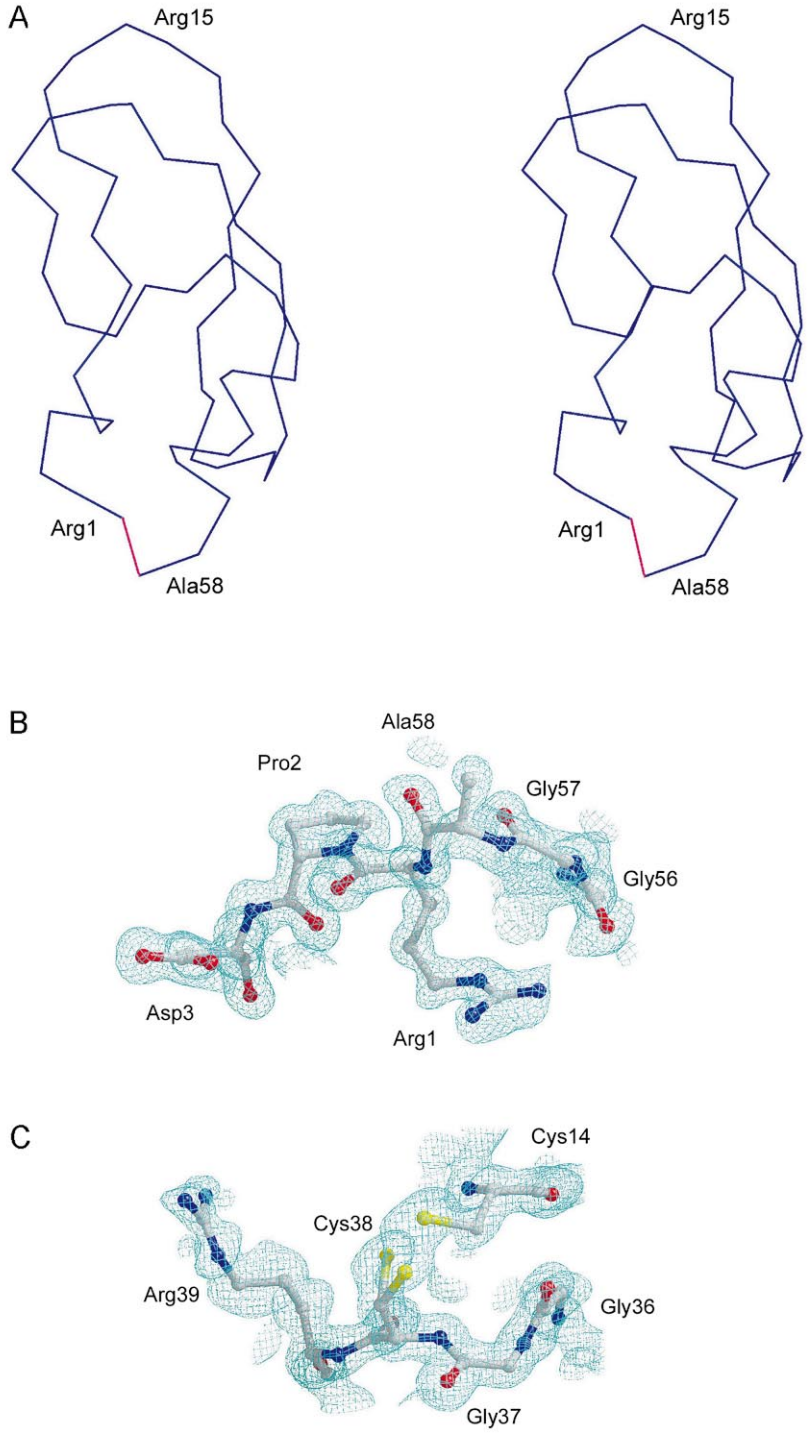

Fig. 2. Structure of the cyclic form of BPTI. A: Stereo tracing of the Cα backbone of the protein. The connection between residues 1 and 58 is marked in red. B: $2F_o-F_c$ electron density contoured at 1.5σ for the peptide bond between Ala58 and Arg1. C: $2F_o-F_c$ electron density contoured at 1.5σ for the peptide bond between Gly37 and Cys38, the site of chemical ligation in the synthesis of the peptide. Illustrations generated with Molscript [23], Bobscript [24] and Raster3D [25].

molecular counterparts, presumably due to increased effective local concentrations of reactants and/or favorable activation entropy or enthalpy in the transition state [20].

2.3. Folding and purification

Folding and disulfide formation was achieved by dissolving 50 mg purified cyclic BPTI polypeptide in 32 ml 6 M GuHCl containing 36 mg cysteine and 7.2 mg cystine, to which 68 ml water containing 840 mg $NaHCO_3$ was added. After gentle stirring overnight at room temperature, the folded product was obtained (Fig. 1). Compared with the reduced cBPTI, the folded protein eluted earlier on the analytical HPLC column, which is consistent with burial of hydrophobic residues during folding. The mass for the folded cBPTI was found to be 6447.3 Da, 6 mass units lower than that of the reduced one, indicative of the formation of three disulfide bonds. Interestingly, refolding of cBPTI was as efficient as the linear counterpart, suggesting that free termini are not a prerequisite for protein folding initiation. In fact Goldenberg and Creighton [11] demonstrated that both forms of BPTI followed an identical folding/unfolding pathway.

2.4. Crystallographic procedures

Crystals of folded cBPTI were grown by the previously described procedure [17] and were isomorphous with those of the mutant of the linear protein [18]. X-ray data were collected at 100 K on beamline X9B, NSLS, Brookhaven National Laboratory, with the ADSC Quantum4 CCD detector, using a wavelength of 0.92 Å. Data were processed using the HKL2000 suite [21], resulting in 32 347 unique reflections with the final *R*-merge of 7.0% (from the initial 728 960 reflections). The structure was refined anisotropically using the program SHELXL [22], resulting in a model with a root mean square deviation (r.m.s.d.) of bond lengths and angles of 0.017 Å and 2.08° respectively; $R = 11.2\%$, $R_{free} = 16.3\%$. The coordinates of cBPTI have been submitted to the Protein Data Bank (PDB) with the accession code 1k6u.

3. Results and discussion

3.1. Stability of cBPTI

Previous studies of cBPTI have shown that the introduction of a chemical bond between the termini preserves the inhibitory function of the protein [7]. It was also reported that the termini do not play an important role in the early stages of folding, but there were significant differences in the stabilities and rates of interconversions of some of the intermediates [11]. Studies utilizing 2D ^{1}H NMR showed that only the residues adjacent to the modification site in cBPTI are perturbed [13]. In the native structure, a salt bridge between Arg1 and Ala58 contributes ~4 kJ/mol to the protein stability, which is of the same magnitude as the stabilizing effect of the Arg1–Ala58 peptide bond. Thus, in the solution structure of cBPTI the salt bridge and the covalent bond have an almost equivalent effect on the stability of the protein [13]. The crystal structure of cBPTI reported here is in good agreement with these data. Superposition of cBPTI and linear BPTI structures shows a virtually identical main chain and larger deviations of terminal residues.

3.2. Comparison to BPTI

The structure of cBPTI (Fig. 2A) is extremely similar to that of the corresponding linear molecule, with the r.m.s.d. 0.1 Å between the Cα coordinates that excluded residues 57, 58 and 1. Both the peptide bond between the former termini (Fig. 2B) and the bond between Gly37 and Cys38 that had to be made in order to cyclize the polypeptide (Fig. 2C) are well ordered. The side chain of Cys38 is observed in a well-conserved dual conformation, in spite of this residue being involved in the cyclization reaction. The refined occupancies of the Sγ atoms from Cys38 are 0.6/0.4 which are identical to the occupancies reported for the same crystal form of the linear protein [18], confirming the static disorder of the two variants of BPTI. The temperature factors of the two S^{γ} atoms are also the same within experimental error. Since the chemical bond introduced between Cys38 and Gly37 is the same as the naturally occurring bond, it does not have an influence on the mobility of the Cys38 side chain. The average B factor of residue Gly37 is the same as the average B factor of the Cys38 main chain.

There is a smaller number of disordered side chains in the

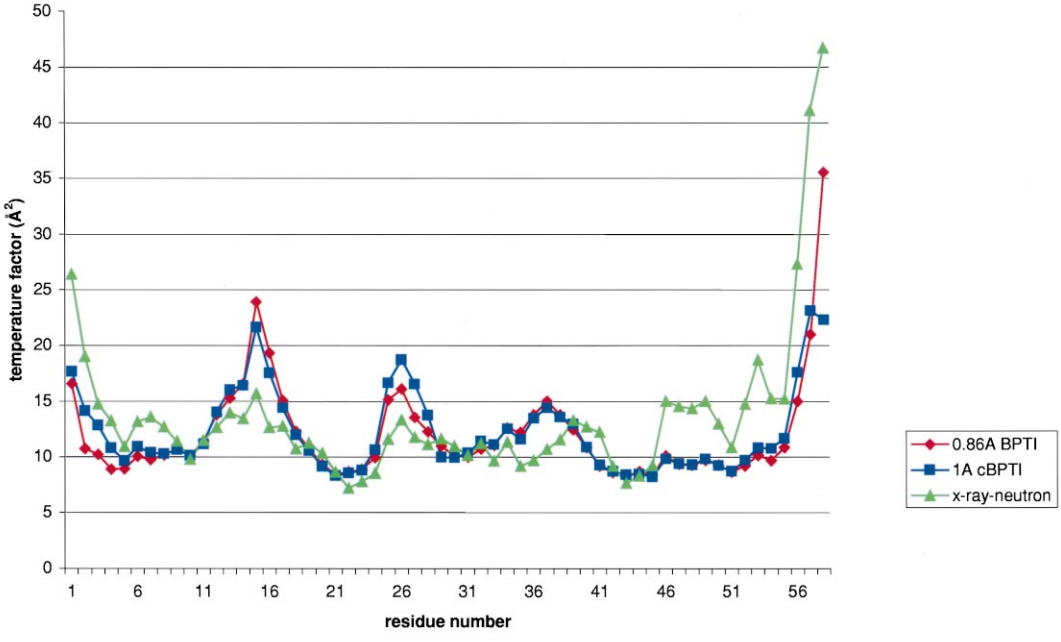

Fig. 3. Comparison of the average main-chain temperature factors for cBPTI (blue) with the 1 Å structure of native BPTI (green) [26] and the 0.86 Å structure of mutant BPTI (red) [18].

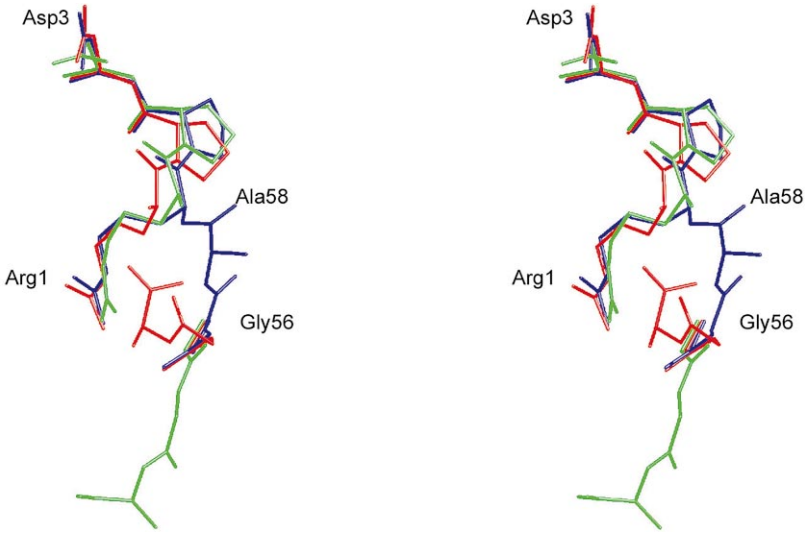

Fig. 4. Stereo representation of terminal residue conformations of cBPTI (blue), 1 Å structure of native BPTI (green) [26], and the 0.86 Å structure of mutant BPTI (red) [18]. The main chain atoms were superimposed using program Align from the CCP4 package [27].

cBPTI structure, all of them being conserved compared to the ultrahigh-resolution structure. However, residues Gly56–Gly57 are in a single, well-defined conformation, unlike the dual conformations in the 0.86 Å structure. This is due to the stabilizing effect of the nearby Arg1–Ala58 peptide bond. In the linear form of BPTI the dual conformation of residues Gly56–Gly57 is synchronized with the double conformation Lys26 side chain and the main chain of Leu29 [18]. In cBPTI Lys26 is slightly farther from both the ordered residues Gly56–Gly57 (more than 3.5 Å) and Leu29 (more than 5.1 Å). The sulfate ions are also well conserved between cBPTI and the 0.86 Å resolution structure. Arg20 has an ion-binding site, harboring a sulfate or a phosphate ion, depending on which ion was present in the crystallization condition.

3.3. Flexibility of the termini

The temperature factors of cBPTI are very similar to those of the linear native protein except for the termini, which are subject to much lower thermal motions than those reported in the high-resolution structures of either native or mutant BPTI (Fig. 3). This linked-terminal region with reduced thermal motion becomes a new flexible loop in cBPTI, with temperature factors comparable to the other loops, such as loop 13–17 and 25–28.

A superposition of the main chain atoms in three high-resolution crystal structures shows the range of different conformations adopted for residues 56–58 and 1–3 (Fig. 4). The N-terminal residues Arg1–Asp3 are fairly well conserved in these structures, N and Cα of Arg1 showing the largest deviations: 1.7 and 1.1 Å relative to the same atoms in cBPTI. The C-terminal residues from Gly56 show much larger deviations. In the 0.86 Å crystal structure, residues 56–58 deviate by an overall 3.5 Å and in the 1 Å native structure Gly57 has an overall deviation of 4.5 Å and Ala58 of 11.0 Å. This is also apparent from the temperature factors of these residues (Fig. 3), with a much higher flexibility of the C-terminus than the N-terminus.

Acknowledgements: The impetus for this work was provided by S.B.H. Kent and we are grateful for his encouragement. We thank Z. Dauter for help with data collection at NSLS.

References

[1] Tang, Y.Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C.J., Ouellette, A.J. and Selsted, M.E. (1999) Science 286, 498–502.
[2] Felizmenio-Quimio, M.E., Daly, N.L. and Craik, D.J. (2001) J. Biol. Chem. 276, 22875–22882.
[3] Gonzalez, C., Langdon, G.M., Bruix, M., Galvez, A., Valdivia, E., Maqueda, M. and Rico, M. (2000) Proc. Natl. Acad. Sci. USA 97, 11221–11226.
[4] Daly, N.L., Koltay, A., Gustafson, K.R., Boyd, M.R., Casas-Finet, J.R. and Craik, D.J. (1999) J. Mol. Biol. 285, 333–345.
[5] Craik, D.J., Daly, N.L., Bond, T. and Waine, C. (1999) J. Mol. Biol. 294, 1327–1336.
[6] Saether, O., Craik, D.J., Campbell, I.D., Sletten, K., Juul, J. and Norman, D.G. (1995) Biochemistry 34, 4147–4158.
[7] Goldenberg, D.P. and Creighton, T.E. (1983) J. Mol. Biol. 165, 407–413.
[8] Yu, Q., Lehrer, R.I. and Tam, J.P. (2000) J. Biol. Chem. 275, 3943–3949.
[9] Iwai, H., Lingel, A. and Plückthun, A. (2001) J. Biol. Chem. 276, 16548–16554.
[10] Camarero, J.A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D.P. and Muir, T.W. (2001) J. Mol. Biol. 308, 1045–1062.
[11] Goldenberg, D.P. and Creighton, T.E. (1984) J. Mol. Biol. 179, 527–545.
[12] Goldenberg, D.P. (1985) J. Cell Biochem. 29, 321–335.
[13] Chazin, W.J., Goldenberg, D.P., Creighton, T.E. and Wüthrich, K. (1985) Eur. J. Biochem. 152, 429–437.
[14] Lu, W., Starovasnik, M.A. and Kent, S.B. (1998) FEBS Lett. 429, 31–35.
[15] Dawson, P.E. and Kent, S.B. (2000) Annu. Rev. Biochem. 69, 923–960.
[16] Wlodawer, A., Deisenhofer, J. and Huber, R. (1987) J. Mol. Biol. 193, 145–156.
[17] Czapińska, H., Otlewski, J., Krzywda, S., Sheldrick, G.M. and Jaskólski, M. (2000) J. Mol. Biol. 295, 1237–1249.
[18] Addlagatta, A., Krzywda, S., Czapińska, H., Otlewski, J. and Jaskólski, M. (2001) Acta Crystallogr. D 57, 649–663.

[19] Schnolzer, M., Alewood, P., Jones, A., Alewood, D. and Kent, S.B. (1992) Int. J. Pept. Protein Res. 40, 180–193.
[20] Jencks, W.P. (1987) Catalysis in Chemistry and Enzymology, Dover Publications, New York.
[21] Otwinowski, Z. and Minor, W. (1997) Methods Enzymol. 276, 307–326.
[22] Sheldrick, G.M. and Schneider, T.R. (1997) Methods Enzymol. 277, 319–343.
[23] Kraulis, P.J. (1991) J. Appl. Crystallogr. 24, 946–950.
[24] Esnouf, R.M. (1999) Acta Crystallogr. D 55, 938–940.
[25] Meritt, E.A. and Murphy, M.E.P. (1994) Acta Crystallogr. D 50, 869–873.
[26] Wlodawer, A., Walter, J., Huber, R. and Sjölin, L. (1984) J. Mol. Biol. 180, 301–329.
[27] Collaborative Computational Project Number 4, (1994) Acta Crystallogr. D 50, 760–763.