Crystal and molecular structure of the inhibitor eglin from leeches in complex with subtilisin Carlsberg

C.A. McPhalen, H.P. Schnebli* and M.N.G. James+

Medical Research Council of Canada in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada and **Ciba-Geigy AG, CH-4002 Basel, Switzerland*

Received 18 June 1985

The crystal structure of the molecular complex of eglin, a serine proteinase inhibitor from leeches, with subtilisin Carlsberg has been determined at 2.0 Å resolution by the molecular replacement method. The complex has been refined by restrained-parameter least-squares. The present crystallographic R factor ($=\Sigma||F_o|-|F_c||/\Sigma|F_o|$) is 0.183. Eglin is a member of the potato inhibitor 1 family, a group of serine proteinase inhibitors lacking disulfide bonds. Eglin shows strong structural homology to CI-2, a related inhibitor from barley seeds. The structure of subtilisin Carlsberg in this complex is very similar to the known structure of subtilisin *novo*, despite changes of 84 out of 274 amino acids.

Potato inhibitor 1 *Serine proteinase* *Molecular replacement method* Hirudo medicinalis *Crystallography*

1. INTRODUCTION

Amino-acid sequence homologies have established several families of serine proteinase inhibitors [1]. Members of 6 of these families have now been studied by X-ray crystallographic methods. Eglin, a small protein from the leech *Hirudo medicinalis*, is a member of the potato inhibitor 1 family [2,3]. Some of the proteins in this family lack the stabilizing disulfide bonds that are thought to contribute to the inhibitory effect of most other serine proteinase inhibitors. Thus, their 3-dimensional structures are of interest in elucidating those features contributing to their inhibitory properties. Eglin is a potent inhibitor of granulocytic elastase and cathepsin G [4] and may have important therapeutic medical applications. Here, we have used the genetically engineered product eglin-c rather than the naturally occurring protein from leeches.

Subtilisin Carlsberg is a bacterial serine proteinase from *Bacillus subtilis*. It is highly homologous to subtilisins *novo* and BPN', whose crystal structures are known [5,6], and has similar catalytic properties. Eglin-c forms a strong complex with subtilisin [4].

The structure of chymotrypsin inhibitor 2 (CI-2) from barley seeds in complex with subtilisin *novo* has been determined and refined at 2.1 Å resolution to a current crystallographic R factor ($R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$ where F_o and F_c are the observed and calculated structure factors, respectively) of 0.193 (McPhalen and James, unpublished). CI-2 is another member of the potato inhibitor 1 family and is homologous to eglin [3].

We present here the 3-dimensional X-ray crystallographic structure of the eglin-c:subtilisin Carlsberg complex, refined to an R factor of 0.183 for data in the resolution range 6.0–2.0 Å, and compare the structure with that of the CI-2:subtilisin *novo* complex.

2. MATERIALS AND METHODS

2.1. *Crystallization*

Purified lyophilized eglin-c (batch no.84, 25-4) was obtained from Ciba-Geigy, Switzerland. Subtilisin Carlsberg was the kind gift of Dr I. Svend-

+ To whom reprint requests should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

sen, Carlsberg Laboratories, Copenhagen. Eglin-c has the same sequence as native eglin and consists of 70 amino acid residues (fig.1) [7]; M_r = 7990. Subtilisin Carlsberg consists of 274 residues [8]; M_r = 27240. Crystals of the complex were grown by the hanging-drop vapor-diffusion method from a solution of 0.7 M KH_2PO_4 and H_3PO_4, buffered to pH 5.6. The crystals have unit cell dimensions of a = 38.31(3) Å, b = 41.41(4) Å, c = 56.50(6) Å, α = 69.5(1)°, β = 83.7(2)°, and γ = 75.3(1)°. The space group is P1, with one molecule of the complex per unit cell.

2.2. *Structure determination and refinement*

X-ray intensity data were collected on an automated diffractometer and processed by standard crystallographic techniques [9]. The method of molecular replacement was used to solve the phase problem for the structure of the eglin-c : subtilisin Carlsberg complex. The search model was the partially refined subtilisin *novo* molecule. The rotation search was performed with the fast rotation function [10]. The eglin-c portion of the complex was recognized in the first map but the molecular structure was not completed until several cycles of refinement and model building were done. The first 7 residues in eglin-c are disordered. The MMS-X interactive graphics system [11] with the macromolecular modelling system M3, developed by Broughton and coworkers [12], was used for map interpretation and model fitting.

Refinement of the model was performed with the restrained-parameter least-squares refinement program of Hendrickson and Konnert [13], modified by Furey et al. [14], and locally by M. Fujinaga for the FPS164 attached processor. After 19 cycles of refinement the R factor was reduced from 0.301 (6.0–2.8 Å resolution) to 0.183 (6.0–2.0 Å resolution). The root-mean-square (rms) deviations from expected stereochemical parameters for cycle 19 are: 0.041 Å for the 2492 covalent bond distances; 0.062 Å for the 3403 interbond angle distances; 0.018 Å from the 432 planar groups of the complex; and 3.2° for the 337 peptide bond torsional angles. No water molecules have been included in the structural model.

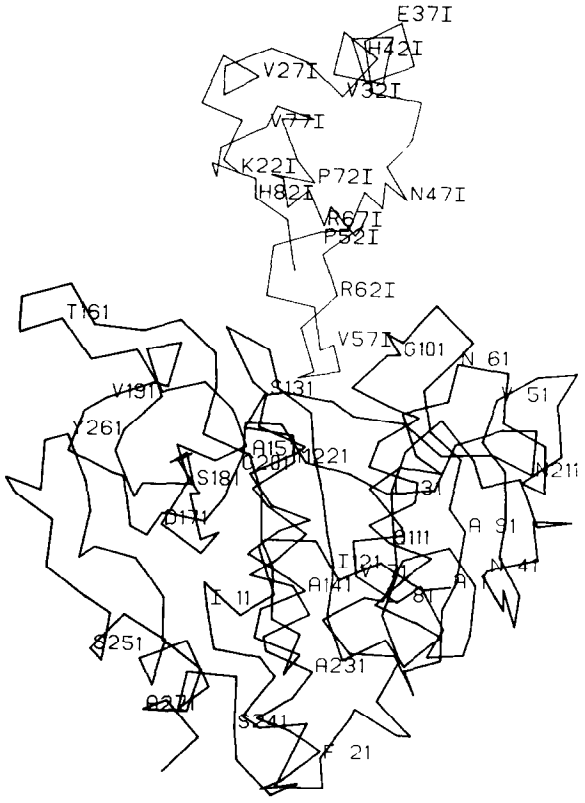

Fig.1. An α-carbon backbone drawing of eglin-c (thin lines) complexed to subtilisin Carlsberg (thick lines). Every 5th amino acid is labelled in the inhibitor, every 10th in the enzyme. An I follows the sequence number of inhibitor residues to distinguish them from those of the enzyme.

3. RESULTS AND DISCUSSION

3.1. *Structure of eglin-c*

The eglin-c molecule is a wedge-shaped disk with the reactive site loop at the narrow end of the wedge. This loop fits into the active site cleft of subtilisin (fig.1). CI-2 and the ovomucoid inhibitors [9] are similarly shaped, although the polypeptide chain foldings of these 2 families are quite different. Eglin-c has 3 major secondary structural elements (fig.2): an α-helix of 3.6 turns, Thr 31I–Tyr 43I; a 4-stranded mixed parallel and antiparallel β-sheet, Lys 22I–Phe 24I, Asn 47I–Leu 51I, Arg 65I–Tyr 70I, and Val 75I–Gly 83I; and a wide loop crossover connection between parallel strands 3 and 4 of the sheet, Pro 56I–Leu 61I, containing the reactive site, Leu 59I–Asp 60I. Reverse turns occur in the molecule at Phe 24I–Val 27I, Val 27I–Lys 30I, Tyr 43I–Ala

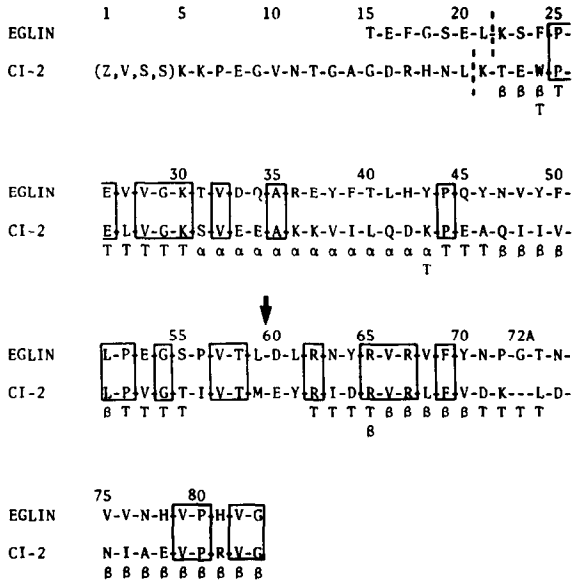

Fig.2. Amino acid sequences of eglin [7] and CI-2 [3], aligned according to 3-dimensional structure homology. Sequence numbering is that of CI-2. The reactive site bond is indicated by the arrow. The 22 identical residues are boxed. Residues before the vertical dashed lines are not seen in the electron density maps of the 2 enzyme:inhibitor complexes. Secondary structural elements: α, α-helix; β, β-sheet; T, reverse turn.

46I, Pro 52I–Thr 55I, Arg 62I–Arg 65I, and Asn 71I–Thr 73I. The α-helix packs against the characteristic curvature of the twisted β-sheet and the interface between these elements is the hydrophobic core of eglin-c.

A least-squares superposition of the 62 structurally equivalent α-carbon atoms of eglin-c and CI-2 (program of W. Bennett) gives an rms deviation of 1.67 Å between the 2 molecules. The largest deviations are at the beginning of the reactive site loop (residues 55I–56I), and near the end of the α-helix (residues 41I–42I) (fig.3). Despite these differences, the mode of binding to their cognate enzymes appears very similar for the 2 inhibitors. This may indicate small adjustments of enzyme and/or inhibitor conformation to achieve an 'induced fit'. The conformation of the inhibitor reactive site loop of both eglin-C and CI-2 similar to that of the Kazal inhibitors [9], although the folding of the remainer of the polypeptide chain is unrelated for the 2 families.

3.2. *Structure of subtilisin Carlsberg*

There are 84 amino acid changes and one deletion between the sequences of subtilisin *novo* and subtilisin Carlsberg [5,8]. Most of the sequence changes occur on exterior loops of the molecule. A least-squares superposition of the 274 structurally equivalent α-carbon atoms of the 2 subtilisins (program of W. Bennett) gives an rms deviation of 0.53 Å. This indicates that the 2 enzymes are extraordinarily similar.

ACKNOWLEDGEMENTS

Koto Hayakawa grew the crystals used in this work. We thank Randy Read and Masao Fujinaga for helpful discussions. C.A.M. is the holder of an

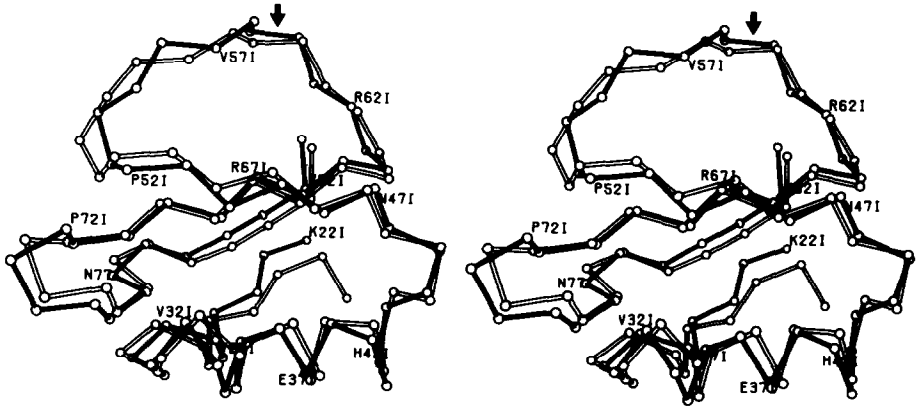

Fig.3. Superposition of α-carbon drawings of eglin-c (filled connections) and CI-2 (open connections). The reactive site bond, Leu 59I–Asp 60I, is indicated by the arrow.

Alberta Heritage Foundation for Medical Research Studentship. This work was funded by grants to the MRC of Canada Group in Protein Structure and Function at the University of Alberta from the Medical Research Council of Canada.

REFERENCES

[1] Laskowski, M. jr and Kato, I. (1980) Annu. Rev. Biochem. 49, 593–626.
[2] Melville, J.C. and Ryan, C.A. (1972) J. Biol. Chem. 247, 3445–3453.
[3] Svendsen, I., Boisen, S. and Hejgaard, J. (1982) Carlsberg Res. Commun. 47, 45–53.
[4] Seemüller, U., Fritz, H. and Eulitz, M. (1981) Methods Enzymol. 80, 804–816.
[5] Wright, C.S., Alden, R.A. and Kraut, J. (1969) Nature 221, 235–242.
[6] Drenth, J., Hol, W.G.J., Jansonius, J.N. and Koekoek, R. (1972) Eur. J. Biochem. 26, 177–181.
[7] Seemüller, U., Eulitz, M., Fritz, H. and Strobl, A. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1841–1846.
[8] Smith, E.L., Markland, F.S., Kasper, C.B., DeLange, R.J., Landon, M. and Evans, W.H. (1966) J. Biol. Chem. 241, 5974–5976.
[9] Read, R.J., Fujinaga, M., Sielecki, A.R. and James, M.N.G. (1983) Biochemistry 22, 4420–4433.
[10] Crowther, R.A. (1973) in: The Molecular Replacement Method (Rossmann, M.G. ed.) International Science Review 13, 173–178, Gordon and Breach, New York.
[11] Barry, C.D., Molnar, C.E. and Rosenberger, F.U. (1976) Technical Memo no.229, Computer Systems Lab., Washington University, St.Louis, MO.
[12] Sielecki, A.R., James, M.N.G. and Broughton, C.G. (1982) in: Crystallographic Computing, Proceedings of the International Summer School (Sayre, D. ed.) pp.409–419, Carleton University, Ottawa, Oxford University Press, Oxford.
[13] Hendrickson, W.A. and Konnert, J.A. (1980) in: Biomolecular Structure, Function, Conformation and Evolution (Srinivasan, R. ed.) vol. I, pp. 43–57, Pergamon, Oxford.
[14] Furey, W. jr, Wang, B.C. and Sax, M. (1982) J. Appl. Crystallogr. 15, 160–166.