Crystal structure of NS-134 in complex with bovine cathepsin B: a two-headed epoxysuccinyl inhibitor extends along the entire active-site cleft

Igor ŠTERN*, Norbert SCHASCHKE†, Luis MORODER† and Dušan TURK*¹

*Department of Biochemistry and Molecular Biology, Josef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia, and †Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, D-82152 Martinsried, Germany

The crystal structure of the inhibitor NS-134 in complex with bovine cathepsin B reveals that functional groups attached to both sides of the epoxysuccinyl reactive group bind to the part of active-site cleft as predicted. The -Leu-Pro-OH side binds to the primed binding sites interacting with the His¹¹⁰ and His¹¹¹ residues with its C-terminal carboxy group, whereas the -Leu-Gly-Meu (-Leu-Gly-Gly-OMe) part (Meu, methoxycarbonylmethyl) binds along the non-primed binding sites. Comparison with the propeptide structures of cathepsins revealed that the binding of the latter part

is least similar to the procatepsin B structure; this result, together with the two-residue shift in positioning of the Leu-Gly-Gly part, suggests that the propeptide structures of the cognate enzymes may not be the best starting point for the design of reverse binding inhibitors.

Key words: cathepsin B, cysteine protease, epoxysuccinyl inhibitor, inhibitor design, NS-134.

INTRODUCTION

Cathepsin B (EC 3.4.22.1) is an intracellular lysosomal cysteine protease primarily exhibiting carboxypeptidyl activity. It is an abundant and ubiquitously expressed cathepsin. Together with other cathepsins, it is involved in the degradation and processing of proteins, including the regulation of prohormone and proenzyme activation, antigen processing, inflammatory responses against antigens, metabolism, tissue remodelling and apoptosis [1,2]. Cathepsin B from the granule-derived cell surface has been suggested to provide self-protection for degranulating cytotoxic lymphocytes [3]. There is evidence for its involvement in the activation of a plasminogen activator [4].

Changes in concentration and activity of cathepsin B were observed in aging [5] and under various pathophysiological conditions including rheumatoid arthritis [6], osteoarthritis [7], tumours and metastasis [8,9]. Cathepsin B also stimulates angiogenesis in cancer tissues and degrades and inactivates the inhibitors of metalloproteases [10]. Furthermore, up-regulated extracellular activity and relocalization of cathepsin B to the cell surface and into the extracellular matrix was observed in malignant tissues, suggesting its role in tumour invasion [11]. Cathepsin B and its zymogen form are secreted, together with trypsinogen and active trypsin, into the pancreatic juice of patients with sporadic pancreatitis or hereditary pancreatitis [12].

The unique exopeptidyl mechanism of cathepsin B soon called for its structural explanation [13]. The crystal structure confirmed the papain fold of the protease and revealed the position of an

insertion specific for cathepsin B (amino acid sequence 110–126 using cathepsin B nomenclature), termed the occluding loop, which endows the enzyme with its carboxypeptidyl activity. It is positioned within the active-site cleft above the primed binding sites. At the bottom of the loop, there are two positively charged residues His¹¹⁰ and His¹¹¹, which are responsible for the binding of the free C-terminal carboxy group of a substrate [13,14]. Without knowing the structure of the enzyme, cathepsin B-selective inhibitors CA030 [EtO-(2S,3S)-tEps-Ile-Pro-OH] and CA074 [nPrHN-(2S,3S)-tEps-Ile-Pro-OH] were developed [15,16], which were derived from E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane], a non-specific inhibitor of papain-like cysteine proteases (Figure 1) [17]. The crystal structure of the CA030–cathepsin B complex demonstrated that CA030 binds with its peptidyl moiety in the primed binding sites in the same direction as a substrate [14], in contrast with the peptidyl moiety of E-64 [18–20], which binds to the non-primed binding sites in the direction opposite to substrate binding and in the same manner as the cathepsin B propeptide [21,22]. The comparison of E-64 and CA030-binding geometries pointed out the internal symmetry of the *trans*-epoxysuccinyl group [14], which contains two terminal carbonyl groups, thus allowing polypeptide chains to be appended at both the carbonyl junctions. Consequently, double-headed inhibitors of cathepsins B [23,24], L, S and K [25] were synthesized, and the structure of the cathepsin L inhibitor, *N*-{L-*trans*-[2-(pyridin-2-yl)ethylcarbamoyl]-oxirane-2-carbonyl]-L-Phe-dimethylamide (referred to as CLIK148), in complex with papain was determined [26].

Abbreviations used: BANA, α -N-benzoyl-DL-arginine β -naphthylamide hydrochloride; CA030, (S)-1-[(S)-2-[[[(2S,3S)-3-ethoxycarbonyl-oxiranecarbonyl]-amino]-3-methyl-pentanoyl]-pyrrolidine-2-carboxylic acid or EtO-(2S,3S)-tEps-Ile-Pro-OH; CA074, (S)-1-[(S)-3-methyl-2-[[[(2S,3S)-3-propylcarbamoyl-oxiranecarbonyl]-amino]-pentanoyl]-pyrrolidine-2-carboxylic acid or nPrHN-(2S,3S)-tEps-Ile-Pro-OH; CLIK148, *N*-{L-*trans*-[2-(pyridin-2-yl)ethylcarbamoyl]-oxirane-2-carbonyl]-L-Phe-dimethylamide; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; MCA, (7-methoxycoumarin-4-yl)-acetyl; Meu, methoxycarbonylmethyl, MeO-Gly; NS-134, (S)-1-[(S)-2-[[[(2S,3S)-3-[(S)-1-[(methoxycarbonylmethyl-carbamoyl)-methyl]-carbamoyl]-3-methyl-butylcarbamoyl]-oxiranecarbonyl]-amino]-4-methyl-pentanoyl]-pyrrolidine-2-carboxylic acid or MeO-Gly-Gly-Leu-(2S,3S)-tEps-Leu-Pro-OH; NS-137, (S)-1-[(S)-2-[[[(2R,3R)-3-[(S)-1-[(methoxycarbonylmethyl-carbamoyl)-methyl]-carbamoyl]-3-methyl-butylcarbamoyl]-oxiranecarbonyl]-amino]-4-methyl-pentanoyl]-pyrrolidine-2-carboxylic acid or MeO-Gly-Gly-Leu-(2R,3R)-tEps-Leu-Pro-OH; pNA, *p*-nitroanilide; Sp, sulphopropyl; Z-, benzyl-oxycarbonyl.

¹ To whom correspondence should be addressed (e-mail dusan.turk@ijs.si).

² The nucleotide sequence data reported have been deposited in the Brookhaven Protein Data Bank with accession code 1SP4.

Figure 1 Schematic representation of the epoxysuccinyl inhibitors E-64, CA030, CA074, NS-134 and CLIK148

In the present study, we present the crystal structure of NS-134 in complex with bovine cathepsin B. The well-resolved structure of the cathepsin B inhibitor bound along the whole active-site cleft clearly demonstrates the correctness of the double-headed inhibitor concept and its potential for achieving higher inhibitor specificity in a drug design process and in studies clarifying the biological roles of papain-like cathepsins.

MATERIALS AND METHODS

Materials

DEAE–Sephadex A-50, Sp (sulphopropyl)-Sephacrose Fast Flow and activated thiol Sepharose were obtained from Amersham Biosciences (Uppsala, Sweden). BANA (α -N-benzoyl-DL-arginine β -naphthylamide hydrochloride) was obtained from Sigma. Z-Arg-Arg-MCA and Z-Phe-Arg-pNA [where Z- stands for benzyl-oxycarbonyl, MCA for (7-methoxycoumarin-4-yl)-acetyl and pNA for *p*-nitroanilide] were from Bachem (Bubendorf, Switzerland). NS-134 was synthesized as described previously [24].

Methods

Enzyme isolation and purification

Cathepsin B was isolated from bovine kidneys by the method described in [27]. In brief, isolation included the following: homogenization 1.5:1 (v/v) [0.1 M sodium acetate/1 mM EDTA/

10% (v/v) Triton X-100/0.3 M NaCl, pH 5.5], acid activation and incubation (pH 4.2; for 2 h at 37 °C), fractional precipitation with ammonium sulphate (0–75% saturation), anion-exchange chromatography on DEAE–Sephadex A-50 (20 mM Bis-Tris and 1 mM EDTA, pH 6.25; elution: gradient NaCl 0–0.3 M), covalent chromatography on activated thiol Sepharose (elution buffer: 0.1 M phosphate buffer/1 mM EDTA/0.3 M NaCl/20 mM cysteine, pH 6.0) and cation-exchange chromatography on Sp-Sephacrose Fast Flow (20 mM sodium acetate and 1 mM EDTA, pH 5.2; elution: gradient NaCl 0–0.25 M).

Enzyme activity assays

Enzyme activity against BANA and Z-Arg-Arg-MCA was tested after each purification step. The assay buffer was 0.1 M phosphate/1.5 mM EDTA (pH 6.0). Assays were performed at 37 °C. Assay conditions for BANA: 100 μ l of diluted samples (usually 1:20) was mixed with 250 μ l of 5 mM cysteine. After a 5 min incubation, 10 μ l of BANA was added. The reaction was blocked after 10 min. Absorption at 520 nm was measured on a UV–visible spectrophotometer (PerkinElmer). Assay conditions for Z-Arg-Arg-MCA: the samples were diluted to 1:1000 and 250 μ l of 5 mM cysteine was added. After 5 min, a 20 μ M solution of Z-Arg-Arg-MCA was added and incubated for another 10 min. The reaction was blocked by the addition of 2 ml of 1 mM iodoacetic acid. Enzyme activity was calculated by excitation at 370 nm and measurement of fluorescence emission at 460 nm [28].

Purity and concentration determinations

Purity of cathepsin B was determined by SDS/PAGE on 8–25% gradient gels (PhastSystem; Amersham Biosciences). Positions of protein and standards were determined by Coomassie Brilliant Blue staining. The concentration was determined spectrophotometrically by measuring the absorbance at 280 nm. An absorption coefficient of 2.2 ml $\cdot$ cm⁻¹ $\cdot$ mg⁻¹ was used, as determined from the primary amino acid structure [29].

Active-site titration

Cathepsin B activity (500 nM) was determined by stoichiometric active-site titration with NS-134 (0–750 nM) in the presence of a chromogenic substrate, Z-Phe-Arg-pNA (100 mM) [28]. The concentration of active cathepsin B was determined from the plot of average velocities against molar ratios of initial enzyme and inhibitor concentrations (I_0/E_0) by the method described in [30].

Preparation of the cathepsin B–NS-134 complex

The purified enzyme, stored at –18 °C, was unfrozen, dialysed against 50 mM sodium acetate/acetic acid and 1 mM EDTA (pH 5.5) and incubated with 10-fold molar excess of the inhibitor in the presence of 20 mM cysteine at 37 °C until the residual activity on Z-Arg-Arg-MCA was zero. The complex solution was dialysed against 10 mM phosphate buffer (pH 6.0), concentrated to 20 mg/ml and stored at 4 °C.

Crystallization and data collection

Crystals were grown using the same conditions as described for the bovine cathepsin B–CA074 complex [31]: the reservoir (1 ml) contained 50 mM sodium citrate buffer (pH 3.5) with 2 M sodium dihydrogen phosphate. The droplet was prepared by mixing 2 μ l of the precipitation buffer with 2 μ l of the sample. Crystals with approximate dimensions 2 $\times$ 2 $\times$ 2 mm appeared within a week. Diffraction data were collected from a hexagonal crystal

Table 1 A summary of the crystallographic data and refinement statistics

(a)	
Data collection	
Space group	P4 ₃ 2 ₁ 2
Cell parameters	$a = b = 73.1 \text{ \AA}$, $c = 142.0 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$
Molecules in asymmetric unit	1
Limiting resolution ($\text{\AA}$)	2.2
Measured reflections	410 575
Independent reflections	20 320
R_{sym} (10–2.2 $\text{\AA}$)	0.177
R_{sym} (2.28–2.2 $\text{\AA}$)	0.473
Completeness	96 %
Completeness (2.28–2.2 $\text{\AA}$)	77 %
No. of solvent molecules	176
(b)	
Refinement	
Resolution range in refinement ($\text{\AA}$)	10–2.2
Reflections used in refinement	19 372
R factor	0.194
RMSD of bond distances	0.0135
RMSD of bond angles	1.635
B factors overall (bond RMSD)	21.8 (2.15)

mounted on a quartz capillary tube (Glas, Schönwalde, Germany) at room temperature (22 °C) using in-house equipment (Cu $K\alpha$ radiation on Rigaku Ru200 rotating anode as the X-ray source and a 345 mm MAR Research image plate detector). Diffraction data were processed and scaled with the HKL programs package [32]. The crystal diffracted to 2.2 $\text{\AA}$ resolution ($1 \text{ \AA} = 10^{-10} \text{ m}$) and belonged to the space group P4₃2₁2 with unit cell dimensions of $a = b = 73.1 \text{ \AA}$ and $c = 142.0 \text{ \AA}$.

Structure determination and refinement

The position of the cathepsin B molecule was determined by the molecular replacement method implemented in the EPMR program [33], using the human cathepsin B structure (1HUC; [13]) as the search model. The best solution with a correlation factor of 0.575 and $R = 0.433$ was found using data in the 15–4 $\text{\AA}$ resolution range. Immediately, the initial $F_o - F_c$ electron density map clearly indicated the positioning of the inhibitor molecule. In subsequent cycles, which involved model rebuilding, addition of solvent molecules and positional and B value refinement with MAIN [34], resolution of the data used was gradually expanded to the final range (10–2.2 $\text{\AA}$). Solvent molecules were generated by the automated procedure in MAIN and then corrected manually. The R value of the final model was 0.19. The final model was validated with MAIN and PROCHECK [35]. The crystallographic data and refinement statistics are summarized in Table 1.

RESULTS AND DISCUSSION

Overall structure

The asymmetric unit of the tetragonal crystal structure contains a single molecule of NS-134 bound to bovine cathepsin B. The cathepsin B molecule consists of 253 residues. Its chain is broken into two parts: the light chain (from Leu¹ to Gly⁴⁸) and the heavy chain (from Arg⁴⁹ to Thr²⁵³). The molecule has seven

Figure 2 Structure of the complex of cathepsin B with NS-134

The cathepsin B molecule is presented in cyan (sticks) and the inhibitor NS-134 in yellow. Oxygen and nitrogen atoms are shown in red and blue respectively. Hydrogen bonds are shown in white dotted lines. Amino acid residues of the cathepsin B molecule around the active-site cleft that interact with the substrate are represented by a ball-and-stick model in green. The Figure was prepared with MAIN [34] and rendered with Raster 3D [46]. The molecular surface was generated with GRASP [47].

disulphide bridges: Cys¹⁴–Cys⁴³, Cys²⁶–Cys⁷¹, Cys⁶²–Cys¹²⁸, Cys⁶³–Cys⁶⁷, Cys¹⁰⁰–Cys¹³², Cys¹⁰⁸–Cys¹¹⁹ and Cys¹⁴⁸–Cys²⁵². The last disulphide bond is unique to bovine cathepsin B. The residues Asn⁴⁷, Gly⁴⁸ and Arg⁴⁹ located at the cleavage site are not well resolved by the electron density maps.

The backbone of the bovine cathepsin B structure is almost identical with the crystal structures of the free human cathepsin B [13], human cathepsin B in complex with CA030 [14] and bovine cathepsin B in complex with CA074 [36] with RMSD = 0.35, 0.35 and 0.16 $\text{\AA}$, calculated for 251, 249 and 240 pairs of $C\alpha$ atoms respectively. The comparisons show that there seems to be no region showing dissimilarity above the statistical noise level, suggesting that, in the cathepsin B structure, there are no changes induced on binding of these small inhibitors.

Binding of inhibitor

The inhibitor NS-134 occupies the active-site cleft from S4 to S2' substrate-binding sites and is covalently linked to the catalytic Cys²⁹ [C2(Epo 4)-S γ (Cys²⁹)] (Figure 2). Numbering of the inhibitor residues is according to the substrate-binding nomenclature [37], although the polarity of the peptide chain is reversed in the non-primed region. The NS-134 chain is composed of Meu 4N (Gly-OMe, as assigned by PDB), Gly 3N, Leu 2N, Epo 1N, Leu 1P and Pro 2P residues and thus has two C-termini (Figure 2) due to the symmetric amide linkages to the central *trans*-epoxysuccinyl group. Here, letters N and P are used to indicate the binding to the primed (P) and the non-primed (N) binding sites of cathepsin B. The carboxy group of the terminal glycine residue is modified into a methyl ester (Meu 4N), whereas the carboxy group of Pro 2P is free. The binding geometry of the residues

constituting the inhibitor NS-134, apart from Gly 3N, is unambiguously defined by the electron density maps. The density at Gly 3N is broadened, indicating multiple conformations, which could, however, not be refined due to the limiting resolution of the diffraction data. Interestingly, the methyl group attached to a glycine residue of Meu 4N is resolved only by the density maps calculated with untruncated data, whereas the density maps truncated to a 10 Å resolution reveal no corresponding electron density.

Binding to the primed binding sites

The Pro 2P is positioned within the S2'-binding site. On the left-hand side of the active-site cleft, the side-chain moiety of the proline residue is packed against the wall formed by the occluding loop residues (Figure 2), whereas on the top right-hand side, the C-terminal group of the proline residue is orientated towards the occluding loop residues His¹¹⁰ and His¹¹¹. The position of each oxygen atom of the carboxyl group is defined by hydrogen bonds, one formed with the ND1 atom of His¹¹⁰ 2.6 Å distant and the other with the NE2 atom of His¹¹¹ 2.7 Å distant. The solvent molecule W47 additionally fixes the position of the carboxy group of Pro 2P by forming a hydrogen bond with the OT2 atom. The solvent molecule W47 is a member of the group of solvent molecules that form a hydration shell around the His¹¹¹ ring. The Leu 1P residue is positioned in the S1'-binding site. In contrast with the S2'-binding site, the S1' site is a pocket formed by hydrophobic side chains of Val¹⁷⁶, Phe¹⁷⁴, Val¹⁷⁶, Phe¹⁸⁰, Met¹⁹⁶ and Trp²²¹ with the imidazole ring of the reactive site residue His¹⁹⁹ at its bottom. The Leu 1P side chain fits well into the shallow pocket of the S1'-binding site. The main-chain carbonyl group of Leu 1P points towards the amide of the Trp²²¹ side chain 3.0 Å distant, forming a hydrogen bond, whereas the amide points towards the carbonyl group of the Gly¹⁹⁸, 3.8 Å distant (Figure 2).

As shown in Figure 3, the position of the Leu-Pro part of NS-134 corresponds to the Ile-Pro part of CA030 [14] and thus of CA074. For 21 structurally equivalent atoms of the epoxysuccinyl moieties and Ile-Pro parts of the first CA030 structure and Leu-Pro of NS-134, the RMSD was 0.55 Å, with a maximal distance of 1.36 Å between the C γ atoms of the proline rings due to different ring puckers. When NS-134 was compared with the second structure of the CA030–cathepsin B complex, it was found that the pucker is the same in both and RMSD = 0.36 Å. Thus, the change of the amino acid residue at position 1' from isoleucine to leucine affects neither the positioning of the C-terminal proline residue nor the equivalent atoms at the binding to the S1-binding site, whereas differences in the proline ring puckers indicate flexibility within the proline ring.

Epoxysuccinyl residue binding

The Epo 1N residue occupies the S1-binding site. During the catalytic reaction, a covalent bond is formed between the C-2 atom of the epoxysuccinyl group and the S γ atom of Cys²⁹ by opening of the epoxy ring. Owing to the Walden inversion at the C-2 atom, the absolute configuration at C-2 changes from *S* to *R*. When comparing the CA030 and NS-134 structures, it became evident that the positions of the C-2 and C-3 atoms differ due to the difference in chirality of both carbon atoms; although the chirality of the two atoms in bound NS-134 is (*R,R*), it was modelled as (*S,S*) in CA030 [this indicates that the (*R,R*) diastereomer of CA030 was selected from the diastereomeric mixture of CA030 present in solution].

The first epoxysuccinyl carbonyl group fills the oxyanion hole, where it forms two hydrogen bonds with the main-chain amide of

Figure 3 Active-site cleft of cathepsin B with NS-134 and the superimposed inhibitors CA030 and E-64

The surface of cathepsin B molecule is presented in white. NS-134 is presented in dark grey, CA030 (1CSB; [14]) is shown in grey and E-64 (1ATK; [39]) in light grey. The Figure was prepared with MAIN [34] and rendered with Raster 3D [46]. The molecular surface was generated with GRASP [47].

Cys²⁹ and the side-chain amide group of Gln²³. The hydroxy group resulting from the epoxy ring opening points in the direction of the side-chain moiety of the S1-binding site, leaving most of it free. The second carbonyl group (O2 atom) is orientated towards the nitrogen atom of Gly⁷⁴ and the edge of the aromatic ring of Trp³⁰. The distance of 3.10 Å indicates the presence of a hydrogen bond between the carbonyl and Gly⁷⁴ amide.

Stereochemistry of the epoxysuccinyl moiety obviously plays an important role in inhibitory activity. Kinetic studies indicated that NS-134 ($k_2/K_1 = 1\,520\,000\text{ M}^{-1} \cdot \text{s}^{-1}$), which is a (2*S*,3*S*) diastereomer, exhibits approx. seven times faster binding when compared with NS-137 ($k_2/K_1 = 214\,600\text{ M}^{-1} \cdot \text{s}^{-1}$), which is a (2*R*,3*R*) diastereomer [24]. Crystals of the complex of NS-137 with the bovine cathepsin B could not be grown. This indicates that NS-137 binds to bovine cathepsin B in a way that disrupts the crystallization surface. Considering that NS-137 differs from NS-134 only in the chirality of the two epoxysuccinyl chiral carbons C-2 and C-3, the difference in chirality imposes a different positioning of the NS-137 Gly-Meu tail when compared with NS-134. Packing of bovine cathepsin B molecules in the crystal is indeed tight. A crystallographic 2-fold axis is positioned in the vicinity of the S2-binding site, so that a tail with a slightly different position would have collided with a neighbouring symmetric image of itself.

Binding beyond the S2-binding site

In contrast with the much shorter CA030 and CA074, which end with a short aliphatic tail, NS-134 extends from the epoxysuccinyl fragment into the non-primed binding sites with the -Leu-Gly-Meu sequence (Figure 3). The S2-binding site is a deep hydrophobic pocket in the R domain. The side chain of Leu 2N does

Figure 4 Structure of the cathepsin B–NS-134 complex and a closer view of the active-site cleft, showing the superimposed propeptide part (3PBH; [22])

NS-134 is represented as a ball-and-stick model in dark grey. The propeptide part is shown in light grey. The Figure was prepared with MAIN [34] and rendered with Raster 3D [46]. The molecular surface was generated with GRASP [47].

not fill it entirely. On one side of it, there is the negatively charged Glu²⁴⁵, a distinctive feature of cathepsin B. To explore the use of electrostatic interactions within the S2 pocket, Agm-Orn-(2*S*,3*S*)-*t*Eps-Leu-Pro-OH and its (2*R*,3*R*) variant were synthesized [23]. It turned out that the (2*S*,3*S*) diastereomer ($k_2/K_i = 197\,000\text{ M}^{-1}\cdot\text{s}^{-1}$) binds to the enzyme about three times slower when compared with the CA030-like [EtO-(2*R*,3*R*)-*t*Eps-Leu-Pro-OH] ($k_2/K_i = 567\,000\text{ M}^{-1}\cdot\text{s}^{-1}$) and is approx. 30 times less selective towards cathepsin L [23]. The lower selectivity indicates that the binding of the ornithine group to the S2 pocket was not as expected. To mimic a substrate residue binding, a simple chain reversal is obviously not sufficient. Apparently, the side chain requires a different positioning; the changed chirality from L to D of an amino acid residue binding to the S2 site appears to be a promising step.

The Gly 3N of NS-134 crosses the S3 binding area, and it does not reach the pocket and form any recognizable electrostatic interactions with the enzyme. The terminal residue Meu 4N is orientated upwards and lies above the active-site cleft of the protein. It is not involved in interactions with the cathepsin B molecule to which NS-134 is covalently attached; however, in the crystal, it is packed against a symmetry-related neighbouring molecule of cathepsin B.

Inhibitor construct validation

The idea of the -Leu-Gly-Meu segment composition originates from the Leu46P, Gly47P, Gly48P part of the propeptide sequence [24]. However, comparison of the NS-134 and procathepsin B

Figure 5 Cathepsin B molecule (white surface) from the cathepsin B–NS-134 complex (the inhibitor is represented as a ball-and-stick model in dark grey), with the superimposed propeptide parts from procatepsins X (1DEU; [42]), B (3PBH; [22]), L (1CJL; [40]) and K (1BY8; [41]) shown in dark grey, light grey and white respectively

The Figure was prepared with MAIN [34] and rendered with Raster 3D [46]. The molecular surface was generated with GRASP [47].

structures revealed that the structurally equipositioned part of the propeptide corresponds to the Thr44P, Phe45P, Leu46P sequence, which precedes this sequence by two residues [21,22]. In addition to the two-residue shift, a closer look revealed that the chain traces differ as well. The positions of C α atoms of Leu 2N and Thr44P are approximately one bond length apart (1.4 Å), whereas the positions of the further two C α atom pairs, Gly 3N–Phe 45P and Meu 4N Gly–Leu47, are 2.9 and 4.2 Å apart respectively. Although the positional shift of the Leu 2N residue by one bond length can be explained by the different number of atoms along the main chain of the epoxysuccinyl fragment (four atoms) and the amino acid residue (three atoms) binding to the S1-binding site, it is clear that the rest of the tail does not follow the path of the cathepsin B propeptide residues (Figure 4).

When superimposing the structures of E-64 in complex with actinidin (1AEC; [38]) and cathepsin K (1ATK; [39]), the complex of CLIK148 with papain (1CVZ; [26]) and propeptides of cathepsins L (1CJL; [40]), K (1BY8; [41]) and X (1DEU; [42]) on the structure of NS-134 in complex with cathepsin B, it became evident that the Gly-Meu (Gly-Gly-OMe) tail follows the chain trace of these structures. In contrast with expectations, these comparisons suggest that future design of this part of the double-headed epoxysuccinyl inhibitors of cathepsin B should be based on the structures of proenzymes of cathepsins L, K and X but not of cathepsin B (Figure 5).

The use of the flexible Gly-Gly tail has been shown to be successful for the attachment of labels for *in vivo* studies [3,43]. The β -cyclodextrin–epoxysuccinyl peptide conjugate can be used as a carrier system for cytotoxic drugs to target cathepsin B in tumour cells [44]. The flexibility of the NS-134 Gly-Gly tail does not appear to be crucial for the binding of the inhibitor con-

structs to a cathepsin surface; however, this does suggest that other amino acid residues can be used for the interactions with S3 and S4 binding areas. Although increased binding affinity and binding selectivity are not required for the design of cathepsin B inhibitors, these positions may be exploited for targeting other papain-like cathepsins.

Conclusions

The presented structure demonstrates that inhibitors based on the thiol-reactive epoxysuccinyl group can span the entire active-site cleft of a papain-like cathepsin. A combination of components responsible for the selectivity of papain-like cathepsins with a flexible tail on the opposite side of the epoxysuccinyl moiety, to which labels are attached, makes this construct a promising tool for the identification of papain-like cathepsins in biological samples. It remains to be determined whether the double-headed epoxysuccinyl construct can compete with urea-derived cathepsin K inhibitors [45] in a drug discovery process.

We thank G. Gunčar for help in preparing the Figures. T. Mather is acknowledged for a critical reading of this paper. This work was supported by the Slovenian Ministry of Education, Science and Sport.

REFERENCES

- Turk, B., Turk, D. and Turk, V. (2000) Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta* **1477**, 98–111.
- Turk, B., Stoka, V., Rozman-Pungerčar, J., Cirman, T., Droga-Mazovec, G., Orešič, K. and Turk, V. (2002) Apoptotic pathways: involvement of lysosomal proteases. *Biol. Chem.* **383**, 1035–1044.
- Balaji, K. N., Schaschke, N., Machleidt, W., Calfamo, M. and Henkert, P. A. (2002) Surface cathepsin B protects cytotoxic lymphocytes from self-destruction after degranulation. *J. Exp. Med.* **196**, 493–503.
- Aisa, M. C., Beccari, T., Costanzi, E. and Maggio, D. (2003) Cathepsin B in osteoblasts. *Biochim. Biophys. Acta* **1621**, 149–159.
- Keppler, D., Walter, R., Perez, C. and Sierra, F. (2000) Increased expression of mature cathepsin B in aging rats. *Cell Tissue Res.* **320**, 181–188.
- Hashimoto, Y., Kakegawa, H., Narita, Y., Hachiya, Y., Hayakawa, T., Kos, J., Turk, V. and Katunuma, N. (2001) Significance of cathepsin B accumulation in synovial fluid of rheumatoid arthritis. *Biochem. Biophys. Res. Commun.* **283**, 334–339.
- Lang, A., Horler, D. and Baici, A. (2000) The relative importance of cysteine peptidases in osteoarthritis. *J. Rheumatol.* **27**, 1970–1979.
- Koblinski, J. E., Ahram, M. and Sloane, B. F. (2000) Unraveling the role of proteases in cancer. *Clin. Chim. Acta* **291**, 113–135.
- Kepgoriski, I. and Sloane, B. F. (2003) Cathepsin B and its role(s) in cancer progression. *Biochem. Soc. Symp.* **70**, 263–276.
- Kostoulas, G., Lang, A., Nagase, H. and Baici, A. (1999) Stimulation of angiogenesis through cathepsin B inactivation of the tissue inhibitors matrix metalloproteinases. *FEBS Lett.* **455**, 286–290.
- Cavallo-Medved, D. and Sloane, B. F. (2003) Cell surface cathepsin B: understanding its functional significance. *Curr. Top. Dev. Biol.* **54**, 313–341.
- Kukor, Z., Mayerle, J., Krüger, B., Tóth, M., Steed, P. M., Halangk, W., Lerch, M. M. and Sahin-Tóth, M. (2002) Presence of cathepsin B in the human pancreatic secretory pathway and its role in trypsinogen activation during hereditary pancreatitis. *J. Biol. Chem.* **277**, 21389–21396.
- Musil, D., Zucic, D., Engh, R. A., Mayr, I., Huber, R., Popovič, T., Turk, V., Towatari, T. and Katunuma, N. (1991) The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* **10**, 2321–2330.
- Turk, D., Podobnik, M., Popovič, T., Katunuma, N., Bode, W., Huber, R. and Turk, V. (1995) Crystal structure of cathepsin B inhibited with CA030 at 2.0-Å resolution: a basis for the design of specific epoxysuccinyl inhibitors. *Biochemistry* **34**, 4791–4797.
- Towatari, T., Nikawa, T., Murata, M., Yokoo, C., Tamai, M., Hanada, K. and Katunuma, N. (1991) Novel epoxysuccinyl peptides. A selective inhibitor of cathepsin B, *in vivo*. *FEBS Lett.* **280**, 311–315.
- Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T. and Katunuma, N. (1991) Novel epoxysuccinyl peptides. Selective inhibitors of cathepsin B, *in vitro*. *FEBS Lett.* **280**, 307–310.
- Hanada, K., Tamai, M., Yamagishi, M., Ohmura, S., Sawada, J. and Tanaka, I. (1978) Isolation and characterization of E-64, a new thiol protease inhibitor. *Agric. Biol. Chem.* **42**, 523–528.
- Varughese, K. I., Ahmed, F. R., Carey, P. R., Hasnain, S., Huber, C. P. and Storer, A. C. (1989) Crystal structure of a papain–E-64 complex. *Biochemistry* **28**, 1330–1332.
- Yamamoto, D., Matsumoto, K., Ohishi, H., Ishida, T., Inoue, M., Kitamura, K. and Mizuno, H. (1991) Refined X-ray structure of papain–E-64-c complex at 2.1 Å resolution. *J. Biol. Chem.* **266**, 14771–14777.
- Kim, M. J., Yamamoto, D., Matsumoto, K., Inoue, M., Ishida, T., Mizuno, H., Sumiya, S. and Kitamura, K. (1992) Crystal structure of papain–E-64-c complex. *Biochem. J.* **287**, 797–803.
- Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A. C. and Mort, J. S. (1996) Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure* **4**, 405–416.
- Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M. and Turk, V. (1996) Crystal structures of human procathepsin B at 3.2 and 3.3 Å resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Lett.* **384**, 211–214.
- Schaschke, N., Assfalg-Machleidt, I., Machleidt, W., Turk, D. and Moroder, L. (1997) E-64 analogues as inhibitors of cathepsin B. On the role of the absolute configuration of the epoxysuccinyl group. *Bioorg. Med. Chem.* **5**, 1789–1797.
- Schaschke, N., Assfalg-Machleidt, I., Machleidt, W. and Moroder, L. (1998) Substrate/propeptide-derived endo-epoxysuccinyl peptides as highly potent and selective cathepsin B inhibitors. *FEBS Lett.* **421**, 80–82.
- Katunuma, N., Murata, E., Kakegawa, H., Matsui, A., Tsuzuki, H., Tsuge, H., Turk, D., Turk, V., Fukushima, M., Tada, Y. et al. (1999) Structure based development of novel specific inhibitors for cathepsin L and cathepsin S *in vitro* and *in vivo*. *FEBS Lett.* **458**, 6–10.
- Tsuge, H., Nishimura, T., Tada, Y., Asao, T., Turk, D., Turk, V. and Katunuma, N. (1999) Inhibition mechanism of cathepsin L-specific inhibitors based on the crystal structure of papain–CLIK148 complex. *Biochem. Biophys. Res. Commun.* **266**, 411–416.
- Zvonar-Popovič, T., Lah, T., Kregar, I. and Turk, V. (1980) Some characteristics of cathepsin B and α -N-benzoylarginine- β -naphthylamide hydrolase from bovine lymph nodes. *Croat. Chem. Acta* **53**, 509–517.
- Barrett, A. J. and Kirschke, H. (1981) Cathepsin B, cathepsin H, and cathepsin L. *Methods Enzymol.* **80**, 535–561.
- Pace, C. N., Vajdos, F., Fee, L., Frimsley, G. and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423.
- Turk, B., Križaj, I., Kralj, B., Dolenc, I., Popovič, T., Bieth, J. G. and Turk, V. (1993) Bovine stefin C, a new member of the stefin family. *J. Biol. Chem.* **268**, 7323–7329.
- Yamamoto, A., Kaji, T., Tomoo, K., Ishida, T., Inoue, M., Kitamura, K. and Murata, M. (1992) Crystallization and preliminary X-ray study of the cathepsin B complexed with CA074, a selective inhibitor. *J. Mol. Biol.* **227**, 942–944.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
- Kissinger, C. R., Gehlhaar, D. K. and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D* **55**, 484–491.
- Turk, D. (1992) Weiterentwicklung eines Programms für Molekülgraphik und Elektronendichte-Manipulation und seine Anwendung auf verschiedene Protein-Strukturaufklärungen. Ph.D. thesis, Technische Universität München, München.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
- Yamamoto, A., Hara, T., Tomoo, K., Ishida, T., Fujii, T., Hata, Y., Murata, M. and Kitamura, K. (1997) Binding mode of CA074, a specific irreversible inhibitor, to bovine cathepsin B as determined by X-ray crystal analysis of the complex. *J. Biochem. (Tokyo)* **121**, 974–977.
- Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
- Varughese, K. I., Su, Y., Cromwell, D., Hasnain, S. and Xiong, N. H. (1992) Crystal structure of actinidin–E-64 complex. *Biochemistry* **31**, 5172–5176.
- Zhao, B., Janson, C. A., Amegadzie, B. Y., D'Alessio, K., Griffin, C., Hanning, C. R., Jones, C., Kurdyla, J., McQueney, M., Qiu, X. et al. (1997) Crystal structure of human osteoclast cathepsin K complex with E-64. *Nat. Struct. Biol.* **4**, 109–111.
- Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S. and Cygler, M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.* **15**, 5492–5503.
- LaLonde, J. M., Zhao, B., Janson, C. A., D'Alessio, K. J., McQueney, M. S., Orsini, M. J., Debouck, C. M. and Smith, W. W. (1999) The crystal structure of human procathepsin K. *Biochemistry* **38**, 862–869.

-
- 42 Sivaraman, J., Nägler, D. K., Zhang, R., Menard, R. and Cygler, M. (2000) Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. *J. Mol. Biol.* **295**, 939–951
- 43 Schaschke, N., Assfalg-Machleidt, I., Lassleben, Th., Sommerhof, C. P., Moroder, L. and Machleidt, W. (2000) Epoxysuccinyl peptide-derived affinity labels for cathepsin B. *FEBS Lett.* **482**, 91–96
- 44 Schaschke, N., Assfalg-Machleidt, I., Machleidt, W., Lassleben, T., Sommerhof, C. P. and Moroder, L. (2000) β -Cyclodextrin–epoxysuccinyl peptide conjugates: a new drug targeting system for tumor cells. *Bioorg. Med. Chem. Lett.* **10**, 677–680
- 45 Thompson, S. K., Halbert, S. M., Bossard, M. J., Thaddeus, A. T., Ley, M. A., Zhao, B., Smith, W. W., Abdel-Meguid, A. A., Janson, C. A., D'Alessio, K. J. et al. (1997) Design of potent and selective human cathepsin K inhibitors that span the active site. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14249–14254
- 46 Merritt, E. A. and Bacon, D. J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524
- 47 Nicholls, A., Sharp, K. A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Gen.* **11**, 281–296

Received 12 February 2004/5 April 2004; accepted 15 April 2004

Published as BJ Immediate Publication 15 April 2004, DOI 10.1042/BJ20040237