Analysis of the black-eyed pea trypsin and chymotrypsin inhibitor-α-chymotrypsin complex

Sonia M. de Freitas[a,*], Luciane V. de Mello[b], Maria Cristina M. da Silva[b], Gerrit Vriend[c], Goran Neshich[b], Manuel M. Ventura[a]

[a] *Department of Cellular Biology, Biophysics Laboratory, University of Brasilia, 70910-900 Brasilia - DF, Brazil*
[b] *CENARGEN-EMBRAPA, Protein Engineering and Bioinformatics Group, SAIN. Parque Rural, Final W5 Norte, 70849-970 Brasilia - DF, Brazil*
[c] *EMBL, Heidelberg, Germany*

Received 23 December 1996; revised version received 2 April 1997

Abstract The black-eyed pea trypsin and chymotrypsin inhibitor (BTCI) is a member of the Bowman-Birk protease inhibitor (BBI) family. The three-dimensional model of the BTCI-chymotrypsin complex was built based on the homology to Bowman-Birk inhibitors with known structures. An extensive theoretical and experimental study of these known structures has been performed. The model confirms the ideas about Bowman-Birk inhibitor structure-function relations and agrees well with our experimental data (circular dichroism, IR and light scattering). The electrostatic potentials at the enzyme-inhibitor contact surface reveal a pattern of complementary electrostatic potentials from which mutations can be inferred that could give these inhibitors an altered specificity.

Key words: Bowman-Birk inhibitor; Serine protease inhibitor; Enzyme-inhibitor complex; Model by homology

1. Introduction

Plant protease inhibitors are frequently found in *Leguminosae* seeds. Most are proteins of low molecular weight that inhibit trypsin and/or chymotrypsin [1]. Their physiological significance in plants has been extensively investigated but has not yet been fully explained.

Plant protease inhibitors are protective agents that are part of the arsenal of defence mechanisms that plants use to protect against attack by microoganisms, and insects. These substances are also involved in the protection of the seeds against environmental hazards during the initial phases of germination and in establishment of the seedling [2].

There is accumulated evidence that consumption of seeds containing protease inhibitors lowers the incidence of breast, colon, prostatic, oral, and pharyngeal cancers [3]. Several studies have been performed which show that the soybean-derived Bowman-Birk inhibitor can prevent carcinogenesis in vivo and malignant transformation in vitro [4,5].

Bowman-Birk double-headed protease inhibitors have molecular weights ranging from 7 to 10 kDa and have a high disulphide bond content. They have two independent, non-overlapping sites for trypsin and chymotrypsin inhibition, respectively [1].

Black-eyed pea trypsin and chymotrypsin inhibitor (BTCI) was isolated and purified to homogeneity from *Vigna unguiculata* (L) Walp seeds (cowpea, cv. Seridó). These cowpea beans provide the main food protein source for most of the North-Eastern Brazilian population. BTCI is a single polypeptide chain of 83 residues [6]. Its amino acid sequence was reported by Morhy and Ventura [7]. The sites Lys26-Ser27 and Phe53-Ser54 in BTCI inhibit trypsin and chymotrypsin, respectively.

The three-dimensional structures of two members of the BBI family, a classical BBI from soybeans and the Bowman-Birk-type protease inhibitor, have been studied by X-ray crystallography [8], and NMR [9]. The coordinates are available from the PDB [10] (PDB id = 1BBI and 1PI2, respectively). An analysis of 1BBI and its complex with trypsin has been reported [11]. The conformation of the free and complexed inhibitors are quite similar.

The loops containing the inhibition sites are very exposed and thus easily accessible to proteolytic enzymes. Several highly conserved residues seem important for fixing the conformation of the inhibition loop in a conformation that allows the inhibitors to bind very tightly to the proteases. The stability and the maintenance of the active conformation of all BBI-like inhibitors are mainly attributed to a large number of (conserved) disulphide bridges [12].

A routine verification of the correctness of the two known BBI structures seemed to indicate that these molecules were not solved correctly. The core of the molecules is far too hydrophilic, and the protein surface is far too hydrophobic. However, analysis of the full crystal packing showed that most of hydrophobic part of the surface is involved in crystal contacts. The experimental work about BTCI (light scattering) indicates that the molecule is oligomeric in solution.

We propose 3-D models for BTCI and a BTCI-α-chymotrypsin complex. These models are built based on the homology of BTCI with the Bowman-Birk inhibitors for which the structures are known. The models are discussed in view of experimental data obtained from light scattering [13], CD and IR spectroscopy [14,15], etc. The models are presently guiding protein engineering experiments aimed at alteration of the inhibitor stability and specificity in order to obtain inhibitors that can aid plant defence.

2. Materials and methods

2.1. Modelling BTCI

Multiple sequence alignments were performed using the PILEUP program in the Genetics Computer Group (GCG) software package. The three-dimensional structure of BTCI has been modelled on the basis of the coordinates of the NMR structure [9] and X-ray crystal structure [8] of homologous inhibitors and the sequence alignment shown in Fig. 1. Molecular modelling and molecular mechanics were done using INSIGHT II, with Discover (Biosym Technologies,

*Corresponding author. Fax: (55) (61) 2721497.
E-mail: sfreitas@cenargen.embrapa.br

PII S0014-5793(97)00419-5

San Diego, 1991) and WHAT IF [16], on Silicon Graphics 4D70GT and INDIGO ELAN workstations.

Models built with WHAT IF were obtained using the protocol described before [17]. This protocol relies heavily on position-specific rotamer distributions [18]. Models built with INSIGHT II were constructed using the HOMOLOGY module of this package. Where these two models disagreed, manual decisions were made based on analyses using Procheck [19] and the structure verification options of WHAT IF.

2.2. Modelling the BTCI-α-chymotrypsin complex

The BTCI-α-chymotrypsin complex has been modelled using the homology with the complex of turkey ovomucoide inhibitor third domain (OMTKY3) for which the coordinates were determined by X-ray crystallography [20] (PDB id=1CHO).

The inhibitor was docked in the enzyme by superimposing the BTCI loop Thr-42-Phe43-Ser44-Ile45 onto the OMTKY3 residues Thr-17-Leu18-Ser19-Tyr20. The side chains of the binding β-hairpin of BTCI were modelled in agreement with the corresponding side chains of residues in the binding loop of OMTKY3. Only minor modifications were needed to avoid Van der Waals clashes. Standard rotamer libraries [21] were used to guide these (small) manual side chain modifications.

This initial complex was energy minimised using Discover. In the minimisation procedure only those residues were incorporated that have at least one atom within 5 Å of a residue that is part of the active site. This was done to avoid prohibitive CPU time requirements. Distance constraints were placed between pairs of atoms at the complex interface that are known to be essential for proper positioning of the inhibitor in the protease active site. These constraints mainly fix hydrogen bonds between the BTCI backbone (Phe-43) and the catalytic triad, and maintain the proper orientation of the active site residues. Energy minimisation was done until the maximum derivative of the energy with respect to the atomic positions was less than 0.001 kcal/mol/Å. The molecular dynamics simulation was run at 500 K in vacuum to avoid restraining the molecule to a local minimum with high energy. The conformation during this molecular dynamics run that had the lowest potential energy was energy minimised (using a conjugate gradient algorithm) to converge below 0.0001 kcal/mol/Å.

2.3. Analysis of the interface

A Connolly-type surface [22] was generated with INSIGHT II using a water probe radius of 1.4 Å. The electrostatic potential was generated at neutral pH using the program GRASP [23]. Calculations were done simulating an aqueous solution with an ionic strength of zero and dielectric constants of 80 and 2 for solvent and protein respectively. The electrostatic potential is interpolated on a solvent-accessible surface. The latter was obtained using a probe radius of 1.6 Å to represent a water molecule. All calculations were performed at 300 K. The positive and negative potentials are contoured at 2 kcal/mol/Å and −2 kcal/mol/Å respectively.

3. Results and discussion

3.1. Sequence comparisons

The multiple sequence alignment of several representative inhibitors from the Bowman-Birk family, is shown in Fig. 1. There is a high level of sequence identity throughout the whole family and the cysteine bridging pattern is entirely conserved. The residues involved in inhibition of trypsin are identical in BTCI and 1BBI, and highly conserved in 1PI2. There is a remarkably high conservation in the distribution of polar and non-polar residues throughout the molecules, including the chymotrypsin inhibition sites.

3.2. The BTCI model

Fig. 2 shows the 3-D model of BTCI with the two inhibitor loops indicated. Phe-43 is the key residue in the α-chymotryptic inhibition site and Lys-16 in the trypsin inhibition site. Phe-43 is substituted by a leucine in 1BBI molecule. Both inhibition sites are located in the tip of the corresponding β-hairpin.

The packing quality control option [24] of the WHAT IF program indicated that the BTCI model had some unusual structural features: aromatic residues are solvent exposed, and several hydrophilic residues are buried. However, this is also the case in the experimentally determined structures that were used as templates for the modelling.

The NMR structure data presented by Werner and Wemmer [9] indicated that many NOEs are observed inside the β-hairpins that contain the inhibition sites, but very few in the rest of the molecule. This indicates that these two β-hair-

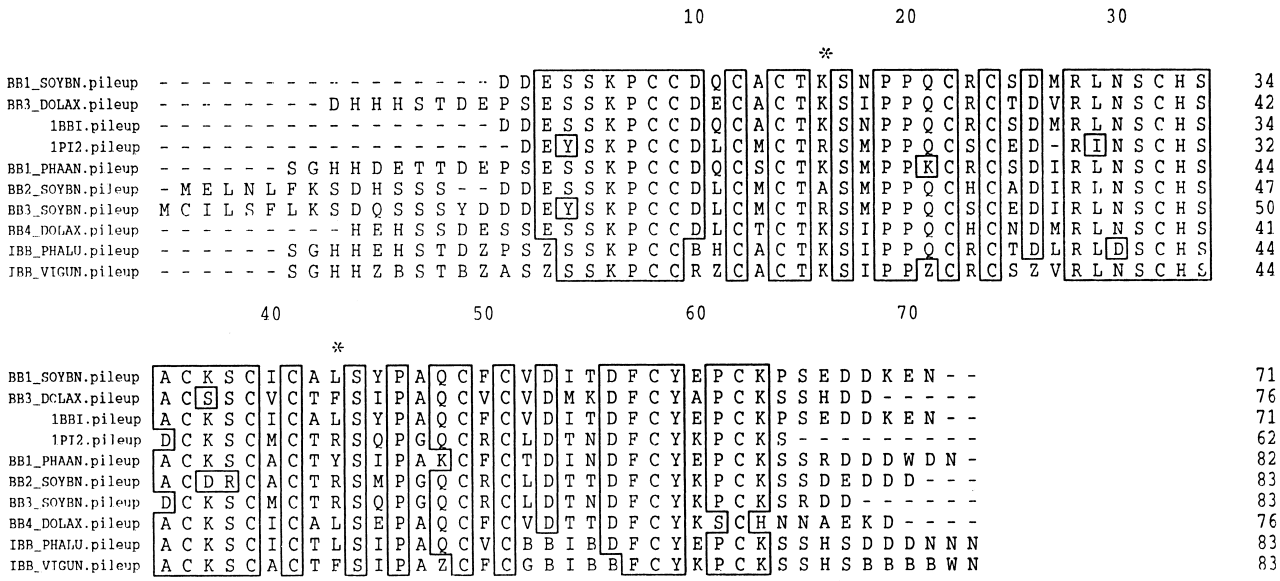

Fig. 1. Alignment of representative Bowman-Birk-type protease inhibitors. Asterisks indicate the interaction sites for trypsin and chymotrypsin. The last sequence represents the BTCI. BB1_SOYBN (*Glycine max*), BB3_DOLAX (*Dolichos axillaris*), 1BBI (*Glycine max*), 1PI2 (*Glycine max*), BB1_PHAAN (*Phaseolus angularis*), BB2_SOYBN (*Glycine max*), BB3_SOYBN (*Glycine max*), BB4_DOLAX (*Dolichos axillaris*), IBB_PHALU (*Phaseolus lunatus*), IBB_VIGUN (*Vigna unguiculata*).

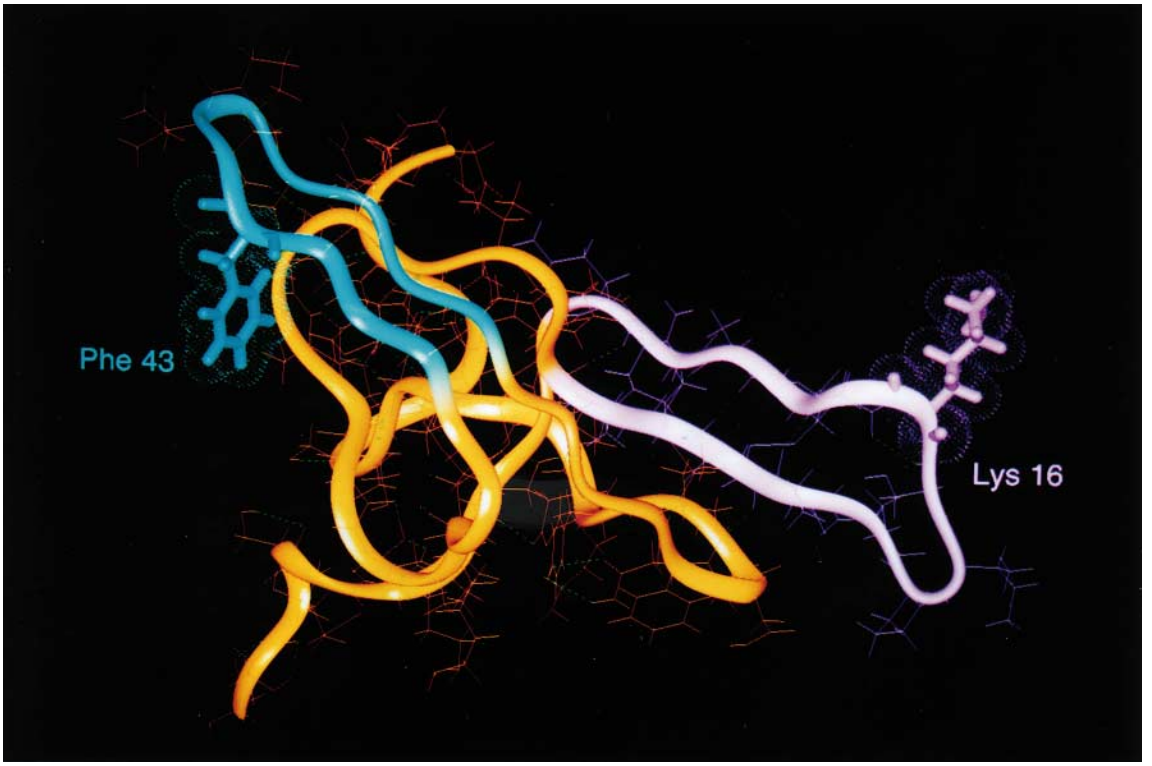

Fig. 2. Model of the 3-D structure of BTCI from *Vigna unguiculata*. The colours in the backbone ribbon indicate which parts of the structure are based on a large (blue or pink) or a small (yellow) number of observed NOEs in the 1BBI structure. No long distance NOEs were observed for residues 1–6 and 64–71; only a few long distance NOEs were observed in the inter domain region [9].

pins are probably determined with high reliability, but their relative orientation, and the residues in the core of the molecule are determined with less precision. In the crystal structure most of the hydrophobic residues that are surface exposed in the monomer are buried due to crystal contacts. The unusual amino acid distribution makes most tools commonly used for model or structure verification useless. We therefore resorted to extensive experimental verification of the models.

3.3. Experimental model verification

The secondary structure content and several aspects of the surface of BTCI have been examined spectroscopically. Fluorescence studies and spectroscopic solvent perturbation studies suggest that the aromatic residues in BTCI are indeed largely exposed to the solvent [25,26]. Experimental data on the accessibility of certain amino acid residues have been reported by Martin and Ventura [27]. They observed that the single tyrosine in BTCI has its ring in contact with the solvent, but the hydroxyl group buried in the core of the protein. Difference absorption spectrometry indicated that all three phenylalanines in BTCI are exposed to solvent [28]. All these characteristics are indeed observed in the BTCI model.

The anomalous distribution of hydrophobic and hydrophilic amino acids is thermodynamically not favourable if BTCI would be monomeric in solution. However, data obtained by light scattering (in solutions with pH 7.0 and ionic strength 0.25 at 25°C) [13] showed that the self-associating system of BTCI could be approximated by a monomer-hexamer equilibrium mechanism. Better approximation was obtained by using schemes of association with intermediate size oligomers: monomer-trimer-hexamer, monomer-dimer-trimer-hexamer and monomer-dimer-tetramer-hexamer. These three models all fit equally well to the experimental data. This suggests that BTCI molecules are in continuous equilibrium between monomers and several forms of multimers. The hydrophobic residues that are exposed to the solvent in the monomeric BTCI molecule can be buried in the multimers, in agreement with the self-association observed in the light scattering experiments. These data explain some of the more puzzling aspects of the BTCI model and the structures reported by Werner and Wemmer [9] and Chen et al. [8] and indicate that the anomalous distribution of hydrophobic and hydrophilic residues is indicative of a functional aspect of these molecules rather than of an error in the model or in the structure determination.

The model predicted in the present work also is in good agreement with available experimental data from circular dichroism and infrared spectroscopic studies on BTCI [29,15]. The far-UV CD spectrum of native BTCI exhibits a deep negative band at 201 nm and small positive ellipticity values at wavelengths above 218 nm. This is quite similar to polypeptides and proteins in constrained unordered conformation, including some protease inhibitors [30] and confirms that BTCI is a non-helical protein. The typical negative dichroic band at 201 nm represents mainly the contribution of disulphide groups [29]. This is in agreement with model compounds such as NN-diacetyl-L-cystine bis-methylamide that have a transition of the disulphide bond resulting in a negative CD band at 199 nm [31]. The reduction of the disulphide bonds in BTCI by dithioerythritol (DTE) resulted in a marked increase of the negative band and a blue-shift from 201 to 199

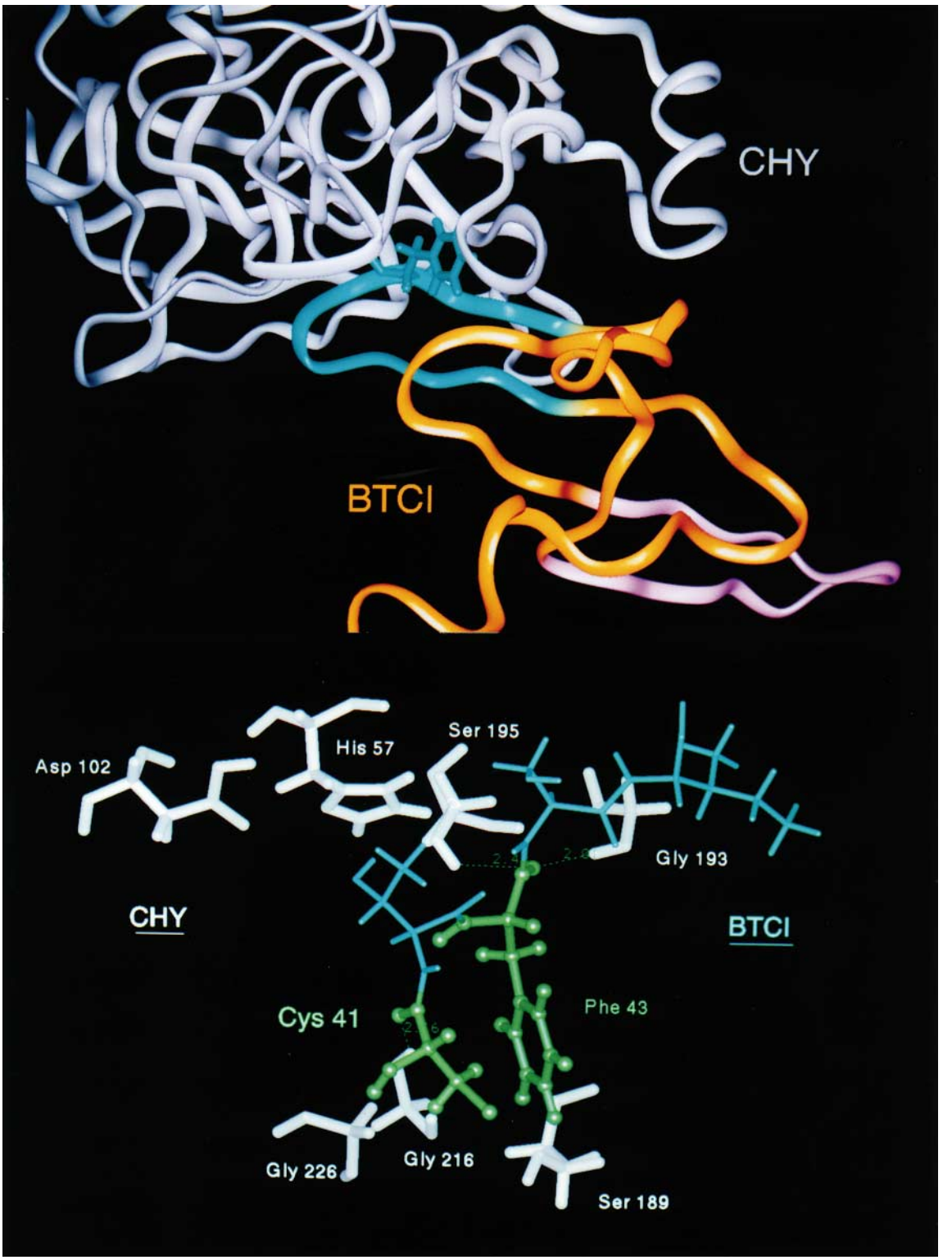

Fig. 3. BTCI-α-chymotrypsin complex. A: The BTCI backbone is shown as a blue, yellow and pink ribbon, the protease backbone is shown as a white ribbon. Phe-43 is depicted in ball and stick representation. Ribbon colours as in Fig. 2; white ribbon: chymotrypsin. B: Active site of α-chymotrypsin (white) with a bound inhibitor (green). This figure illustrates how Phe-43 in BTCI is oriented relative to the catalytic triad, the substrate specificity pocket, the oxyanion hole and the non-specific substrate binding region. Hydrogen bonds between BTCI and chymotrypsin are dashed.

nm. The position and the sign of the 199 nm peak correspond perfectly with the signal expected from the C-S-S-C dihedral angle [31]. The amide I–III and amide V regions in the IR spectrum of BTCI in a solid film suggest the occurrence of the unordered, antiparallel β-strand and β-turn structures [14]. The infrared spectrum in D_2O, shows prominent bands at

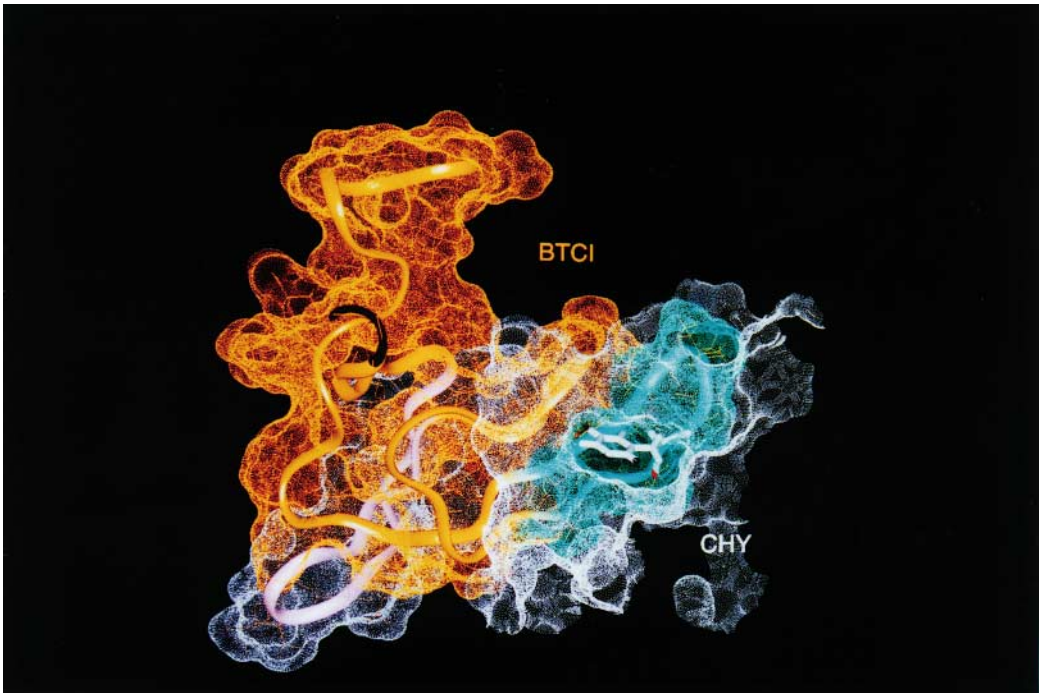

Fig. 4. Connolly surface of the BTCI:α-chymotrypsin complex. The inhibitor surface and ribbon colours as in Fig. 2 and protease surface is in white. A ball and stick model of the Phe-43 (white) is shown to aid with the identification of the active site pocket.

1643 cm^{-1} and 1632 cm^{-1}, indicative of unordered conformation (28%) and antiparallel β-strands (32%), respectively [15].

3.4. *BTCI-α-chymotrypsin complex*

One important goal of the present study is to provide a model for the interaction between α-chymotrypsin and BTCI. Fig. 3A shows the model of this complex. The catalytic triad, the chymotrypsin active site and the inhibition loop are indicated in Fig. 3B. Even after energy minimisation the OMTKY3 and BTCI inhibitor loops display a remarkably similarity (the backbone RMS displacement after optimal superpositioning is 0.48 Å). The conformation of the enzyme-inhibitor complex has a good steric arrangement for hydrogen bonding of interacting amino acids, as is shown in Fig. 3B. This apparent ease of the docking procedure strongly suggests that the inhibitor, the enzyme and the complex are modelled reasonably correctly.

3.5. *The active site*

The solvent-accessible surface [22] of the interface in the complex is presented in Fig. 4. The fit of Phe-43 in the α-chymotrypsin active site is remarkable. In the BTCI monomer the P1 residue (Phe-43) is solvent exposed. The conformation of the P3-P3′ loop agrees with the conformations that are proposed in several studies of enzyme-inhibitor complexes [9,32]. The primary recognition event occurs between the P1 residue of the inhibitor and the S1 pocket of the protease. Huber and Bode [33] have noted that the position and orientation of the Ser-195 Oγ, relative to the scissile peptide carbonyl, are consistent with a minimum in free energy. It has been demonstrated that the structure of the serine protease inhibitor complex may be consistent with a 'locked conformation' [34]. A scheme illustrating BTCI according to this model is shown in Fig. 4 and Fig. 3B. The convex hydrophobic surface of the β-hairpin make most of the contacts. In the substrate specificity pocket several direct hydrogen bonds are observed between the inhibitor and protease residues that surround the oxyanion hole. This includes the important serine in the S1 pocket (Ser-189), which does not interfere with the binding of inhibitor [34]. Phe-43 in BTCI fills the mainly hydrophobic specificity pocket allowing tetrahedral intermediate formation by hydrogen bond interaction at the oxyanion hole. Details of this hydrogen bonding network are indicated in Fig. 3B.

3.6. *The electrostatic potential*

The electrostatic potential of the BTCI-α-chymotrypsin complex can be used to understand some of characteristics of the complex and can shed further light on the mechanism of inhibition. Fig. 5 shows electrostatic potentials around the interface of the complex. GRASP picture shows peculiar complementarity of the localized electrostatic potential on active site of the enzyme and inhibitor. Complementarity seems to be more emphasized on the surface ring around the active site. The inhibitor and enzyme both have a net charge of +3, but they can nevertheless bind because of the presence of this localized charge complementarity.

4. Conclusions

Lacking an experimentally determined 3-D structure, we have built a model for BTCI, and a model for the active site area of α-chymotrypsin. This model has the same anomalous distribution of hydrophobic and hydrophilic residues as observed in the experimentally determined BBI structures; the core is rather hydrophilic and the surface is very hydrophobic. However, experimental data on the location of aromatic residues in BTCI made clear that this exposure of aromatic side chains to solvent only occurs in monomers, but BTCI is mainly observed as multimers in solution. In the BBI X-ray

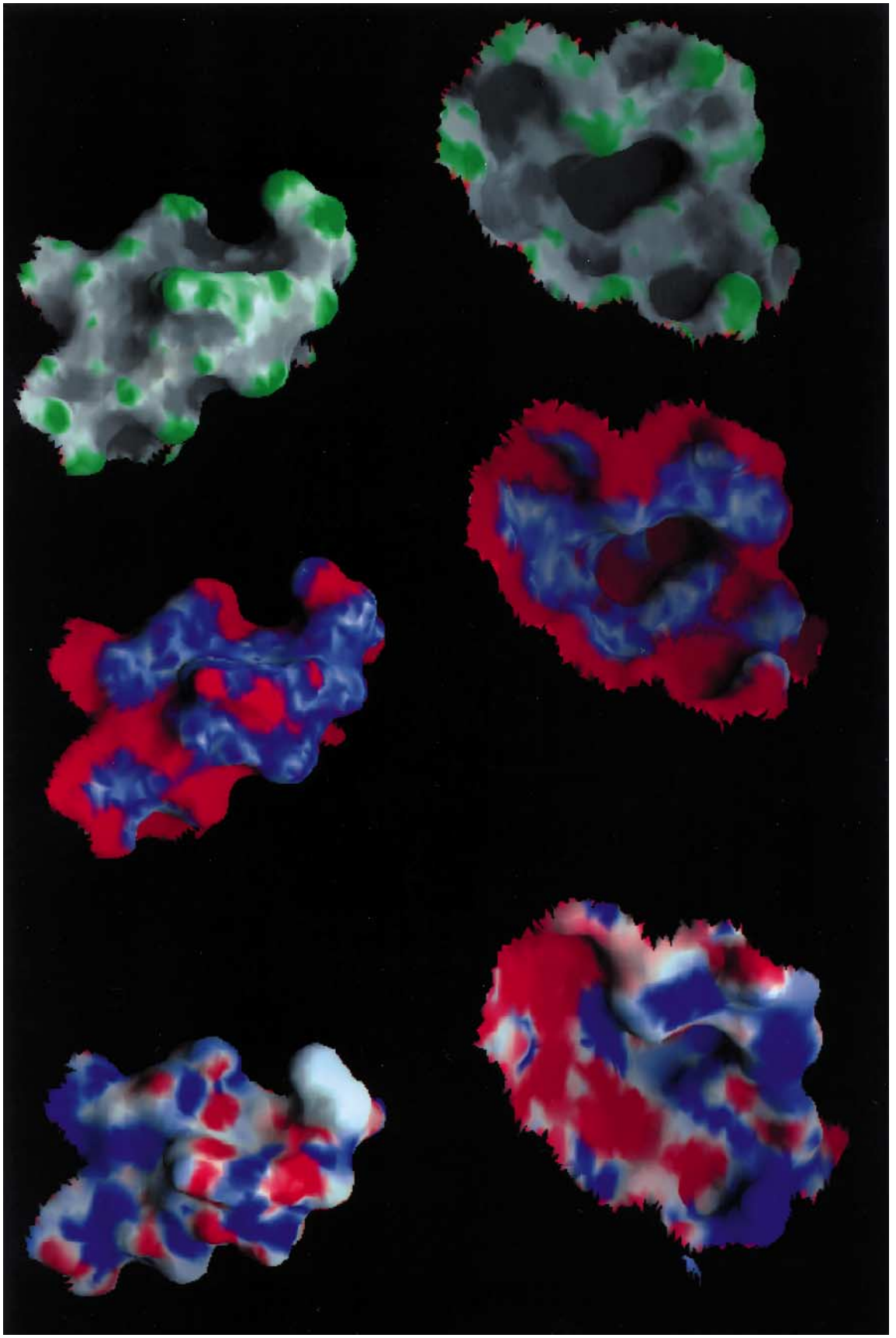

Fig. 5. Molecular interface of the BTCI (left) and chymotrypsin (right) in vicinity of the active site. On the top is presented surface curvature: convex vertices are green and flat to concave are white to gray colored, respectively. In the middle is presented the surface to surface distance map. White color represents more proximal interface distances, blue and red are color codes for increasingly distant surfaces. At the bottom is presented electrostatic surface potential. The surface is colored according to the local electrostatic potential, ranging from dark blue (most positive region) to deep red (most negative region).
←

structure most hydrophobic surface residues are involved in crystal contacts. The same effects have previously been determined by Millar et al. [35] and Harry and Steiner [36]. They reported that the surface of BBI is likely to be non-polar with a tendency to self-associate. Bode and Huber [37] noted that it is quite common to find non-polar residues with relatively high solvent accessibilities in many of the small serine protease inhibitor families.

Our models give insight into many structural aspects of the inhibitor and of the complex, including a good fit at the binding interface. The identification of the interaction surface in the complex, allows us to design serine protease inhibitors that can strengthen the plants defence potential against specific insects/parasites. The peculiar positive potential around the tip of the BTCI inhibition site interacts with the negative potential near the α-chymotrypsin active site. This explains how two molecules with net positive charges can bind to each other and it opens the possibility to alter the specificity of BBIs without the need to make mutations in the inhibitor loop.

The Bowman-Birk inhibitor will be synthesized and altered through recombinant DNA techniques in order to evaluate its potential as anticarcinogenic and seed protection agents. Moreover, the trypsin and chymotrypsin inhibitor genes from cowpea, and other legume seeds, are promising candidates for engineering plants resistant against arthropod parasites in an economically and environmentally sensible way.

Acknowledgements: We thank the Centro Nacional de Pesquisa de Recursos Genéticos e Biotecnologia - Empresa Brasileira de Pesquisa Agropecuária (CENARGEN/EMBRAPA) for the use of the computer graphics systems. We thank Dr. Chris Sander for hospitality and stimulating discussions.

References

[1] M. Laskowski Jr., I. Kato, Ann. Rev. Biochem. 49 (1980) 593–625.
[2] J. Xavier-Filho, M.M. Ventura, Comments Agric. Food Chemistry 1 (1988) 239–264.
[3] P. Correa, Cancer Res. 28 (41) (1981) 3685–3690.
[4] W. Troll, A.R. Kennedy, Cancer Res. 49 (1989) 499–502.
[5] A.R. Kennedy, B.F. Szuhaj, P.M. Newberne, P.C. Billings, Nutr. Cancer 19 (1993) 281–302.
[6] M.M. Ventura, J. Xavier Filho, An. Acad. Brasil. Ciênc. 38 (1966) 553–566.
[7] L. Morhy, M.M. Ventura, An. Acad. Brasil. Ciênc. 59 (1987) 71–81.
[8] P. Chen, J. Rose, R. Love, C.H. Wei, B.C. Wang, J. Biol. Chem. 267 (1992) 1990–1994.
[9] M.H. Werner, D.E. Wemmer, J. Mol. Biol. 225 (1992) 873–889.
[10] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, J. Mol. Biol. 112 (1977) 535–542.
[11] A. Susuki, Y. Tsunogae, I. Tanaka, T. Yamani, T. Ashida, S. Norioka, S. Hara, T. Ikenaka, J. Biochem. Tokyo 101 (1987) 267–274.
[12] T. Ikenaka, S. Norioka, In: Barret, A.J. and Salvesen, G., Eds. Bowman-Birk Family Serine Proteinase Inhibitors. Amsterdam: Elsevier, 1986:361–374.
[13] M.M. Ventura, K. Mizuta, H. Ikemoto, An. Acad. Brasil. Ciênc. 53 (1981) 195–201.
[14] J. Aragão, M.M. Ventura, An. Acad. Brasil. Ciênc. 58 (1986) 339–343.
[15] M. Sônia de F, Manuel M. Ventura, An. Acad. Brasil. Ciênc. 68 (1996) 165–174.
[16] G. Vriend, WHAT IF, J. Mol. Graph. 8 (1990) 52–55.
[17] G. Vriend, V.G.H. Eijsink, J. Comp. Aided Mol. Des. 7 (1993) 367–396.
[18] T.A. Jones, S. Thirup, EMBO J. 5 (1986) 819–822.
[19] R.A. Laskowski, M.W. MacArthur, J.M. Thornton, Programs to Check the Sterochemical Quality of Protein Structure. Oxford: Oxford Molecular Ltd., 1992:1–38.
[20] M. Fujinaga, A.R. Sielecki, R.J. Read, W. Ardelt, M. Laskowski Jr, N.G. James, J. Mol. Biol. 195 (1987) 397–418.
[21] H. Schrauber, F. Eisenhaber, P. Argos, J. Mol. Biol. 230 (1993) 592–612.
[22] M.L. Connolly, Science 28 (221) (1983) 709–713.
[23] A. Nicholls, K.A. Sharp, B. Honig, Prot. Struct. Funct. Genet. 11 (1991) 281–286.
[24] G. Vriend, C. Sander, J. Appl. Cryst. 26 (1993) 47–60.
[25] M.M. Ventura, J.B. Aragão, An. Acad. Brasil. Ciênc. 49 (1977) 329–336.
[26] M.M. Ventura, K. Mizuta, H. Ikemoto, An. Acad. Brasil. Ciênc. 56 (1984) 217–220.
[27] C.O. Martin, M.M. Ventura, An. Acad. Brasil. Ciênc. 58 (1986) 297–302.
[28] K. Mizuta, M.M. Ventura, An. Acad. Brasil. Ciênc. 55 (1983) 129–131.
[29] M.M. Ventura, J.B. Aragão, An. Acad. Brasil. Ciênc. 50 (1978) 587–596.
[30] B. Jirgensons, Biochym. Biophys. Acta 328 (1973) 314–322.
[31] D.L. Coleman, E.R. Blout, J. Am. Chem. Soc. 90 (1968) 2405–2415.
[32] D.S. Katz, D.W. Christianson, Prot. Eng. 6 (1993) 701–709.
[33] R. Huber, W. Bode, Acc. Chem. Res. 11 (1978) 114–122.
[34] R.J. Read, M.N.G. James, In: Barret, A.J. and Salvesen, G., Eds. Protease Inhibitors. Amsterdam: Elsevier, 1986:301–336.
[35] D.B.S. Millar, G.E. Willick, R.F. Steiner, V. Frattali, J. Biol. Chem. 244 (1969) 281–284.
[36] J.B. Harry, R.F. Steiner, Biochemistry 8 (1969) 5060–5064.
[37] W. Bode, R. Huber, Curr. Opin. Struct. Biol. 1 (1991) 45–52.