Structure of the complex of carboxypeptidase B and *N*-sulfamoyl-L-arginine

Valery Akparov,^{a*} Nikolay Sokolenko,^b Vladimir Timofeev^{c,d} and Inna Kuranova^{c,d}

^aProtein Chemistry Department, State Research Institute for Genetics and Selection of Industrial Microorganisms, 1yi Dorozhnyi Proezd 1, 117545 Moscow, Russian Federation, ^bChemical Synthesis Department, Innovation Research Centre 'Solaris', Tallinskaya str. 32-2-44, 123458 Moscow, Russian Federation, ^cX-ray Analysis Methods and Synchrotron Radiation Laboratory, Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninsky Prospekt 59, 119333 Moscow, Russian Federation, and ^dNBICS Center, National Research Centre 'Kurchatov Institute', Akad. Kurchatova Sq. 1, 123182 Moscow, Russian Federation. *Correspondence e-mail: valery@akparov.ru

Received 23 June 2015

Accepted 8 September 2015

Edited by J. Newman, Bio21 Collaborative Crystallization Centre, Australia

Keywords: carboxypeptidase B; *N*-sulfamoyl-L-arginine; X-ray analysis.

PDB reference: carboxypeptidase B, complex with *N*-sulfamoyl-L-arginine, 4z65

Supporting information: this article has supporting information at journals.iucr.org/f

Porcine pancreatic carboxypeptidase B (EC 3.4.23.6) was complexed with a stable transition-state analogue, *N*-sulfamoyl-L-arginine, in which an S atom imitates the *sp*³-hybridized carbon in the scissile-bond surrogate. Crystals were grown in a form belonging to the same space group, *P*₄₁₂₁₂, as the uncomplexed enzyme. X-ray data were collected to a resolution of 1.25 Å. The molecule was refined and the positions of non-H atoms of the inhibitor and water molecules were defined using difference Fourier maps. The enzyme–inhibitor complex and 329 water molecules were further refined to a crystallographic *R* factor of 0.159. The differences in conformation between the complexed and uncomplexed forms of carboxypeptidase B are shown. The inhibitor is bound in a curved conformation in the active-site cleft, and the sulfamide group is bound to the Zn ion in an asymmetric bidentate fashion. The complex is stabilized by hydrogen bonds between the N1/N2 guanidine group of the inhibitor and the Asp255 carboxyl of the enzyme. The side-chain CH₂ groups of the inhibitor are in van der Waals contact with Leu203 and Ile247 in the enzyme. This study provides useful clues concerning how the transition state of arginine may bind to carboxypeptidase B and therefore provides an insight into the structural basis of carboxypeptidase B selectivity, which is useful for the rational design of a carboxypeptidase with improved selectivity for industrial recombinant pro-insulin processing.

1. Introduction

Pancreatic porcine metallocarboxypeptidase B (EC 3.4.17.2; CPB) is a zinc-dependent metallocarboxypeptidase (MPC) that participates in food digestion and preferably cleaves off positively charged residues from the C-terminal ends of peptides and proteins (Adler *et al.*, 2005). It is used in the insulin industry and in drug development as a model for blood carboxypeptidase B (Bunnage *et al.*, 2007). We used carboxypeptidase B as a prototype to investigate the structural basis of the wide substrate selectivity of carboxypeptidase T from *Thermoactinomyces vulgaris* (Grishin *et al.*, 2008; Akparov *et al.*, 2015). Because these investigations were conducted for the rational design of metallocarboxypeptidase B with improved or changed selectivity, they required detailed information about the active site in order to provide a three-dimensional structure of the complex with the transition-state analogue of an arginine substrate. Although A/B-type MCPs are among the most thoroughly studied proteolytic enzymes, no structures have been published. In this work, we synthesized and crystallized porcine carboxypeptidase B with *N*-sulfamoyl-L-arginine (SArg), which may be considered a stable

transition-state arginine analogue. The similar inhibitor *N*-sulfamoyl-L-phenylalanine has been used with carboxypeptidase A (CPA; EC 3.4.17.1; Park *et al.*, 2002).

2. Materials and methods

2.1. Macromolecule production

Porcine pancreatic carboxypeptidase B was purchased from Sigma (USA).

2.2. *N*-Sulfamoyl-L-arginine synthesis

2.2.1. Benzyl-*N*²-([[(benzyloxy)carbonyl]amino}sulfonyl)-*N*⁵-[imino(nitroamino)methyl]-L-ornithinate (MW 522.53; Fig. 1). Benzyl alcohol (1.2 ml, 11.5 mmol) was slowly added at 0°C to chlorosulfonylisocyanate (1 ml, 11.5 mmol) solution in anhydrous dichloromethane (10 ml) under stirring, and the stirring was continued for 30 min at 0°C. *N,N*-Diisopropylethylamine (DIPEA; 6 ml, 34 mmol) in anhydrous dichloromethane was added to the solution. The resulting mixture was then added dropwise to an ice-chilled suspension of *N*^ω-nitro-L-arginine benzyl ester *p*-toluenesulfonate (5.30 g, 11 mmol) and DIPEA (2.5 ml, 14.3 mmol) in anhydrous dichloromethane (25 ml). The obtained solution was stirred for 2 h at room temperature and evaporated under reduced pressure. The residue was dissolved in a mixture of ethyl acetate (50 ml) and water (50 ml); the ethylacetate layer was successively washed with a 5% solution of H₂SO₄ (50 ml × 2) and brine

(50 ml × 2) and dried over anhydrous Na₂SO₄. The ethyl acetate solution was filtered from Na₂SO₄ and concentrated *in vacuo* to produce an oily residue. The crude product was purified using flash chromatography on silica gel (CHCl₃:ethyl acetate gradient from 2:1 to 1:3) to produce an oily product. The oil was dissolved in chloroform and the crystalline product precipitated overnight at 4°C. The crystals were filtered, washed with cold chloroform (20 ml) and petroleum ether 40–70 (30 ml × 2) and dried *in vacuo*. The yield was 3.188 g (6.1 mmol, 55.5%).

TLC: toluene:acetone:acetic acid (100:50:1), *R*_f 0.32; CHCl₃:methanol:acetic acid (45:5:1), *R*_f 0.55. Melting point: 113.4–118.0°C. [α]_D²⁵ = +4.0 (*c* 1.0, methanol). MS: [*M* + *H*⁺], 523.14. ¹H NMR (DMSO-*d*₆, 400 MHz): δ 11.4 (s, 1H, NH), 7.6–8.5 (m, 4H, 4 NH), 7.35–7.37 (m, 10H, CH₂Ph), 5.08 (s, 2H, CH₂Ph), 5.07 (s, 2H, CH₂Ph), 4.0–4.05 (m, 1H, CH), 3.05–3.20 (m, 2H, CH₂), 1.40–1.78 (m, 4H, 2 CH₂).

2.2.2. *N*-Sulfamoyl-L-arginine (Fig. 2; MW 253.28). Benzyl *N*²-([[(benzyloxy)carbonyl]amino}sulfonyl)-*N*⁵-[imino(nitroamino)methyl]-L-ornithinate (3.136 g, 6.0 mmol) was dissolved in a mixture of methanol (60 ml) and cyclohexene (10 ml, ~100 mmol). A suspension of 10% Pd on carbon (1 g) in a mixture of 95% ethanol (6 ml) and water (3 ml) was added to the solution. The reaction mixture was stirred for 6 h at 55°C. The reaction mixture was cooled to room temperature and filtered using Pd on carbon. The Pd on carbon was washed with methanol (30 ml × 2). The collected solution was evaporated under reduced pressure. The oily residue was dissolved in acetone (30 ml) and the crystalline product precipitated overnight at 4°C. The crystals were filtered, washed with cold acetone (20 ml) and petroleum ether 40–70 (30 ml × 2) and dried *in vacuo*. Yield, 0.692 g (2.732 mmol, 45.5%).

TLC: *n*-butanol:acetic acid:H₂O (4:1:1), *R*_f 0.40. Melting point: 195°C (decomposing). [α]_D²⁵ = 3.5 (*c* 0.5, H₂O). MS: [*M* + *H*⁺], 254.08. ¹H NMR (DMSO-*d*₆, 400 MHz): δ 9.10 (d, 1H, OH), 7.3–7.8 (s, 3H, NH, NH₂), 6.54 (s, 2H, NH₂), 5.78 (s, 11H, NH), 3.52 (s, 1H, CH), 3.01–3.09 (m, 2H, CH₂), 1.37–1.70 (m, 4H, CH₂CH₂).

Melting points were determined on an Optimelt MPA100 (Stanford Research Systems, UK) capillary melting-point apparatus and are uncorrected. The ¹H NMR spectra were recorded with a Bruker FT 400 (400 MHz) instrument using tetramethylsilane as an internal standard. Silica gel 60 (230–

Figure 1
Synthesis scheme of benzyl-*N*²-([[(benzyloxy)carbonyl]amino}sulfonyl)-*N*⁵-[imino(nitroamino)methyl]-L-ornithinate.

Figure 2
Synthesis scheme of *N*-sulfamoyl-L-arginine.

Table 1
Crystallization.

Method	Counter-diffusion
Plate type	Capillary
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	12
Buffer composition of protein solution	Water
Composition of precipitant	16% PEG 8000, 0.1 M sodium cacodylate, 0.1 M zinc acetate
Volume of protein solution in the capillary	8 µl protein solution (no reservoir solution)
Volume of reservoir solution (µl)	90

400 mesh) was used for flash chromatography, and thin-layer chromatography (TLC) was performed on silica-coated glass sheets (Merck silica gel 60 F₂₅₄). Optical rotations were measured on an ADP 440 polarimeter (Bellingham & Stanley, UK). Mass spectra were obtained at Orekhovich's Institute of Biomedical Chemistry, Moscow. Elemental analyses were performed at Zakusov's Institute of Pharmacology, Moscow and the results were within 0.4% of the theoretical values.

2.3. Crystallization

Crystals of the CPB–SArg complex were grown in a capillary in microgravity using the counter-diffusion technique (Fig. 3, Table 1). The equipment and technology developed by the Aerospace Agency of Japan (JAXA) were used for crystal growth, as described by Takahashi *et al.* (2010) and Kuranova *et al.* (2011). The protein concentration was 12 mg ml⁻¹ in water. The precipitating solution consisted of 16% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.1 M zinc acetate, 10 mM *N*-sulfamoyl-L-arginine.

A glass capillary (0.5 mm in diameter) containing the protein solution served as a crystallization device. One end of the capillary was hermetically sealed; a silicon pipe (0.5 mm), which was filled with 1% agarose gel and plunged into a cylinder with a precipitating agent, was connected to the other end. Several crystallization devices placed in special boxes were delivered to the International Space Station, where crystals were grown at 20°C. It is well known that the slow diffusion of a precipitant into a protein solution through a gel layer under zero-gravity conditions improves the diffraction properties of crystals (McPherson, 1996). To collect diffraction data using a synchrotron source, the crystal was removed from the capillary, placed into precipitant solution and subsequently a cryosolution, and cooled in a nitrogen-vapour flow. In addition to the precipitant components, the cryosolution also contained 20%(v/v) glycerol.

2.4. Data collection and processing

Diffraction data sets were collected at the BL41XU station of the Spring-8 synchrotron, Japan at a temperature of 100 K. A MAR Mosaic 225 CCD device was used as a detector. The diffraction data were obtained using a single crystal by rotation. The wavelength was 0.8 Å, the crystal-to-detector distance was 100 mm, the oscillation angle was 0.5° and the angle of rotation was 180°. The experimental intensities were

Table 2
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Spring-8, Japan
Wavelength (Å)	0.8
Temperature (K)	100
Detector	MAR Mosaic 225 CCD
Crystal-to-detector distance (mm)	100
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	0.5
Space group	<i>P</i> ₄ ₁ ₂ ₁ ₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.380, 79.380, 100.370
α , β , γ (°)	90.00, 90.00, 90.00
Mosaicity (°)	0.09
Resolution range (Å)	30–1.25 (1.32–1.25)
Total No. of reflections	1195243
No. of unique reflections	88471
Completeness (%)	99.4 (96.4)
Multiplicity	13.51 (7.44)
$\langle I/\sigma(I) \rangle$	16.1 (3.2)
<i>R</i> _{r.i.m.}	0.111 (0.498)
Overall <i>B</i> factor from Wilson plot (Å ²)	8.3

Table 3
Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	30–1.25 (1.28–1.25)
Completeness (%)	99.4 (91.9)
No. of reflections, working set	83959 (4426)
No. of reflections, test set	5341 (325)
Final <i>R</i> _{cryst}	0.159 (0.180)
Final <i>R</i> _{free}	0.180 (0.218)
No. of non-H atoms	
Protein	2436
Ion	2
Ligand	16
Water	329
Total	2783
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.318
Average <i>B</i> factors (Å ²)	
Protein	10.833
Ion	7.427
Ligand	7.548
Water	19.747
Ramachandran plot	
Most favoured (%)	97
Allowed (%)	3

processed using *iMosflm* (Battye *et al.*, 2011). The data set was processed to 1.25 Å resolution. The crystals belonged to space group *P*₄₁₂₁₂. A single molecule of the enzyme was present in the independent part of the cell. The statistical characteristics of the data set are given in Table 2.

2.5. Structure solution and refinement

The structure of the CPB–SArg complex was determined at 1.25 Å resolution by the molecular-replacement method using *Phaser* (McCoy *et al.*, 2007) with the atomic coordinates of CPB (PDB entry 3wc6; Yoshimoto *et al.*, 2013) as a starting model. For anisotropic structure refinement, *REFMAC* (Murshudov *et al.*, 2011) was used. Manual correction of the models was performed using *Coot* (Emsley & Cowtan, 2004),

and electron-density maps were calculated with $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ coefficients. Water molecules and calcium ions were located in the electron-density maps, and difference Fourier synthesis revealed electron density in the active site which was identified as a ligand. The ligand was refined with an occupancy of 100%. The coordinates of the structure have been deposited in the PDB (PDB entry 4z65). The refinement statistics are presented in Table 3.

3. Results and discussion

The final model was refined at 1.25 Å resolution with R and R_{free} values of 0.159 and 0.180, respectively. The asymmetric unit contains one protein molecule (molecule A; residues 6–309), two zinc ions per protein molecule and 329 water molecules. The polypeptide chain of CPB can be clearly and completely traced in the electron-density map (Gly6–Leu306). The asymmetric unit of the crystal contains one polypeptide chain. The active centre of the enzyme contains electron density that was interpreted as an *N*-sulfamoyl-L-arginine molecule (Fig. 4). The conformational changes upon ligand binding are presented in Fig. 5. The side chains of Ile247 and Tyr248 change their orientations.

Carboxypeptidase B has the α/β fold, which is typical of hydrolases, with a β -sheet which is formed by eight β -strands and is surrounded by nine α -helices (Fig. 6). This entire fold replicates the fold of a close homologue of CPB, CPA (Teplakov *et al.*, 1992).

The active centre of CPB contains an *N*-sulfamoyl-L-arginine molecule, which is bound to a catalytic zinc ion (Fig. 6).

The sulfamide moiety of *N*-sulfamoyl-L-arginine has an ill-shaped tetrahedral configuration, which mimics the sp^3 -hybridized configuration of the C atom of a scissile bond. The binding mode of the sulfamide moiety is notably similar to that in the complex of CPA with *N*-sulfamoyl-L-phenylalanine. The environment of the *N*-sulfamoyl-L-arginine in the active site of CPB is shown in Fig. 7. One (O4) of the two O atoms of the sulfamide moiety of *N*-sulfamoyl-L-arginine is engaged in a hydrogen bond to an N atom of the guanidinium group in

Arg127, with a bond distance of 2.78 Å. The other O atom in the sulfamide moiety is in an open space that is exposed to bulk water and is separated from the O atoms of the Glu270 carboxylate by 4.07 and 4.21 Å. The Glu270 carboxylate forms

Figure 4

Electron density for *N*-sulfamoyl-L-arginine. The electron density was calculated with $|F_o| - |F_c|$ coefficients at the 2.0σ level. The ligand was excluded in the calculation of the electron-density map.

Figure 5

The conformational changes in the active site of carboxypeptidase B upon binding *N*-sulfamoyl-L-arginine. The amino-acid residues in the free enzyme are coloured blue and those in the CPB-SAArg complex are in red. *N*-Sulfamoyl-L-arginine is shown in stick representation. C atoms are coloured green, O atoms red, N atoms blue and the S atom yellow.

Figure 3

Crystal of CPB complexed with *N*-sulfamoyl-L-arginine in the capillary.

Table 4

Hydrogen-bond net between fixed water molecules that surround the guanidine group of the ligand and the protein.

Water molecule	Distance (Å)	Atom
519	3.60	SArg N13
	3.69	SArg N16
	2.63	Ser207 OG
	2.85	Asp255 OD2
	2.85	Asp255 OD1
	3.38	Asp255 OD2
	2.80	Asp256 OD1
	3.56	Tyr204 O
	3.35	Ser207 N
670	3.57	SArg N13
	2.95	Asn144 ND2
	3.38	Asn144 OD1
	3.73	Ser254 OG
	3.27	Asp255 OD1
	3.43	Asp255 OD2
	2.88	Thr268 OG1
655	2.91	H ₂ O532
	2.88	SArg N15
	3.33	SArg N16
	2.86	Asn144 OD1
	2.96	Ala251 O
	3.14	Ala250 O
	3.33	Arg145 NE1
	3.25	Gly253 N
591	3.39	H ₂ O591
	3.39	H ₂ O655
	2.95	SArg N16
	2.78	Asp256 OD2
	2.89	Gly253 N
627	2.77	Ala250 O
	3.32	Gly252 N
	3.74	Tyr208 N
	2.87	SArg N15
	2.81	Ser207 OG
	2.83	Tyr248 O

a hydrogen bond to N1 of the sulfamide: one of the Glu270 carboxylate O atoms is separated from N1 of the sulfamide moiety by 2.77 Å. It is also noteworthy that instead of the O atom, the terminal N atom (N2) of the sulfamide moiety coordinates to the zinc ion in the active site of CPA with a bond distance of 1.97 Å.

The carboxylate of the ligand forms a salt bridge with the guanidinium moiety of Arg145 with bond distances of 2.92 and 2.81 Å. The amide group of the side chain of Asn144 forms a hydrogen bond (2.99 Å) to O2 of the carboxylate. The aromatic side chain of Tyr248 is found in the so-called ‘down’ position, and its phenolic O atom is separated from one of the carboxylate O atoms of the inhibitor by 2.61 Å, forming a hydrogen bond. The side chain of *N*-sulfamoyl-L-arginine is anchored in the S1’ subsite, which is the primary substrate-recognition pocket of CPB. The guanidine group of the ligand forms a salt bridge and two hydrogen bonds (3.00 and 2.71 Å) to Asp255. Five fixed water molecules surround the guanidine group and form a kind of ‘ice jacket’ between the guanidine group and the protein. These groups also form water-mediated hydrogen bonds to Ser207, Tyr208, Asn144, Tyr248, Ala250, Ala251, Gly252, Gly253, Asp255 and Asp256 (see Table 4). An identical net of hydrogen bonds is found in the complex of CPB with 5-[[amino(imino)methyl]amino]-2-(sulfanylmethyl)-pentanoic acid (PDB entry 1zg9; Adler *et al.*, 2005), but the

conformations of the *N*-sulfamoyl-L-arginine side chain differ in these two complexes. The guanidine groups of both inhibitors are near the Asp255 carboxyl group, but the corre-

Figure 6

Folding of CPB according to the data for the complex with *N*-sulfamoyl-L-arginine. *N*-Sulfamoyl-L-arginine in the active site of CPB is shown in stick representation.

Figure 7

Environment of *N*-sulfamoyl-L-arginine in the active site of CPB.

Figure 8
Comparison of the environments of 5-[[amino(imino)methyl]amino]-2-(sulfanylmethyl)pentanoic acid and *N*-sulfamoyl-L-arginine in their complexes with CPB.

sponding N atoms are shifted by 0.2–0.63 Å. The lengths of these hydrogen bonds are also different (see Fig. 8).

Acknowledgements

The work was supported by the Russian Foundation for Basic Research (Project 14-08-01245) and the Central Research Institute for Mechanical Engineering, ROSKOSMOS. We

thank our Japanese colleagues K. Ohta, H. Tanaka and K. Inaka for loading and assembling the JCB crystallization box and for their assistance in collecting X-ray diffraction data sets at the SPring-8 synchrotron-radiation facility.

References

- Adler, M., Bryant, J., Buckman, B., Islam, I., Larsen, B., Finster, S., Kent, L., May, K., Mohan, R., Yuan, S. & Whitlow, M. (2005). *Biochemistry*, **44**, 9339–9347.
- Akparov, V. K., Timofeev, V. I., Khaliullin, I. G., Švedas, V., Chestukhina, G. G. & Kuranova, I. P. (2015). *FEBS J.* **282**, 1214–1224.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Bunnage, M. E. *et al.* (2007). *J. Med. Chem.* **50**, 6095–6103.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Grishin, A. M., Akparov, V. K. & Chestukhina, G. G. (2008). *Protein Eng.* **21**, 545–551.
- Kuranova, I. P., Smirnova, E. A., Abramchik, Y. A., Chupova, L. A., Esipov, R. S., Akparov, V. K., Timofeev, V. I. & Kovalchuk, M. V. (2011). *Crystallogr. Rep.* **56**, 884–891.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McPherson, A. (1996). *Crystallogr. Rev.* **6**, 157–305.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Park, J. D., Kim, D. H., Kim, S.-J., Woo, J.-R. & Ryu, S. E. (2002). *J. Med. Chem.* **45**, 5295–5302.
- Takahashi, S., Tsurumura, T., Aritake, K., Furubayashi, N., Sato, M., Yamanaka, M., Hirota, E., Sano, S., Kobayashi, T., Tanaka, T., Inaka, K., Tanaka, H. & Urade, Y. (2010). *Acta Cryst.* **F66**, 846–850.
- Teplyakov, A., Polyakov, K., Obmolova, G., Strokopytov, B., Kuranova, I., Osterman, A., Grishin, N., Smulevitch, S., Zagnitko, O., Galperina, O., Matz, M. & Stepanov, V. (1992). *Eur. J. Biochem.* **208**, 281–288.
- Yoshimoto, N., Itoh, T., Inaba, Y., Ishii, H. & Yamamoto, K. (2013). *J. Med. Chem.* **56**, 7527–7535.