Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray crystallographic analysis of the cysteine protease inhibitor clitocypin

Katja Galeša, Jože Brzin, Jerica Sabotič and Dušan Turk*

Department of Biochemistry and Molecular Biology, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

Correspondence e-mail: dusan.turk@ijs.si

Received 4 October 2005
Accepted 21 November 2005
Online 16 December 2005

Clitocypin is a cysteine protease inhibitor from the mushroom *Clitocybe nebularis*. The protein has been purified from natural sources and crystallized in a variety of non-isomorphous forms belonging to monoclinic and triclinic space groups. A diffraction data set to 1.55 Å resolution was obtained from a crystal belonging to space group $P2$, with unit-cell parameters a = 38.326, b = 33.597, c = 55.568 Å, $\beta = 104^\circ$. An inability to achieve isomorphism forced the use of MAD and SAD phasing methods. Phasing is in progress.

1. Introduction

In recent decades, several families of cysteine protease inhibitors have been discovered (reviewed in Turk *et al.*, 2005; Dubin, 2005) and classified (MEROPS; Rawlings *et al.*, 2004). Cystatins were the first to be discovered and are the best characterized. The three-dimensional structure of the complex formed between human stefin B and papain demonstrated the inhibitory mechanism, which is based on the N-terminal trunk and two hairpin loops that bind to the active-site cleft of papain (Stubbs *et al.*, 1990).

The second best characterized family of cysteine protease inhibitors are the thyropins. Their structure is based on that of the thyroglobulin type-1 domain (Thyr-1), a cysteine-rich structural element present as a single or repetitive module in a variety of functionally unrelated proteins. Inhibition by thyropins is mainly confined to the papain family of cysteine proteases (Lenarčič & Bevec, 1998). The general mechanism of Thyr-1 domain interactions with cysteine proteases was revealed by the crystal structure of a complex of the inhibitory fragment p41 form of the invariant chain and cathepsin L (Gunčar *et al.*, 1999).

Staphostatins are extremely selective inhibitors from *Staphylococcus aureus* and *S. epidermis*. Apart from staphopains, no other proteases tested from the CA clan were inhibited (Rzychon *et al.*, 2003; Dubin *et al.*, 2004). The crystal structure of the complex staphostatin B–staphopain B from *S. aureus* demonstrated a new mechanism of action previously not observed in cysteine protease inhibitors (Filipek *et al.*, 2003).

The chagasin family of papain-like cysteine protease inhibitors has been identified in protozoan parasites as well as in bacterial pathogens. The proteins have low sequence similarity to each other and have no significant similarity to cystatins or other known cysteine peptidase inhibitors (Sanderson *et al.*, 2003; Monteiro *et al.*, 2001). Cysteine proteases are further inhibited by proteins that are similar to classical inhibitors of serine proteases (Križaj *et al.*, 1993; Schick *et al.*, 1998) or homologous to proteinase propeptide regions (Yamamoto *et al.*, 1999).

Clitocypin from *Clitocybe nebularis* is the first protein inhibitor of cysteine proteases to have been isolated and characterized from higher fungi and has no similarity to any known cysteine protease inhibitors (Brzin *et al.*, 2000). For this reason, it was classified in a new family (I48) of the MEROPS inhibitor classification (Schick *et al.*, 1998). Its sequence of 150 amino-acid residues contains no cysteines, no methionines and no histidines. The protein is not glycosylated. The targets of clitocypin are not known. Control of endogenous proteases seems unlikely, given the high amounts of the inhibitor produced.

Table 1
Data-collection statistics for crystals of native and derivatized clitocypin.

Values in parentheses are for the highest resolution shell.

Crystal form	Native	ClitIr1-1 (MAD)	ClitIr1-2 (MAD)	ClitIr1-3 (MAD)	ClitIr2-1 (MAD)	ClitIr2-2 (MAD)	ClitIr2-3 (MAD)	ClitOs1 (SAD)	ClitOs2 (SAD)
Space group	*P*2	*P*2	*P*2	*P*2	*P*1	*P*1	*P*1	*P*2	*P*1
Wavelength (Å)	1.000	1.10565	1.10520	0.9500	1.10565	1.10520	1.0500	1.10697	1.10697
Unit-cell parameters									
a (Å)	38.33	46.55	46.55	46.55	34.36	34.36	34.36	33.75	34.25
b (Å)	33.60	58.00	58.00	58.00	45.99	45.99	45.99	57.44	45.70
c (Å)	55.57	58.26	58.26	58.26	47.19	47.19	47.19	34.78	47.19
α (°)	90	90	90	90	108.92	108.92	108.92	90	109.05
β (°)	104.00	111.21	111.21	111.21	91.30	91.30	91.30	111.91	90.90
γ (°)	90	90	90	90	98.81	98.81	98.81	90	98.81
Resolution range (Å)	50–1.55 (1.65–1.55)	50–2.27 (2.40–2.27)	50–2.27 (2.40–2.27)	50–1.95 (2.05–1.95)	50–1.66 (1.80–1.66)	50–1.67 (1.80–1.67)	50–1.65 (1.75–1.65)	50–2.44 (2.55–2.44)	50–2.01 (2.15–2.01)
Mosaicity (°)	1.14	0.44	0.44	0.44	0.42	0.42	0.45	1.38	0.51
Multiplicity	3.8 (3.5)	7.2 (6.8)	7.3 (6.9)	7.2 (6.9)	7.7 (6.2)	7.7 (6.6)	2.0 (2.0)	10.7 (7.2)	3.6 (2.8)
Completeness (%)	99.4 (99.8)	98.1 (89.3)	98.0 (88.7)	97.5 (83.9)	94.0 (71.5)	96.4 (93.9)	88 (94.5)	98.0 (86.1)	95.9 (80.5)
No. of molecules per AU	1	1–2	1–2	1–2	2	2	2	1	2
$\langle I/\sigma(I)\rangle$	23.65 (6.92)	49.88 (25.29)	50.76 (26.47)	41.83 (16.19)	40.06 (13.01)	40.26 (13.16)	22.92 (11.52)	61.67 (28.85)	22.20 (12.24)
$R_{\rm merge}$† (%)	7.2 (18.7)	2.2 (6.1)	2.1 (5.9)	2.4 (10.2)	3.1 (21)	3.3 (19.2)	2.5 (5.5)	3.8 (7.7)	2.9 (9.2)

† $R_{\rm merge} = \sum|I_i - \langle I_i\rangle|/\sum\langle I_i\rangle$, where I_i is the observed intensity and $\langle I_i\rangle$ is the average intensity over symmetry-equivalent reflections.

Protection against pathogen infection or predation by insects appears to be more probable. The mushroom protein inhibits papain, bromelain and cathepsins B and L with K_i values in the nanomolar range (Brzin *et al.*, 2000). Clitocypin is a monomeric protein with a molecular weight of 16.8 kDa as determined by analytical centrifugation and gel-exclusion chromatography (Galeša *et al.*, 2004). Infrared spectroscopy shows that clitocypin contains a high proportion of antiparallel β-sheet and little or no helical structure (Kidrič *et al.*, 2002). The high activation energy of unfolding, slow deuterium exchange and the thermal unfolding temperature indicate the high stability of the structure (Kidrič *et al.*, 2002). In this report, we present the crystallization of clitocypin and crystallographic data analysis of its crystals.

2. Crystallization

Clitocypin was isolated from the mushroom *C. nebularis* as described by Brzin *et al.* (2000). The protein was concentrated to 8–20 mg ml^{-1} in 15 m*M* MES pH 6.0 using Amicon Centricon 3 kDa filters as estimated from the absorption at 280 nm (ε = 21 620 M^{-1} cm^{-1}). Sitting-drop vapour diffusion was used to screen for suitable crystallization conditions using Crystal Screen, Crystal Screen 2, Crystal Screen Cryo and PEG/Ion Screen (Hampton Research). Drops were prepared by combining equal volumes of protein solution and reservoir solution (1 µl each) and were equilibrated against 900 µl reservoir solution. Crystals were obtained when the protein was equilibrated against 50 m*M* monopotassium dihydrogen phosphate, 20%(*w*/*v*) polyethylene glycol 8000 pH 3.76. Several non-isomorphous crystals appeared within a period of several days to a few weeks, depending on the protein concentration. The size and morphology of the crystals varied. The monoclinic crystals often look like rhomboid prisms with lengths from 50 to 350 µm, while larger (200–400 µm) triclinic crystals obey no geometrical laws. 50–70% of the crystals were useless for crystallographic analysis. We tried to optimize crystals by the use of additives (Additive Screens 1, 2 and 3, Detergent Screen, Hampton Research), pH screening, alteration of protein concentration and buffer composition and the use of different plates (24-well hanging-drop plates, 24-well sitting-drop plates, microbatch 96-well plates). The variability and irreproducibility of crystallization most probably reflects the microheterogeneity of the inhibitor from the natural source, as indicated by the presence of several genome and cDNA sequences (J. Sabotič *et al.*, manuscript in preparation).

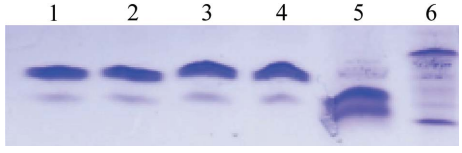

Figure 1
Isoelectric focusing of derivatization reactions. 2 µl protein solution at a concentration of 2 mg ml^{-1} was mixed with 2 µl derivatization solution. The mixtures were left at room temperature for 10 min and then loaded onto Phastgel IEF 3-9 (Amersham Pharmacia Biotech). [Lanes 1–6: native protein, gadolinium(III) chloride hydrate, sodium tungstate dihydrate, ethylmercuric phosphate, potassium osmate(IV) dihydrate, IEF standard.]

3. Data collection and analysis

X-ray diffraction data from a native crystal of clitocypin was collected at beamline XRD-1 at Elettra in Trieste. The crystal was inserted directly in the cryostream without the use of an additional cryoprotectant. The data were indexed, scaled and integrated using *HKL*2000 (Otwinowski & Minor, 1997). The resolution of diffraction was 1.55 Å and the space group was identified as *P*2, with unit-cell parameters a = 38.33, b = 33.60, c = 55.57 Å, β = 104°. A value of $V_{\rm M}$ of 2 Å^3 Da^{-1} was calculated according to Matthews (1968), assuming the presence of one molecule in the asymmetric unit. From this calculation the solvent content is 39%. No peak was observed in the self-rotation function that would indicate the presence of non-crystallographic symmetry.

The absence of cysteine, histidine and methionine residues in the protein and the use of phosphate in the crystallization conditions seriously reduce the number of potential derivatives. Our inability to produce isomorphous crystals of clitocypin prevented MIR and SIR approaches. Cocrystallization of the protein with heavy-metal derivatives resulted in non-isomorphous low-quality crystals that were useless for further experiments. Finally, crystals soaked in cryobuffer [50 m*M* monopotassium dihydrogen phosphate, 20%(*w*/*v*) polyethylene glycol 8000 pH 3.76, 15% ethylene glycol] enriched with heavy-atom derivatives at concentrations of 1–10 m*M* were used. The

crystals were left in solutions of bromide, caesium chloride, ethylmercuric phosphate, potassium hexachloroiridate(IV), rubidium chloride, cobalt chloride and strontium nitrate for 1–30 min. The soaking was stopped immediately when crystals started dissolving or cracking. MAD data sets at three different wavelength were collected on the Max-Planck beamline (BW6) at the Deutsches Elektronensynchrotron in Hamburg. The wavelengths were based on the fluorescence absorption spectrum near the iridium L_{III} absorption edge. The crystal ClitIr1 (data sets ClitIr1-1, ClitIr1-2 and ClitIr1-3) belongs to the monoclinic space group with unit-cell parameters a = 46.55, b = 58.00, c = 58.26 Å, β = 111.21°, whereas the crystal CllitIr2 (data sets ClitIr2-1, ClitIr2-2 and ClitIr3-3) was triclinic (a = 34.36, b = 45.99, c = 47.19 Å, α = 108.92, β = 91.30, γ = 98.81°). The data-collection statistics of the best data sets are given in Table 1.

The search for heavy-atom compounds was subsequently speeded up by pre-screening for derivatization by band shift on isoelectric focusing (Fig. 1). A change in protein mobility, indicating phasing-atom derivatization, was observed for potassium osmate(IV). Soaked crystals were used for SAD data collection. SAD data sets from a monoclinic crystal (ClitOs1; a = 33.75, b = 57.44, c = 34.78 Å, β = 111.91°) and a triclinic crystal isomorphous to ClitIr2 (ClitOs2; a = 34.25, b = 45.70, c = 47.19 Å, α = 109.05, β = 90.90, γ = 98.81°) were collected at the XRD-1 beamline at Elettra in Trieste. The wavelengths used for X-ray diffraction were chosen to be on the high-energy side of the osmium absorption edge.

The number of molecules in the asymmetric unit was estimated by calculation of the Matthews coefficient and self-rotation function with *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 program package. For ClitOs1 and ClitIr1 data no peak in the self-rotation function was observed. The Matthews coefficient indicated the presence of one molecule in the asymmetric unit for ClitOs1 data, whereas the absence of a peak from the ClitIr1 data could indicate parallel orientation of the non-crystallographic and twofold crystallographic axes. The large peak in the self-rotation functions of isomorphous ClitIr2 and ClitOs2 data sets suggests that the two molecules in the asymmetric units of triclinic crystals are related by a non-crystallographic twofold symmetry axis, which is in agreement with the Matthews coefficient, suggesting approximately 40% solvent content for each crystal form.

Solution of the phase problem is in progress. The structure is expected to provide insights into the key determinants of the function of clitocypin.

We thank Wladek Minor, Maksymilian Chruszcz and Marcin Cymborowski for their assistance and support with data collection and phasing. We thank Hans D. Bartunik for access to the Max-Planck beamline (BW6) at Deutsches Elektronensynchrotron in Hamburg. We thank Gleb Bourenkov for his assistance during data collection and processing. We thank Maurizio Polentarutti and Alberto Cassetta for their assistance during data collection. We thank Roger H. Pain for critical reading of the manuscript. The Agency of Research of the Republic of Slovenia is acknowledged for financial support and for encouragement of publication.

References

Brzin, J., Rogelj, B., Popovič, T., Štrukelj, B. & Ritonja, A. (2000). *J. Biol. Chem.* **275**, 20104–20109.
Dubin, G. (2005). *Cell. Mol. Life Sci.* **62**, 653–669.
Dubin, G., Stec-Niemczyk, J., Dylag, T., Silberring, J., Dubin, A. & Potempa, J. (2004). *J. Biol. Chem.* **385**, 543–546.
Filipek, R., Rzychon, M., Oleksy, A., Gruca, M., Dubin, A., Potempa, J. & Bochtler, M. (2003). *J. Biol. Chem.* **278**, 40959–40966.
Galeša, K., Thomas, R. M., Kidrič, M. & Pain, R. H. (2004). *Biochem. Biophys. Res. Commun.* **324**, 576–578.
Gunčar, G., Pungerčič, G., Klemenčič, I., Turk, V. & Turk, D. (1999). *EMBO J.* **18**, 793–803.
Kidrič, M., Fabian, H., Brzin, J., Popovič, T. & Pain, R. H. (2002). *Biochem. Biophys. Res. Commun.* **297**, 962–967.
Križaj, I., Drobnič-Kosorok, M., Brzin, J., Jerala, R. & Turk, V. (1993). *FEBS Lett.* **333**, 15–20.
Lenarčič, B. & Bevec, T. (1998). *Biol. Chem.* **379**, 105–111.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Monteiro, A. C., Abrahamson, M., Lima, A. P., Vannier-Santos, M. A. & Scharfstein, J. (2001). *J. Cell Sci.* **114**, 3933–3942.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rawlings, N. D., Tolle, D. P. & Barrett, A. J. (2004). *Nucleic Acids Res.* **32**, D160–D164.
Rzychon, M., Sabat, A., Kosowska, K., Potempa, J. & Dubin, A. (2003). *Mol. Microbiol.* **49**, 1051–1066.
Sanderson, S. J., Westrop, G. D., Scharfstein, J., Mottram, J. C. & Coombs, G. H. (2003). *FEBS Lett.* **542**, 12–16.
Schick, C., Pemberton, P. A., Shi, G. P., Kamachi, Y., Cataltepe, S., Bartuski, A. J., Gornstein, E. R., Bromme, D., Chapman, H. A. & Silverman, G. A. (1998). *Biochemistry*, **37**, 5258–5266.
Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarčič, B. & Turk, V. (1990). *EMBO J.* **9**, 1939–1947.
Turk, B., Turk, D. & Salvesen, G. S. (2005). *Med. Chem. Rev. Online*, **2**, 283–297.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Yamamoto, Y., Watabe, S., Kageyama, T. & Takahashi, S. Y. (1999). *FEBS Lett.* **448**, 257–260.