PAWEŁ TŁOMAK and KORNEL NOWAK *

A PREDICTED SECONDARY STRUCTURE OF PROTEIN PROTEASE INHIBITORS **

Department of Biophysics, Department of Biochemistry *, Institute of Physiology and Biochemistry, Medical School; Chalubińskiego 10; 50-368 Wroclaw, Poland

Basing on the Chou & Fasman method (Biochemistry, 13, 222-245; 1974) and the known amino acid sequences, the α-helical, β-sheet, β-turn and random-coil regions were predicted for protein protease inhibitors. It appears that in all the inhibitors examined there is a region conserved in the vicinity of the active site, composed of one β-turn with an adjacent β-sheet structure, and a second β-turn situated in the other part of the polypeptide chain and linked with the first one by a disulphide bridge. Two disulphide bridges between Cys2 and Cys25, and between Cys10 and Cys21 were proposed for the squash seed inhibitor. It is suggested that the two β-turns play an essential role in the process of trypsin inhibition by protein protease inhibitors.

Protein protease inhibitors are common in Nature; they are components of various plant and animal tissues as well as of microorganisms (Laskowski & Sealock, 1971). Studies on the mechanism of protease inhibition, and the structure/function relationship of the protease inhibitors, are of importance both for elucidation of the physiological role of this type of inhibitors, and for their clinical application (Rzeszotarska & Wiejak, 1976).

The amino acid sequences of a large number of the protein protease inhibitors are known, and on this basis the inhibitors have been classified into homologous families (Hokama et al., 1976; Bartelt et al., 1977; Szilagyi & Szilagyi, 1978). On the other hand, the three-dimensional structure has been determined for only three protein protease inhibitors,

** This work was supported by the Polish Academy of Sciences within the project MR.I.12.
namely: BPTI\(^1\) (of Kunitz type), (Huber et al., 1970) and its complex with bovine trypsin (Rühlmann et al., 1973; Huber et al., 1974), STI in complex with porcine trypsin (Sweet et al., 1974; Blow et al., 1974) and SSI (Mitsui et al., 1979) and its complex with subtilisin BPN\(^{\prime}\) (Hirono et al., 1979).

It is commonly accepted that the primary structure of a protein constitutes sufficient information for folding of the polypeptide chain into a biologically active native conformation. In the present work, advantage was taken of the method of Chou & Fasman (1974) for prediction of the secondary structure of protein trypsin inhibitors on the basis of their amino acid sequences. The structural arrangement of active site was proposed.

METHODS

\(\alpha\)-Helical and \(\beta\)-sheet structures. These structures in trypsin inhibitors were predicted taking into account the conformational parameters given by Chou & Fasman (1977a). The fragments of the inhibitor polypeptide chains that were not predicted as \(\alpha\)-helical or of \(\beta\)-sheet structure, were referred to as random coil regions.

\(\beta\)-Turns. Independently of the prediction of \(\alpha\)-helical and \(\beta\)-sheet structures (Chou & Fasman, 1977b), the \(f_i\), \(f_{i+1}\), \(f_{i+2}\) and \(f_{i+3}\) parameters were used to estimate the probability of occurrence of the particular amino acid residues in each of the four consecutive positions of the \(i\)-th \(\beta\)-turn, as well as the set of conformational parameters \(P_i\) which express the "conformational potential" of the amino acid residues in \(\beta\)-turn (Chou & Fasman, 1979a). The tetrapeptide is assumed to form \(\beta\)-turn when it fulfills the following conditions: the probability value of \(\beta\)-turn in position no. \(i\), \(P_i = f_i \times f_{i+1} \times f_{i+2} \times f_{i+3}\), is higher than \(0.75 \times 10^{-4}\), the mean value \(\langle P_i \rangle\) of the conformational parameter \(P_i\) fulfills the conditions: \(\langle P_{\alpha} \rangle < \langle P_{\beta} \rangle < \langle P_{\beta'} \rangle\) and \(\langle P_\gamma \rangle > 1\), and when, moreover, it is not included in the regions predicted either as \(\alpha\)-helix or \(\beta\)-sheet. In the case when some residues of the assumed \(\beta\)-turn coincided with the residues included in the \(\alpha\)-helix or \(\beta\)-sheet, they were distinctly designated in Figs. 1, 2 and 3. On the other hand, two or more coinciding \(\beta\)-turns were considered as a single region of high probability of occurrence. The predicted \(\beta\)-turns were considered as conserved when they were in the same position of the amino acid sequence of the analysed group of inhibitors. The

\(^1\) Abbreviations used: BPTI, basic pancreatic trypsin inhibitor (Kunitz); STI, soybean trypsin inhibitor; SSI, Streptomyces subtilisin inhibitor; PSTI, pancreatic secretory trypsin inhibitor (Kazal); Bull SI, bull seminal plasma inhibitor; Boar SI, boar seminal plasma inhibitor; BBSI, Bowman Birk's inhibitor from soybean; GBPI-II, Component II of garden bean inhibitor; LBPI-IV, Component IV of lima bean inhibitor, and CPPI, inhibitor from chick pea.
predicted positions of α-helices, β-sheets and β-turns were denoted by the consecutive numbers of the initial and terminal residues. The mean values of the conformational parameters $\langle P_\alpha \rangle$, $\langle P_\beta \rangle$ and $\langle P_\tau \rangle$ for the predicted α-helical, β-sheet and β-turn structures were computed using ODRA 1304. The predicted position of the particular secondary structures was calculated without the use of the computer.

Amino acid sequence of the inhibitors. For prediction of the elements of secondary structure, the data concerning the amino acid sequences of the following 15 protein proteinase inhibitors were used: trypsin inhibitor E from the venom of *Dendroaspis polyolepis polyolepis* (Joubert & Spydrom, 1978), trypsin inhibitors from the venoms of *Nemachatus haemachatus* and *Naja nivea* (Hokama et al., 1976), pancreatic secretory trypsin inhibitors from human (Bartelt et al., 1977), porcine (Bartelt & Greene, 1971; Tschesche et al., 1969), bovine (Greene & Bartelt, 1969) and ovine (Hochstrasser et al., 1969) pancreases, trypsin-acrosin inhibitors from bull and boar seminal plasma (according to Čechova & Meloun, 1979), inhibitor from canine submandibular glands (Hochstrasser et al., 1975), Bowman Birk’s inhibitor from soybean (Odani & Ikenaka, 1972). Component II of garden bean inhibitor (Wilson & Laskowski, 1975), Component IV of lima bean inhibitor (Tan & Stevens, 1971), inhibitor from chick pea (Belew & Eaker, 1976) and inhibitor from squash seeds (our results).

The amino acid residues of the above inhibitors were numbered in such a way that Lys. or Arg of the active sites had the same number in all the sequences examined. For designation of the particular residues, the one-letter system was used (IUPAC-IUB Commission on Biochemical Nomenclature, 1969; 1970).

RESULTS

The 15 protein trypsin inhibitors examined have been classified into three homologous groups (Laskowski & Seacock, 1971): animal inhibitors of the Kunitz type, animal secretory inhibitors of the Kazal type, and plant inhibitors which all (except the inhibitor from squash seeds) show homologous sequences.

Secondary structures of these inhibitors were predicted using the method of Chou & Fasman (1974), and presented in Figs. 1, 2 and 3 in a way which enables direct comparison of the vicinity of the active sites of the predicted structures.

The predicted secondary structures of three trypsin inhibitors of the Kunitz type (Fig. 1) were compared with that of BPTI suggested by Chou & Fasman (1978) and that derived from crystallographic data (Huber et al., 1970). In this group of inhibitors, 36% of the amino acid residues are invariant. The predicted β-turns in positions 12-15 and 41-44 of all the inhibitors of this group were conserved despite differences in the
Fig. 1. Predicted secondary structures in the following trypsin inhibitors: a. BPTI (known X-ray structure: Huber et al., 1970); b. BPTI (predicted by Chou & Fasman, 1978); c. from venom of Hemachatus haemachatus; d. from venom of Naja nivea, and e. inhibitor E.

α-helix: \( \gamma \)-sheet: \( \delta \)-turn: Lys or Arg residues incorporated in the active site of the inhibitor; Cys residues of the disulphide bridges (see text).

amino acid sequence of the respective tetrapeptides. β-Turns of the inhibitors from the venoms of Hemachatus haemachatus and Naja nivea were predicted in position 12 - 15 despite low value of p<sub>T</sub> (0.28 × 10⁻⁴) because in the immediate vicinity of the 12 - 15 region the p<sub>T</sub> values are much lower than 0.28 × 10⁻⁴, making the occurrence of β-turns in this region highly probable. β-Sheets are located probably in the immediate vicinity (at the C-terminal) of β-turns in the 12 - 15 region. Arg and Lys residues of the active sites of all the four inhibitors of this group occupy a boundary position of the adjacent above-mentioned β-turns and β-sheets, and are located at the C-terminal of β-turns. Thus, in the four inhibitors discussed, the parts of the polypeptide chains linked by a disulphide bridge in position Cys14 - Cys38, assumedly form β-turns.
Fig. 2. Predicted secondary structures in the following trypsin inhibitors: a, human PSTI; b, porcine PSTI; c, bovine PSTI; d, ovine PSTI; e, Bull SI; f, Boar SI, and g, from canine submandibular glands. Designations are the same as in Fig. 1.
The comparison of the secondary structures of the four protein trypsin inhibitors of the Kunitz type and that of BPTI established on the basis of X-ray analysis, suggests topological homology of the structures in the vicinity of the active sites, apart from the homology of their amino acid sequences.

In the predicted secondary structures of seven secretory trypsin inhibitors of the Kazal type (Fig. 2), only 18% of the amino acid residues were invariant. Despite differences in their sequences, the predicted β-turns in positions 9 - 12, 13 - 16, 14 - 17 and 25 - 28, as well as 31 - 34, were conserved in all the four pancreatic inhibitors examined; the same was observed for β-sheets in position 18 - 24. The results obtained for the remaining three secretory inhibitors were similar, except that the adjacent β-turns and β-sheets predicted in the vicinity of the active site of Bull SI and Boar SI were positioned in the reverse order as compared with the pancreatic inhibitors. The disulphide bridges in positions Cys16 - Cys35 link these regions of the polypeptide chains of secretory inhibitors which form β-turns (Fig. 2). Arg or Lys residues of the active sites occupy the boundary position in the adjacent β-turns and β-sheets. Thus, a topological homology in the vicinity of the active site is also possible in the case of these inhibitors. Arg or Lys residues of the active sites are located at the N-terminal of β-turn in Bull SI and Boar SI, or at the C-terminal of β-turn in the remaining secretory inhibitors.

As concerns plant inhibitors, the structures of five trypsin inhibitors were predicted (Fig. 3). In the case of GBPI-II and CPPI, only fragments

Fig. 3. Predicted secondary structures in the following trypsin inhibitors: a. BBSI: b. GBPI-II: c. LBPI-IV: d. CPPI and e. from squash seeds. For GBPI-II and CPPI fragments of the amino acid sequences are shown. Designations are the same as in Fig. 1.
of the secondary structures are shown. In the immediate vicinity of the active sites of these inhibitors β-turns are probably adjacent to β-sheet structures. Similarly as with the Kunitz- and Kazal-type inhibitors, disulphide bridges link the regions of the polypeptide chains predicted as β-turns: in the inhibitors BBSI and LBPI-IV the bridges are in positions Cys12 - Cys58, in GBPI at Cys5 - Cys12, and in the inhibitor from squash seed, probably at Cys13 - Cys36 (the position of the disulphide bridge for this fragment of the amino acid sequence of CPPI is unknown; Belew & Eaker, 1976). Arg or Lys of the active sites of these inhibitors are located, respectively, at the C- and N-terminal of β-turn, in the squash seed inhibitor and the remaining four inhibitors. Also in this case topological homology in the vicinity of the active sites can be reasonably suggested for these inhibitors.

The secondary structures predicted for the protein trypsin inhibitors can be also presented in the graphical form of Chou & Fasman (1979b), visualizing close proximity of those fragments of polypeptide chains which form the active site. The predicted secondary structures of three trypsin inhibitors of the known position of disulphide bridges, and the predicted structure of the squash seed inhibitor, with the proposed position of these bridges, are illustrated in Fig. 4.

On the basis of the predicted secondary structures of 15 protein trypsin inhibitors it can be concluded that the three-dimensional structure of

---

Fig. 4. Schematic diagrams of predicted secondary structures in a, BPTI; b, human PSTI; c. BBSI and d. inhibitor from squash seeds (see text). Arrows indicate the hydrolysed peptide bonds.
these inhibitors includes, in the vicinity of the active site, a β-turn with an adjacent β-sheet, and another β-turn situated in close proximity to the first β-turn due to the disulphide bridge. Moreover, no α-helices could be expected in the immediate vicinity of the active sites. The occurrence of the two β-turns in the active site of trypsin inhibitors, and the predicted secondary structure of the squash seed inhibitor suggest that the first disulphide bridge of this inhibitor was positioned at Cys2 - Cys25, and the second at Cys10 - Cys21. The former links the two predicted β-turns, which form the active site containing a peptide bond (hydrolysed by trypsin) between Arg and Ile residues. (Numbering of the disulphide bridges proposed for the squash seed inhibitor is consistent with the numbering of the amino acid residues of this inhibitor).

DISCUSSION

A search for similarities and differences in the three-dimensional structure of protein protease inhibitors is essential for elucidation of the mechanism of inhibition.

Although the amino acid sequence is not homologous in the 15 protein inhibitors discussed, the predicted secondary structures of all these inhibitors contain a region composed of two β-turns linked by a disulphide bridge. Thus, a topological homology of these inhibitors pertains to the vicinity of the active sites. In each of the inhibitors, the Lys or Arg residues of the active site are positioned either at the N- or C-terminal of the predicted β-turn. It should be noted that an identical region composed of two β-turns linked by a disulphide bridge occurs in BPTI of a known three-dimensional structure (Huber et al., 1970). Thus, the predicted secondary structures of the protein trypsin inhibitors include a region homologous to that established for BPTI by X-ray analysis. Concerning the inhibitory activity of protein protease inhibitors, the results presented point to an important role of the structure of that region which is conserved irrespective of the differences in the sequence of the amino acid residues (i.e. the vicinity of the active site). Conservation of β-turns was also demonstrated by Chou & Fasman (1979b) for six other protein protease inhibitors homologous with BPTI. The results of the present work confirm their results in relation to a larger group of inhibitors of non-homologous amino acid sequences. Also the rigid structure of the protein protease inhibitors in the vicinity of the active site was confirmed; it is due to the presence both of β-sheet and the disulphide bridge. Since, in all the protein trypsin inhibitors analysed, the Lys or Arg residues of the active site are present at the N- or C-terminal of the predicted β-turns, two groups of inhibitors can be distinguished showing topological similarity but differing in the position of Lys or Arg with respect to β-turns. It should, however, be noted that this suggestion is based solely on prediction of
secondary structures of the inhibitors, and cannot be considered as a basis for their unequivocal classification. Moreover, there are difficulties in interpretation because of slight differences in the amino acid sequences—in spite of similar α-helical and β-sheet potentials—in the homologous regions of amino acid sequences of the inhibitors; for example, homologous regions in positions 19-35 of the inhibitor from venom of *Hemachatus haemachatus* and from *Naja nivea* form two different secondary structures.

The results obtained suggest that the presence of two β-turns in the active site plays a much greater role in the inhibitory activity of the trypsin inhibitors, than the presence of the disulphide bridge linking these β-turns. This suggestion is supported by the presence of two β-turns in the active site of STI and the lack of a disulphide bridge in this inhibitor (Sweet *et al.*, 1974). Moreover, the disulphide bridge Cys14-Cys38 is not indispensable for stabilization of BPTI conformation (Creighton, 1975).

Generally, interaction of a protein substrate with a protease occurs in three steps: binding of substrate to protease; hydrolysis of the appropriate peptide bond; and dissociation of the hydrolysis products. In the presence of inhibitor, the first step is not affected, the occurrence of the second is possible, while the third step does not take place. Thus, this last step would be decisive for the action of the inhibitor on protease, and distinguishes the protein inhibitor from the substrate hydrolysed in the non-denatured form. Taking into account the results of the present work, it could be suggested that the two β-turns situated in close proximity play an important role in formation of a steric hindrance preventing dissociation of the inhibitor from protease after a possible hydrolysis of the peptide bond in the active site.

Although the prediction of the secondary structures of the protein trypsin inhibitors was based on a statistical method which does not permit an unequivocal positioning of the particular structures, this method can serve as a starting point for studies on the mechanism of interaction of protein inhibitors with proteases.

**REFERENCES**


**PRZEWIDYWANE STRUKTURY DRUGORZĘDOWE BIAŁKOWYCH INHIBITORÓW PROTEAZ**

**Streszczenie**

Korzystając z metody Chou i Fasmana przewidywania struktury drugorzędowej białek oraz znanych sekwencji reszt aminokwasowych inhibitorów proteaz, szczególnie trypsyny, obliczono prawdopodobne występowanie α-helików, struktur β, β-skrętów oraz obszarów nieuporządkowanych w ich liniowej strukturze pierwszorzędowej. Z obliczeń tych wynika, że w otoczeniu centrów reaktywnych wszystkich zbadanych inhibitorów zachowany jest obszar składający się z β-skrętu sąsiadującego ze strukturą β oraz z drugiego β-skrętu położonego w innej części łańcucha polipeptydowego a połączonemu z pierwszym mostkiem dwusiarczkowym. Zaproponowano dwa mostki dwusiarczko w pomiędzy Cys2 i Cys25 oraz Cys10 i Cys21 dla inhibitora z nasion dyni. Przedstawiono hipotezę o istotnej roli dwóch β-skrętów w procesie hamowania trypsyny przez białkowe inhibitory proteaz.

Received 5 December, 1980.