Acta Crystallographica Section F Structural Biology and Crystallization Communications

ISSN 1744-3091

### Shunfu Piao, Ji Young Jung, Ji Won Park, Jaewon Lee, Bok Leul Lee and Nam-Chul Ha\*

College of Pharmacy and Research Institute for Drug Development, Pusan National University, Jangjeon-dong, Geumjeong-gu, Busan 609-735, South Korea

Correspondence e-mail: hnc@pusan.ac.kr

Received 30 March 2006 Accepted 4 July 2006



© 2006 International Union of Crystallography All rights reserved

# Preliminary X-ray crystallographic analysis of the catalytic domain of prophenoloxidase activating factor-I

Clip-domain serine proteases (SPs) have been identified in invertebrates as crucial enzymes that are involved in diverse extracellular signalling pathways. Prophenoloxidase (proPO) activating factor-I (PPAF-I), a catalytically active clip-domain SP, cleaves proPO. To date, no crystal structures of a catalytically active clip-domain SP have been determined. Here, the results of crystallization and preliminary X-ray analysis of the SP domain of PPAF-I are reported. The crystal of the PPAF-I SP domain was obtained using the hanging-drop vapour-diffusion method in a precipitant solution containing 0.15 *M* lithium sulfate, 30% polyethylene glycol 4000 and 0.1 *M* Tris–HCl pH 8.0. The crystal diffracts X-rays to 1.7 Å resolution using a synchrotron-radiation source. The crystal belongs to space group  $P2_12_12_1$ , with one molecule in the asymmetric unit and unit-cell parameters a = 38.3, b = 53.3, c = 116.6 Å,  $\alpha = \beta = \gamma = 90^{\circ}$ . A molecular-replacement solution has been found using kallikrein as a starting model, resulting in an interpretable electron-density map.

#### 1. Introduction

Both vertebrates and invertebrates employ serine protease (SP) cascades to mediate rapid and local reactions in response to diverse physiological or pathological stimuli in the extracellular region. Clipdomain SPs, which contain one or two clip domain(s) at the N-terminus of the SP domain, have been identified in the immune response and embryonic development of invertebrates. Easter from *Drosophila melanogaster* is a well characterized clip-domain SP and is expressed and secreted during oogenesis as an inactive zymogen (Anderson, 1998; Morisato & Anderson, 1995).

Clip-domain SPs can be divided into catalytic and non-catalytic groups according to their proteolytic activity. The non-catalytic group members do not exhibit any proteolytic activity owing to replacement of the serine residue at the active site by glycine. The overall structure of the SP domain of the non-catalytic group is similar to those of chymotrypsin-like SPs (Piao et al., 2005). In contrast, the SP domain of the catalytic group is distinct from those of chymotrypsin-like SPs. One feature of the SP domain of the catalytic group is the replacement of two cysteine residues (Cys137 and Cys202 in the chymotrypsin numbering) which form a disulfide-bond linkage by other amino acids (Fig. 1). Another distinctive feature is a short insertion between the His and Asp residues of the catalytic triad (Fig. 1). The short insertion contains two cysteine residues which are expected to form a disulfide-bond linkage. Moreover, the short insertion has been predicted to contain a calcium-binding site (Rose et al., 2003). The clip-domain SPs containing the short insertion are referred to as easter-type SPs because Drosophila easter exhibits the features mentioned above. However, no crystal structures of a catalytically active clip-domain SP have been determined.

Two easter-type SPs have been cloned and characterized from the large beetle *Holotrichia diomphalia* (Kwon *et al.*, 2000; Lee *et al.*, 1998). ProPO activating factors-I (PPAF-I) and III (PPAF-III) cleave proPO and PPAF-II, a non-catalytic group member, respectively (Kwon *et al.*, 2000; Lee *et al.*, 1998). Here, we report the crystallization and preliminary X-ray crystallographic analysis of the SP domain of PPAF-I from *H. diomphalia*.

#### 2. Materials and methods

#### 2.1. DNA construction and protein purification

Overexpression and purification of the SP domain of PPAF-I (residues 88–365) have been described previously (Piao *et al.*, 2005). The purified protein was concentrated to 5–10 mg ml<sup>-1</sup> for crystal-lization using an ultrafiltration device (Millipore).

#### 2.2. Crystallization, data collection and X-ray analysis

The crystals of the SP domain of PPAF-I were obtained by the vapour-diffusion technique at 287 K. Initial crystallization screening was performed by the sitting-drop method with Crystal Screen HT, a high-throughput sparse-matrix screening kit (Hampton Research, USA), in a 96-well plate using 0.5  $\mu$ l protein solution (10 mg ml<sup>-1</sup>) mixed with 0.5 µl of each precipitant solution. Crystals of various shapes were obtained in a few days in 12 of the 96 conditions. One condition (0.2 M lithium sulfate, 30% polyethylene glycol 4000, 0.1 M Tris-HCl pH 8.0), in which many tiny diamond-shaped crystals grew, was chosen for optimization (Fig. 2). The crystallization conditions were optimized to produce high-quality single crystals (0.2 mm in diameter; Fig. 2) in droplets containing 1 µl protein solution (5 mg ml<sup>-1</sup>) and 1  $\mu$ l precipitant solution consisting of 0.2 M lithium sulfate, 30% polyethylene glycol 4000, 0.1 M Tris-HCl pH 8.5. The droplets were equilibrated by the hanging-drop vapour-diffusion method against 1 ml of the same precipitant solution at 287 K for one week.

#### 2.3. Crystallographic data collection

For X-ray data collection, a single crystal was briefly immersed into precipitation solution containing 10% glycerol as a cryoprotective agent. The crystal was flash-frozen in a stream of nitrogen gas at 100 K. Diffraction data were collected from a single crystal on beamline 4A of Pohang Accelerator Laboratory (Korea) at a wavelength of 0.9794 Å using an ADSC Q210 CCD detector with an exposure time of 2 s, a rotation angle of 1° and a crystal-to-detector distance of 130 mm. Diffraction was observed to a maximum resolution of 1.6 Å; however, data beyond 1.7 Å were weak and were not included in the processing. A complete data set was indexed, processed and scaled with *DENZO* and *SCALEPACK* from the *HKL*-2000 program suite (Otwinowski & Minor, 1997).

#### 3. Results and discussion

Crystallization of the full-length PPAF-I was unsuccessful despite extensive screening of crystallization conditions. We suspected that free motion between the two domains of PPAF-I prevents the crystallization of full-length PPAF-I. In order to obtain crystals, we separately expressed the clip and SP domains of PPAF-I according to the boundary obtained from the proteolysis experiment reported previously (Piao *et al.*, 2005). The SP domain of PPAF-I was successfully overexpressed for crystallization screening. However, crystallization was not successful for the clip domain of PPAF-I.

Preliminary analysis of the X-ray diffraction data using the autoindexing routine in *DENZO* (Otwinowski & Minor, 1997) indicated that the crystal belongs to the primitive system, point group 222, with

Hd-PPAF-I	(100 a.a including	g a clip do	omain)CGYQ	VEADK II	LNG DDTV PEE	FPWTAMIGYKN		SLINNRYIVT	AAHCVAGRVLR-
Dm-PAE-I	(127 a.a including	g a clip do	omain)CGPH	SFSNK VI	NGNDTAIDE	FNWMALLEYVD	NRGRR-ELSCGG	SLINNRYVLTA	AAHCVIGAVET-
Dm-easter	(119 a.a including	g a clip do	omain)CG-N	ILSNR IN	GGMKTKIDE	EFFWMALIEYTK		SLISTRYVITA	ASHCVNGKALP-
Dm-SPE	(126 a.a including	g a clip do	omain)CG-F	LFADR II	GGTNTTLWE	EFPWMVLLQYKK	LFSETYTFNCGG	ALLNSRYVLTA	AGICLASRELDK
Dm-Snake	(178 a.a including	a clip do	main)CV	PSVPL IV	GGTPTRHGI	FPHMAALGWTQGS	GSKDQDIKWGCGG	ALVSELYVLTZ	AAHCATSGSKP-
Hd-PPAF-II	(138 a.a including	g a clip do	main)CGIRNE	RGLDFK IT	GQTNEAEYGEE	PWMVAVLKAN-VI	PG SGEEQLV-CGG	SLIAPSVVLTO	GAHCVNSYQSN
Bt-trypsin			FIFLALLGAAV	AF PVDDDDK IV	GGYTCGANI	VPYQVSLN	-SGYHFCGG	SLINSQWVVSZ	AAHC-YKSGIQ-
Bt-chymo			CGVPAI	QPVLSGLSR IV	NGEEAVPGS	SWPWQVSLQDK	-TGFHECGG	SLINENWVVT/	AAHCGVTTSDV-
			1	11 "	21	31	41	51	61
		S-S						+	and the second
Hd-PPAF-I	VVGALNKVRLGEWNTAT	DPDCYGAVR-	VCVPDKPIDLGI	EETIQHPDYVDG	SKDRYHDIALI	IRLNRQVEFTNYI	RPVCLPQPN-EE	VQVGQRLTVVG	SWGRTETGQ
Dm-PAE-1	EVGHLIIVRLGEYDISE		CN-QPILQLGI	EQAIVHPQYDPA	NKNRIHDIALI	SRLDRPVVLNEYIG	2PVCLPLVS-IRMA	INIGELLVVSG	SWGRIITAR
Dm-easter	I DWRLSG VRLGEWDINI	NPDCEVDVRG	MKDCA-PPHLDVPV	ERIIPHPDYIPA	ASKNOVNDIALI	RLAQQVEYIDEVE	CPICLPLDVNLRSA	TEDGIIMDVAG	SWGKIEQLS
Dm-SPL	5 GAVLES VELGEWDIK	DPDGIIQMNG	QRIGA-PRHIDIEV.	EKGIIHEMIAPI	15V DQRINDIALV	REKRIVSIIDIVE	CELCTEIDG-PACK	NEVDIGMDVAG	SWGLIENMQ
Dm-Snake	PDMVRLGAROLNE	TSATO	ODIKI	LIIVLHPKYRSS	SAYYHDIALL	KLTRRVKFSEOVE	RPACLWOLP	ELOIPTVVAA	GWGRT-E-FLGA
Hd-PPAF-II	LDAIKIRAGEWDTLT	EKERL	PYOERKI	ROVIIHSNENPP	TVVNDVALL	LLDRPLVQADNIG	TICLPOOSO	I FOSTEC FASO	GWGKKEFGSRHR
Bt-tyrpsin	VRLGEDNIN	VEGNE	QFISA	SKSIVHPSYNSM	TLNNDIMLI	KLKSAASLNSRVA	SISLPTSCA	-SAGTQCLISC	GWGNTKS-SGTS
Bt-chymo	VVAGEF DQGS	SSEKI	OKLKI	AKVFKNSKYNSI	TINNDITLI	KLSTAASFSOTVS	AVCLPSASDD	FAAGTTCVTTC	GWGLTRY-TNAN
	71		81	91	101	111	121	131 1	141 151
					•		_		
Hd-PPAF-I	YSTIKQKLAVPVVHAEQ	AKTFGAAG	VRVRSSQLCAGG	-EKAKDSCGGD	GGPLLAEF	RANQQFFLEGLVSH	GAT-CGTEGWPGI	YTKVGKYRDW.	IEGNIRP
Dm-PAE-I	KSTIKQRLDLPVNDHDYC	ARKFATRN	IHLISSQLCVGG	-E FYRDSCDGDS	GGPLMRRG	FDQAWYQEGVVSH	GNR-CGLE GWPGV	YTRVADYMDW.	IVETIRP
Dm-easter	ASNLKLKAAVEGSRMDEC	QNVYSSQD	ILLEDTQMCAGG	-KEGVDSCRGD	GGPLIGLDINK	WNT YYFLAGVVSH	GPTPCGLAGWPGV	YTLVGKYVDW.	IQNTIES
Dm-SPE	PSAIKLKITVNVWNLTSC	QEKYSSFK	VKLDDSQMCAGG	-QLGVDTCGGD	GG PLMVPI STG	GRDVFYIAGVIS	GTKPCGLKGWPGV	YTRIGAFIDWI	IKQKLEP
Dm-Snake	KSNALRQVDLDVVPQMT	KQIYRKERR-	LPRGIIEGQFCAGY	LPGGRDTCQGD	GG PIHALL P-E	YNCVAFVVGITS	GKF-CAAPNAPGV	YTRLYSYLDW	IEKIAFKQH
Hd-PPAF-II	YSNILKKIQLPTVDRDKC	QADLENTELG	SLKFVLDQTFVCAGG	-EQGKDTCTGDO	GSPLFCPD-PF	RNPSRYMQMG IVAN	GIG-CGDENVPGV	YANVAHERNW	IDQEMQAKGLSTTPYV
Bt-trypsin	YPDVLKCLKAPILSDSSC	KSAYPG	QITSNMFCAGY	LEGGKDSCQGD	GG PVVCS	GKLQGIVSV	GSG-CAQKNKPGV	YTKVCNYVSW	IKQTIASN
Bt-chymo	TPDRLQQASLPLLSNTNC	KKYWGT	KIKDAMICAG-	-ASGVSSCMGD	GG PLVCKK	NGAWTLVGIVS	GSSTCS-TSTPGV	YARVTALVNW	VQQTLAAN
1000 million (2010)	161	171	181	191	201	211	221	231	241
Figure 1									
Sequence align	ment of the SP domains	of clin domai	n SPe with trypein	and chymotryne	in Four easter	r type SPc PPAF	II and howing try	nein and chyn	otrypsin (from top t

Sequence alignment of the SP domains of clip-domain SPs with trypsin and chymotrypsin. Four easter-type SPs, PPAF-II and bovine trypsin and chymotrypsin (from top to bottom; Hd, *H. diomphalia*; Dm, *Drosophila melanogaster*; Bt, *Bos taurus*; chymo, chymotrypsin) are shown. The alignment was performed using *ClustalW* (Thompson *et al.*, 1997) and then adjusted based on the conserved cysteine residues (shaded in cyan). The cleavage sites for zymogen activation are indicated by double lines and the catalytic triads (His, Asp, Ser) at the active sites are highlighted. The calcium-coordinating residues of PPAF-II and trypsin are underlined. The easter-type SP-specific insertion, indicated by a rectangle, contains two cysteine residues forming a putative disulfide bond. An invariant disulfide-bond linkage of SPs (Cys137 and Cys202 of chymotrypsin) is absent from all the catalycally active clip-domain SPs and is marked with diamonds. Every ten amino acids of chymotrypsin are underlined with a residue number.

## crystallization communications



Figure 2

A crystal of the SP domain of PPAF-I. Approximate dimensions are  $0.2 \times 0.1 \times 0.1$  mm. Eight divisions on the scale represent 0.1 mm.

unit-cell parameters a = 38.3, b = 53.3, c = 116.6 Å,  $\alpha = \beta = \gamma = 90^{\circ}$ . Analysis of the X-ray diffraction pattern showed that along the *h*, *k* and *l* axes reflections were only present if *h*, *k* and *l* = 2*n*, identifying the space group as  $P2_12_12_1$ . The data set was 98.4% complete in the resolution range 100–1.70 Å, with  $R_{\text{merge}} = 6.7\%$ . The crystals contain one molecule in the asymmetric unit, corresponding to a Matthews coefficient ( $V_{\text{M}}$ ) of 1.9 Å<sup>3</sup> Da<sup>-1</sup> and a solvent content of 34.4% (Matthews, 1968). The crystallographic parameters and data-collection statistics are summarized in Table 1.

To date, many crystal structures of serine proteases have been deposited in the PDB. Kallikrein, a plasma serine protease (PDB code 2any; Tang et al., 2005), was chosen as a search model for molecular replacement (MR) because it shares the highest sequence identity (34%) with the SP domain of PPAF-I. The MR procedure was performed with the program AMoRe (Navaza, 2001). The rotation with the highest correlation coefficient was applied to the search model and was used in the subsequent translation-function calculation. The calculation of the translation function gave one peak with a correlation coefficient of 24.8%, while the next solution exhibited a correlation of 20.5%. Rigid-body refinement with the best solution yielded a correlation coefficient of 35.5% and an R factor of 51.5% in the resolution range 10-3.5 Å. The solution produced an interpretable electron-density map, although it gave a relatively high Rfactor. The atomic model was refined using the program CNS to an  $R_{\rm free}$  of below 30%. From the present model, the unique short insertion containing two cysteine residues is visible in the electrondensity map, which will provide a clue as to how the insertion is involved in the function of easter-type SPs.

In conclusion, we obtained a high-quality crystal of the SP domain of PPAF-I and we are refining the structure of the SP domain of PPAF-I using the diffraction data set from the crystal. The crystal

#### Table 1

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	Beamline 4A, Pohang Accelerator Laboratory				
Wavelength (Å)	0.9794				
Temperature (K)	100				
Unit-cell parameters (Å, °)	a = 38.3, b = 53.3, c = 116.6,				
	$\alpha = \beta = \gamma = 90$				
Space group	$P2_{1}2_{1}2_{1}$				
$V_{\rm M}$ (Å <sup>3</sup> Da <sup>-1</sup> )	1.9				
Solvent content (%)	34.4				
Resolution (Å)	100-1.70 (1.76-1.70)				
Unique reflections	27064				
Redundancy	7.7 (4.3)				
Completeness (%)	98.4 (96.0)				
$R_{\text{merge}}$ (%)	6.7 (25.8)				
Average $I/\sigma(I)$	34.0 (3.0)				

structure of PPAF-I will serve as a representative model of eastertype SPs to elucidate the molecular mechanism by which the clip-domain SPs recognize and catalyze the substrates in various biological processes.

We thank the staff members at beamline 4A of Pohang Accelerator Laboratory (Korea) for the data collection. This project was supported by Programs of the National Research Laboratory (M10400000028-04J0000-02) grants to BLL and N-CH from the Korea Ministry of Science and Technology. This research was partly supported by Korea Research Foundation Grant (KRF-2004-041-C00247) to N-CH and Pusan National University Research Grant 2004 to N-CH.

#### References

- Anderson, K. V. (1998). Cell, 95, 439-442.
- Kwon, T. H., Kim, M. S., Choi, H. W., Joo, C. H., Cho, M. Y. & Lee, B. L. (2000). Eur. J. Biochem. 267, 6188–6196.
- Lee, S. Y., Cho, M. Y., Hyun, J. H., Lee, K. M., Homma, K. I., Natori, S., Kawabata, S. I., Iwanaga, S. & Lee, B. L. (1998). *Eur. J. Biochem.* 257, 615– 621.
- Matthews, B. W. (1968). J. Mol. Biol. 33, 491-497.
- Morisato, D. & Anderson, K. V. (1995). Annu. Rev. Genet. 29, 371-399.
- Navaza, J. (2001). Acta Cryst. D57, 1367-1372.
- Otwinowski, Z. & Minor, W. (1997). Methods Enzymol. 276, 307-326.
- Piao, S., Song, Y. L., Kim, J. H., Park, S. Y., Park, J. W., Lee, B. L., Oh, B.-H. & Ha, N.-C. (2005). *EMBO J.* 24, 4404–4414.
- Rose, T., LeMosy, E. K., Cantwell, A. M., Banerjee-Roy, D., Skeath, J. B. & Di Cera, E. (2003). J. Biol. Chem. 278, 11320–11330.
- Tang, J., Yu, C. L., Williams, S. R., Springman, E., Jeffery, D., Sprengeler, P. A., Estevez, A., Sampang, J., Shrader, W., Spencer, J., Young, W., McGrath, M. & Katz, B. A. (2005). J. Biol. Chem. 280, 41077–41089.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). Nucleic Acids Res. 25, 4876–4882.