FEBS Letters 351 (1994) 397–400

FEBS 14506

Sequence and structure similarities of cathepsin B from the parasite *Schistosoma mansoni* and human liver

Mo-Quen Klinkert[a,*], Donato Cioli[a], Elliott Shaw[b], Vito Turk[c], Wolfram Bode[d], Richard Butler[a]

[a]*Institute of Cell Biology, Consiglio Nazionale delle Ricerche, Rome, Italy*
[b]*Friedrich-Miescher Institute, Basel, Switzerland*
[c]*Jozef Stefan Institute, Ljubljana, Slovenia*
[d]*Max Planck Institute fur Biochemie, Martinsried, Germany*

Received 29 July 1994; revised version received 5 August 1994

Abstract A three-dimensional structure of *Schistosoma mansoni* cathepsin B was modelled using the coordinates of the crystal structure of the human liver enzyme. Both enzymes appear to share remarkable structural similarity. However, an examination of the models complexed with two synthetic inhibitors revealed differences in inhibitor binding, as confirmed by differences in the 50% inhibitory concentration of the same inhibitor.

Key words: Cathepsin B; *Schistosoma mansoni*; Human liver; Structural comparison; Inhibitor modelling

1. Introduction

Cathepsin B, a widely studied cysteine proteinase, is known to function in the degradation of tissue proteins [1]. It has also been implicated in a number of pathological conditions such as tumour metastasis [2] and dystrophy [3], as well as in the processing of protein precursors [4]. The molecular structure of the human liver cathepsin B has been determined by X-ray analysis at 2.15 Å resolution [5]. The deduced amino acid sequence corresponding to a *Schistosoma mansoni* protein of 31 kDa (termed Sm31) exhibits nearly 70% homology to cathepsin B from human liver [6]. Taking advantage of the sequence similarity, a 3-dimensional model was constructed. The schistosome proteinase is believed to be one of several digestive enzymes [7] and thus, may be of critical importance to the life cycle of the parasite. Therefore, we analyzed more closely the structure of the two cathepsins for a better understanding of enzyme–inhibitor interactions, in order to aid in the design of parasite-specific inhibitors as novel chemotherapeutic agents.

2. Materials and methods

Determination of inhibitory activities. The inhibitors used were a peptidyldiazomethane (Z-Trp-MetCHN$_2$) [8] and a E-64 derivative, compound *N*-(L-3-trans-carboxyoxirane-2-carbonyl)-L-isoleucyl-L-proline (CA-074) [9]. The latter was a kind gift of Taisho Pharmaceutical Co. Ltd., Japan. The inhibitory activities are shown as their 50% inhibitory concentrations (IC_{50}), as carried out by Murata et al [9]. Activity was assayed at pH 6.0 using benzyloxycarbonyl-phenylalanyl-arginyl-7-amino-4-methylcoumarin (Z-Phe-Arg-AMC) as a substrate (5 μM final concentration) and Sm31 or human liver cathepsin B (150 pM final concentration), as described by Barrett and Kirschke [1]. Each IC_{50} value is the mean of at least three experiments. K_m values as calculated under our assay conditions with Sm31 and human liver cathepsin B are 33.0 ± 7.1 (S.D.) μM and 112 ± 39 (S.D.) μM, respectively.

Tertiary structure modelling. The structure of the schistosome Sm31 protease was modelled directly from the coordinates of the human cathepsin B enzyme [5]. Appropriate amino-acid replacements, insertions and deletions were performed using the programs Biopolymer and Builder (Biosym Technologies) followed by successive stages of energy-minimization, using the conjugate gradients method and the CVFF forcefield of the program Discover (Biosym Technologies). The final minimization was to 0.01 kcal/Å maximum derivative.

The binding of the diazomethane inhibitor Z-Trp-MetCHN$_2$ [8] to the two proteases was modelled by copying the Z-Phe-methylenyl-Ala inhibitor from the structure of papain [10] of the 6PAD entry in the Protein Data Bank at Brookhaven National Laboratory [11,12] to the human or schistosome protease. The inhibitor was then converted to Z-Trp-MetCHN$_2$ as above with energy-minimization of the inhibitor portion at each step. Finally the inhibitor and all atoms within 10 Å of the inhibitor, including a 5 Å water layer, were energy-minimized to a final maximum derivative of 0.01 kcal/Å. The CA-074 inhibitor [9] was 'grown' stepwise onto Cys^{29} of each protein again minimizing the energy of the inhibitor portion at each stage. A final energy-minimization step was performed as for the diazomethane inhibitor.

3. Results and discussion

Fig. 1 shows an alignment of the amino acids of cathepsin B from the human liver and from *Schistosoma mansoni*. 148 of the 252 amino acids are identical, while an additional 25 are conservative changes. The numbering of amino acids follows that introduced for human cathepsin B [5]. Thus Met^{66} of the human enzyme (shown as superscript) is deleted in the schistosome and a lysine (shown as subscript and numbered 124 A in Fig. 1) is inserted between positions 124 and 125. Both changes occur on the surface of the molecule. Even though it is difficult to predict the exact conformation due to the change, we can assume that the insertion has little or no effect on the binding site. In the human, the dipeptide Ala^{48}-His^{49} is excised during processing of the enzyme to yield one light and one heavy chain. In agreement with the fact that a double-chain form has never been detected in the schistosome [13], we find two amino acids, Gly and Lys, at positions 48 and 49, respectively. Furthermore, the last two C-terminal amino acids of the human enzyme are absent in the schistosome enzyme, for which there is no evidence of carboxyterminal processing, and termination of the protein is characterized by an authentic stop codon in the cDNA [6]. The overall result of the amino acid substitutions is an increase in the net charge from −3 in the human enzyme to +12 in the schistosome. These new positive charges appear to be fairly randomly distributed throughout the molecule, but with a particular concentration in one section corresponding to residues 122–137.

*Corresponding author.

SSDI 0014-5793(94)00893-0

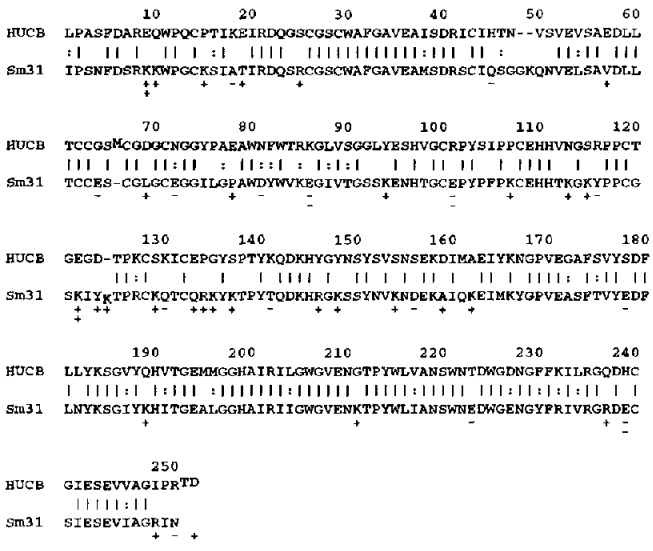

Fig. 1. Comparison of the Sm31 protease from *Schistosoma mansoni* and human liver cathepsin B. Errors in the deduced amino acid sequence of Sm31 as published in [6] are corrected here. + and − underneath the alignment indicate where an amino acid change causes a charge difference. All His residues are considered as protonated at the physiological pH 5.

The 2.15 Å crystal structure of the human enzyme reveals a 30 Å thick disc-shaped molecule, the diameter of which is 50 Å. In between the left-hand and right-hand domains, on top of the molecule is a V-shaped incision where the active-site cleft is found. In the center of the cleft is the active Cys^{29} and the rear of the cleft is closed by a loop of peptide containing two His residues (110 and 111). Because of the extensive sequence identities, we were able to model the entire molecule using the X-ray coordinates of the human enzyme. Fig. 2 shows a backbone trace stereo diagram of the schistosome enzyme. Like the crystalline human cathepsin B, the main features are three helical segments and the 'occluding' loop in the L-domain, and six extended strands forming an open barrel-like structure in the R-domain. Residues 122–137 where many of the new positively charged residues are concentrated, corespond to an external loop on the protein surface originating in the left rim of the binding cleft. There are no changes in the positions of the cysteines in the sequence, so it is expected that the six disulphide pairs in both molecules are the same. Thus, there are no significant changes in the model structure of the schistosome cathepsin B compared with the crystal structure of the human proteinase. The structural imposition of the schistosome and human enzymes is shown in Fig. 3.

Since we were interested in using the model to make predictions about peptide inhibitors, we first focussed our attention on comparing amino acids comprising the active sites. One part of the left-hand side of the active site cleft is formed by the segment Cys^{119}-Glu^{122}. A Lys at 122 in the schistosome instead of Glu makes this part positively charged. The amino acids contributing to the rest of the left rim Gly^{24}-Trp^{30} and Asn^{72}-Tyr^{75} are fairly conserved with respect to the human enzyme. Noteworthy are three further changes, namely that of Gly^{24} to Ser, Ser^{25} to Arg and Asn^{72} to Glu. Overall, the presence of Lys and Arg renders this region more hydrophilic than the human counterpart. Next, we examined how inhibitors interacted by modelling the synthetic diazomethane inhibitor Z-Trp-MetCHN$_2$ into the active sites of the two enzymes (Fig. 4A,B). The amino acid residues of the inhibitor Met and Trp are found on the S1 and S2 subsites, respectively. The P1 Met residue was observed to project out of the cleft in the general direction of Glu^{122} in the human enzyme, and appears to occupy the same position in the schistosome enzyme, even though Glu^{122} is replaced by Lys.

The right-hand side of the cleft is formed by Phe^{180}-Leu^{182} and Met^{196}-His^{199}. While the change from Leu^{182} to Asn creates a more hydrophilic environment in the schistosome, the Met^{196} to Leu change makes no significant contribution. The amino acids forming the floor of the cleft are Trp^{221}, Gln^{23}, His^{199}, Ala^{200}, Ala^{173} and Cys^{29}, all of which are conserved, except for a change from Ala^{173} to Ser, which represents part of the basement of the S2 depression. Consequently, the region is more hydrophilic due to the presence of the hydroxyl group in Ser. In fact, this part of the cleft appears wider, where Ile replaces Tyr^{75} and Ser replaces Ala^{173}. Hence, the P2 Trp residue of the inhibitor lies sideways in the cleft and is able to form a hydrogen bond between the indole N of the Trp and the OH of Ser^{173}. Such a bond is likely to contribute to the stability of the enzyme-inhibitor interaction (Fig. 4A). On the other hand, the cleft of the human enzyme is comparatively narrow with the Trp residue being sandwiched vertically between Tyr^{75} on the left rim and Ala^{173} and Ala^{200} on the floor (Fig. 4B). Experimentally, Z-Trp-MetCHN$_2$ was found to exhibit stronger inhibitory activity for the schistosome enzyme (IC_{50} value of 2 ± 0.2 nM) than for the human enzyme (33 ± 6 nM).

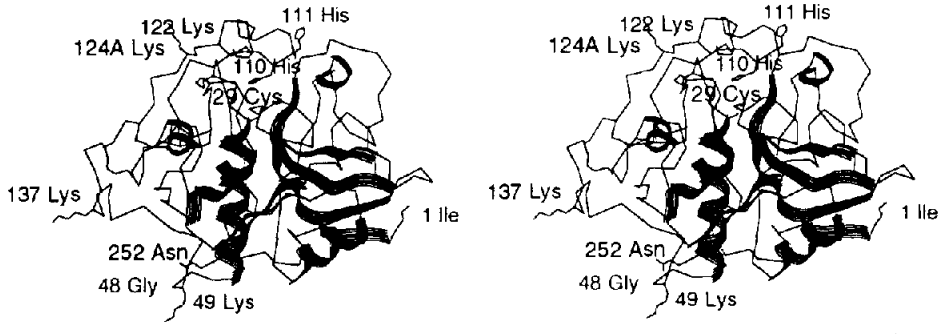

Fig. 2. A backbone trace stereo diagram of *S. mansoni* Sm31 polypeptide chain. Some of the amino acids residues are given to assist in the orientation of the molecule, which follows that shown for human cathepsin B [5]. The active site cleft (V-shaped) is on 'top' of the molecule and the 'occluding loop' at His^{110} and His^{111} is 'behind' the active site. The L and R domains are on the left- and right-hand side, respectively. Disulfide connections are as observed in the human enzyme and are indicated by lines drawn perpendicularly to the main strand.

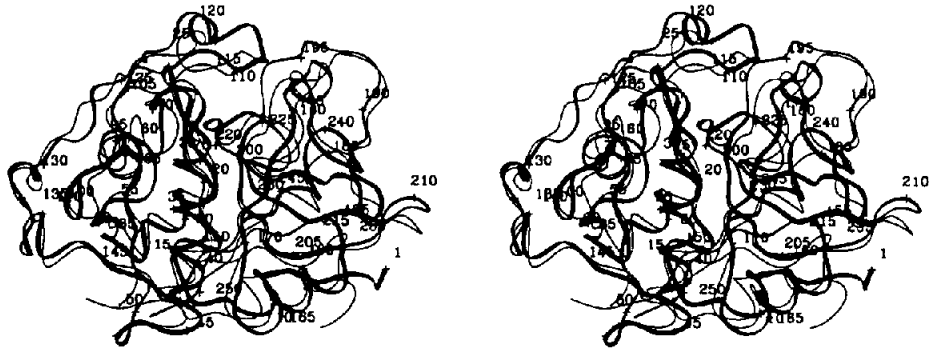

Fig. 3. Carbon atom structure of schistosome Sm31 (in heavier lines) superimposed with human cathepsin B. The view is identical to that in Fig. 2.

The second inhibitor CA-074 was observed to occupy a different position in the active site cleft in the human enzyme, such that the amino acid residues of the inhibitor Ile and Pro are found in the S1′ and S2′ positions, respectively (D. Turk, pers. comm.). Thus, the P1′ side chain of Ile projects towards the right rim and is surrounded by the side chains of His^{199}, Leu^{181}, Leu^{196} and Trp^{221}, while the P2′ side chain of Pro points towards the left rim (Fig. 4C). The schistosome cleft appears to be narrow, probably caused by the change in the side chain positions from Asn^{72} to Glu and from Met^{196} to Leu. Thus the inhibitor appears to be more constricted and lies extended along the floor and towards the back of the cleft. In comparison, CA-074 in the human enzyme is found in a wider cleft than in the schistosome enzyme and contact surface area between inhibitor and human enzyme is greater, thus probably assisting its binding (Fig. 4D). This could be a possible explanation as to why CA-074 is more inhibitory for the human enzyme (IC_{50} of 0.43 ± 0.11 nM) than the schistosome enzyme (2.3 ± 0.6 nM).

The present report has outlined the remarkable overall simi-

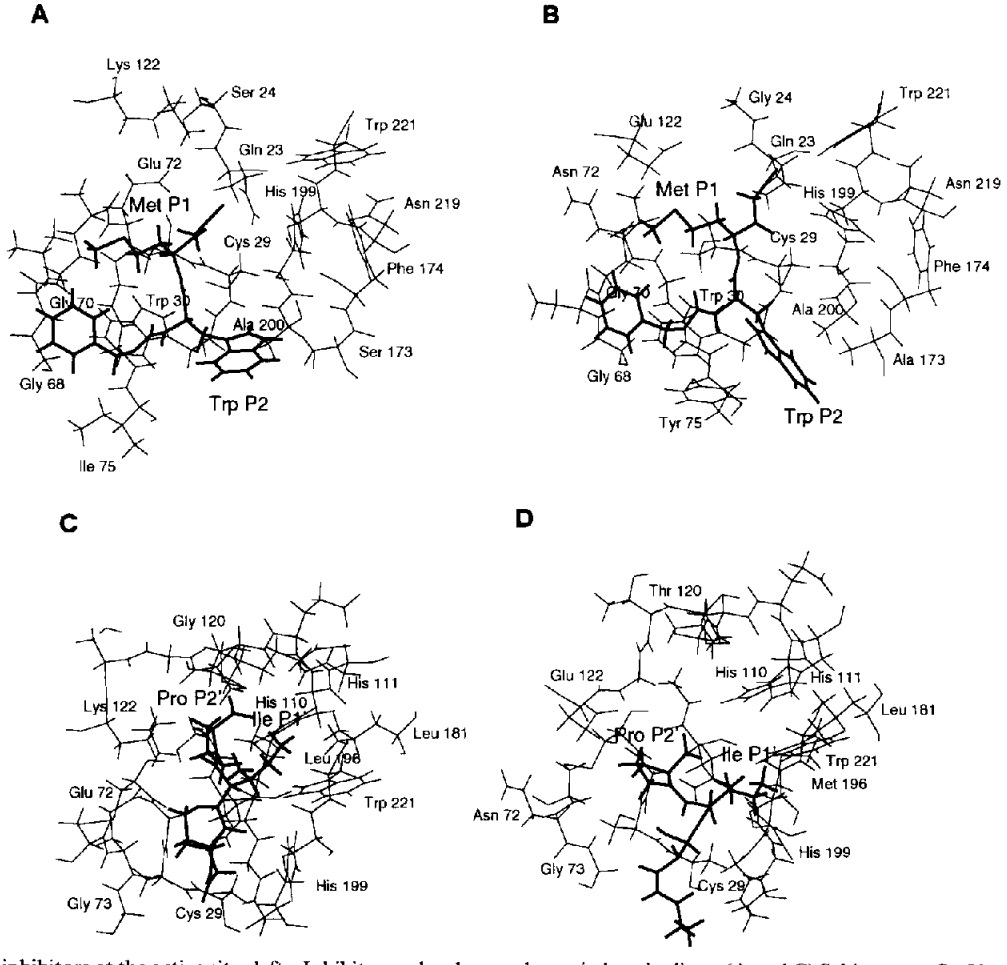

Fig. 4. Models of inhibitors at the active site clefts. Inhibitor molecules are shown in heavier lines. (A and C) Schistosome Sm31 and (B and D) human cathepsin B are complexed with Z-Trp-MetCHN$_2$ and CA-074, respectively. Differences in binding of inhibitors to the two enzymes are discussed in the text.

larity between the human and the schistosome cathepsin B, in terms of both primary and three-dimensional structures. Conflicting with this finding is the detection of a strong and very early immune response against the parasite enzyme in *S. mansoni*-infected humans [14]. A possible explanation for this observation is the involvement of regions of considerable difference, such as the external loop of residue 122–137, but it could also be due to the lysosomal confinement of the human enzyme as opposed to the location of parasite cathepsin B in the intestinal lumen of schistosomes, such that it is continuously regurgitated into the host bloodstream.

In spite of their overall similarities, we have shown that the active sites of the two proteins present distinct differences. We have also proposed that these structural differences may account for the experimentally observed divergence in susceptibility to protease inhibitors. Further combined crystallographic and biochemical studies will be useful in the design of parasite-specific inhibitors and possibly in the elucidation of its role(s) in the schistosome.

Acknowledgments: We would like to thank Taisho Pharmaceutical Co. Ltd. Japan for its kind gift of CA-074. This work received the support of the C.M.T. – Italian Ministry for Foreign Affairs – D.G.C.S.

References

[1] Barrett, A.J. and Kirschke, K. (1981) Methods Enzymol. 80, 535–561.
[2] Sloane, B.F. and Honn, K.V. (1984) Cancer Metastasis Rev. 3, 249–263.
[3] Ii, K., Hizawa, K., Nonaka, I., Sugita, H., Kominawi, E. and Katanuma, N. (1986) Am. J. Pathol. 122, 193–198.
[4] Chan, S.J., San Segundo, B., McCormick, M.B. and Steiner, D.F. (1986) Proc. Natl. Acad. Sci. USA 83, 7721–7725.
[5] Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katanuma, N. and Bode, W. (1991) EMBO J. 10, 2321–2330.
[6] Klinkert, M.Q., Felleisen, R., Link, G., Ruppel, A. and Beck, E. (1989) Mol. Biochem. Parasitol. 33, 113–122.
[7] Timms, A.R. and Bueding, E. (1959) Brit. J. Pharmacol. Chem. 14, 68–73.
[8] Shaw, E., Mohanty, S., Colic, A., Stoka, V. and Turk, V. (1993) FEBS Lett. 334, 340–342.
[9] Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T. and Katanuma, N. (1991) FEBS Letts. 280, 307–310.
[10] Drenth, J., Kalk, K.H. and Swen, H.M.(1976) Biochemistry 15, 3731–3738.
[11] Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542.
[12] Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. and Weng, J. (1987) in: Crystallographic Databases – Information Content, Software Systems, Scientific Applications (Allen, F.H., Bergerhoff, G. and Sievers, R. Eds.) pp. 107–132, Data Commission of the Int'l Union of Crystallography, Bonn/Cambridge/Chester.
[13] Felleisen, R. and Klinkert, M.Q. (1990) EMBO J. 9, 371–377.
[14] Ruppel, A., Rother, U., Vongerichten, U., Lucius, R. and Diesfeld, H.J. (1985) Exp. Parasitol. 60, 195–206.