crystallization communications

Michikazu Tanio,[a] Shin Kondo,[b] Shigetoshi Sugio[b]* and Toshiyuki Kohno[a]*

[a]Mitsubishi Kagaku Institute of Life Sciences (MITILS), Minamiooya 11, Machida, Tokyo 194-8511, Japan, and [b]ZOEGENE Corporation, 1000 Kamoshida, Aoba, Yokohama 227-8502, Japan

Correspondence e-mail:
ssugio@rc.m-kagaku.co.jp,
tkohno@libra.ls.m-kagaku.co.jp

Overexpression, purification and preliminary crystallographic analysis of human M-ficolin fibrinogen-like domain

Ficolins, which are comprised of a collagen-like domain and a fibrinogen-like domain, are a kind of pattern-recognition molecule for pathogens in the innate immunity system. To investigate the molecular mechanism of the discrimination between self and non-self by ficolins, human M-ficolin fibrinogen-like domain (FD1), which contains the ligand-binding site, was overexpressed in *Pichia pastoris*, purified and crystallized using the vapour-diffusion method at 293 K. The crystals belong to the monoclinic space group $P2_1$, with unit-cell parameters $a = 55.16$, $b = 117.45$, $c = 55.19$ Å, $\beta = 99.88°$, and contain three molecules per asymmetric unit. An X-ray data set was collected to 1.9 Å resolution using synchrotron radiation at beamline BL24XU at the SPring-8 facility in Japan.

1. Introduction

Innate immunity systems are present in all multicellular organisms, including humans, and play a crucial role in the first line of defence against pathogens. Ficolins are one of the most important groups of pattern-recognition molecules in the innate immunity system and have been identified in both vertebrates and invertebrates (Fujita *et al.*, 2004). Ficolins are composed of a collagen-like domain at the N-terminus and a fibrinogen-like domain, which is the sugar-binding site, at the C-terminus (Le *et al.*, 1998; Teh *et al.*, 2000). Three human ficolins, L-ficolin and H-ficolin in serum and M-ficolin in cells, have been characterized. These ficolins are associated with the mannose-binding lectin-associated serine proteases and the complexes activate the lectin-complement pathway (Fujita *et al.*, 2004; Liu *et al.*, 2005; Frederiksen *et al.*, 2005). L-ficolin (synonymous with ficolin-2 or ficolin/P35) binds to *N*-acetyl-D-glucosamine (GlcNAc; Matsushita *et al.*, 1996; Le *et al.*, 1997) and *N*-acetyl-D-galactosamine (GalNAc; Le *et al.*, 1998). The binding ability is inhibited by acetylated compounds, indicating that this protein specifically recognizes acetyl groups (Krarup *et al.*, 2004). L-ficolin activates the lectin-complement pathway upon binding to lipoteichoic acid, a cell component found in all Gram-positive bacteria (Lynch *et al.*, 2004). H-ficolin (synonymous with ficolin-3 or Hakata-antigen), the primary structure of which is 48% identical to that of L-ficolin, binds to GlcNAc, GalNAc and D-fucose (Sugimoto *et al.*, 1998). Unlike L-ficolin, the GlcNAc-binding activity of H-ficolin is not inhibited by acetyl compounds (Krarup *et al.*, 2004). M-ficolin (synonymous with ficolin-1 or ficolin/P35-related protein), the primary structure of which is 80 and 48% identical to those of L-ficolin and H-ficolin, respectively (Endo *et al.*, 1996; Sugimoto *et al.*, 1998), binds to GlcNAc, GalNAc and sialic acid (Teh *et al.*, 2000; Liu *et al.*, 2005). This protein also recognizes acetyl groups, like L-ficolin (Frederiksen *et al.*, 2005). Unlike the serum ficolins, M-ficolin has been detected on the surfaces of peripheral blood monocytes and promonocytic U937 cells (Lu *et al.*, 1996; Teh *et al.*, 2000; Frederiksen *et al.*, 2005). An antibody against the recombinant M-ficolin fibrinogen-like domain inhibits the phagocytosis of *Escherichia coli* by U937 cells, suggesting that M-ficolin acts as a phagocytic receptor or adaptor on circulating monocytes for pathogens (Teh *et al.*, 2000). Recently, M-ficolin has been reported to be
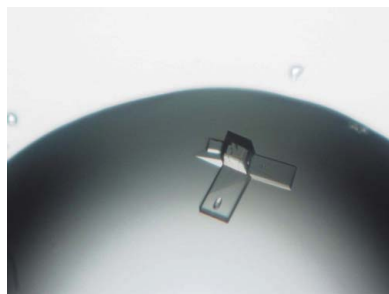
located in secretory granules in the cytoplasm of neutrophils, monocytes and type II alveolar epithelial cells in the lung (Liu *et al.*, 2005).

The ligand-binding site on the fibrinogen-like domain of ficolins is not yet known. However, the crystal structures of human fibrinogen γ fragment (Pratt *et al.*, 1997) and of the fibrinogen-like domain of *Tachypleus tridentatus* tachylectin-5A (Kairies *et al.*, 2001) have shown that the P domain (Fig. 1) contributes to ligand binding. In angiopoietin-2, it was found that the P domain of the fibrinogen-like domain relates to receptor binding (Barton *et al.*, 2005). These findings suggest that the P domain of ficolins includes the sugar-binding site. Interestingly, although the sugars recognized by ficolins exist on the surface of the host cell, these proteins can discriminate between pathogens and the host cell. Since the detailed mechanism of the discrimination between self and non-self remains unclear, the crystal structures of ficolins complexed with ligands should clarify the molecular mechanism of the discrimination and advance drug design for general pathogens. Here, we report the overexpression, crystallization and preliminary crystallographic analysis of the M-ficolin fibrinogen-like domain (FD1).

2. Cloning, overexpression and purification of FD1

The cDNA encoding FD1, corresponding to residues 115–326 of M-ficolin (NCBI database accession No. O00602), was amplified using a human cDNA library (Clontech) as a template with a forward (5′-GAATTCCCACGCAACTGCAAGGACCTGC-3′) primer and a reverse (5′-TCTAGAAAGGCGGGCCGCACCTTCATC-3′) primer. The *Eco*RI to *Xba*I fragment containing the FD1 gene was cloned in the *Eco*RI and *Xba*I sites of the pPICZαA vector (Invitrogen). This resulted in a vector containing in frame the *Saccharomyces cerevisiae* α-factor secretion signal sequence (89 amino acids), the FD1 gene and both a *c-myc* epitope and a six-histidine tag originally present in pPICZαA. To generate multiple copies of the FD1 expression cassette, the *Eco*RV to *Bam*HI and the *Bgl*II to *Eco*RV fragments containing the FD1 gene of the plasmid were ligated, resulting in a vector containing two FD1 expression cassettes. Finally, a plasmid containing four FD1 expression cassettes was made in the same manner and was electroporated into the yeast *Pichia pastoris* strain X-33 (Invitrogen). The protein expressed from this construct undergoes processing of the α-factor signal sequence, resulting in the secretion of the mature FD1 protein [237 amino acids consisting of an EF sequence derived from the *Eco*RI site at the N-terminus, residues 115–326 of M-ficolin and a FLEQKLISEEDLNSAVDHHHHHH sequence (23 residues) including the *c-myc* epitope and the histidine tag at the C-terminus]. The transformed yeast cells were initially grown for 72 h at 298 K in 1 l BMGY medium containing 1%(w/v) yeast extract, 2%(w/v) peptone, 0.1 M potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base with ammonium sulfate without amino acids, 4×10^{-5}%(w/v) biotin and 1%(v/v) glycerol. For the expression of FD1, the harvested cells were resuspended in 100 ml BMMY medium [the same as BMGY except that the glycerol was replaced by 0.5%(v/v) methanol] and cultured for 48–54 h with the addition of 0.5% methanol every 24 h. The secreted FD1 was isolated from the cell-culture supernatant. The supernatant containing FD1
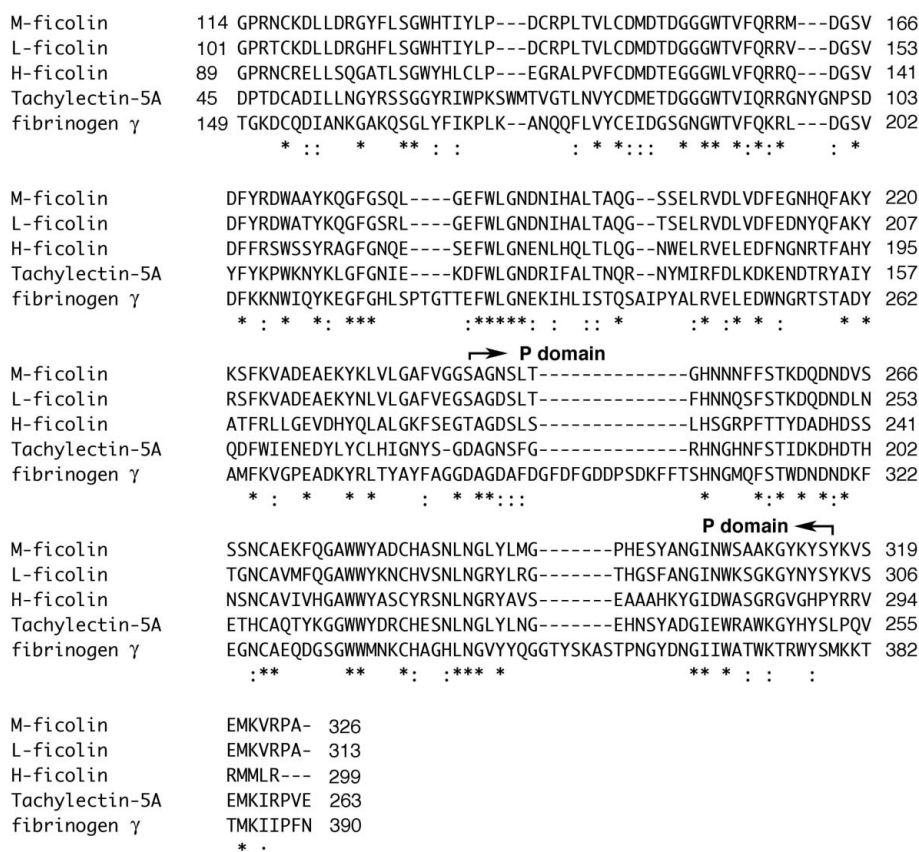
```
M-ficolin       114 GPRNCKDLLDRGYFLSGWHTIYLP---DCRPLTVLCDMDTDGGGWTVFQRRM---DGSV 166
L-ficolin       101 GPRTCKDLLDRGHFLSGWHTIYLP---DCRPLTVLCDMDTDGGGWTVFQRRV---DGSV 153
H-ficolin        89 GPRNCRELLSQGATLSGWYHLCLP---EGRALPVFCDMDTEGGGWLVFQRRQ---DGSV 141
Tachylectin-5A   45 DPTDCADILLNGYRSSGGYRIWPKSWMTVGTLNVYCDMETDGGGWTVIQRRGNYGNPSD 103
fibrinogen γ    149 TGKDCQDIANKGAKQSGLYFIKPLK--ANQQFLVYCEIDGSGNGWTVFQKRL---DGSV 202
                     * ::   *   ** : :           : * *::: * ** *:*:*    : *

M-ficolin           DFYRDWAAYKQGFGSQL----GEFWLGNDNIHALTAQG--SSELRVDLVDFEGNHQFAKY 220
L-ficolin           DFYRDWATYKQGFGSRL----GEFWLGNDNIHALTAQG--TSELRVDLVDFEDNYQFAKY 207
H-ficolin           DFFRSWSSYRAGFGNQE----SEFWLGNENLHQLTLQG--NWELRVELEDFNGNRTFAHY 195
Tachylectin-5A      YFYKPWKNYKLGFGNIE----KDFWLGNDRIFALTNQR--NYMIRFDLKDKENDTRYAIY 157
fibrinogen γ        DFKKNWIQYKEGFGHLSPTGTTEFWLGNEKIHLISTQSAIPYALRVELEDWNGRTSTADY 262
                     * : *  *: ***       :*****: : :: *      :* :* * :     * *

                                          ┌──▶ P domain
M-ficolin           KSFKVADEAEKYKLVLGAFVGGSAGNSLT--------------GHNNNFFSTKDQDNDVS 266
L-ficolin           RSFKVADEAEKYNLVLGAFVEGSAGDSLT--------------FHNNQSFSTKDQDNDLN 253
H-ficolin           ATFRLLGEVDHYQLALGKFSEGTAGDSLS--------------LHSGRPFTTYDADHDSS 241
Tachylectin-5A      QDFWIENEDYLYCLHIGNYS-GDAGNSFG--------------RHNGHNFSTIDKDHDTH 202
fibrinogen γ        AMFKVGPEADKYRLTYAYFAGGDAGDAFDGFDFGDDPSDKFFTSHNGMQFSTWDNDNDKF 322
                     * :  *   * *         : * **::*                *    *:* * *:*

                                                        P domain ◀──┐
M-ficolin           SSNCAEKFQGAWWYADCHASNLNGLYLMG-------PHESYANGINWSAAKGYKYSYKVS 319
L-ficolin           TGNCAVMFQGAWWYKNCHVSNLNGRYLRG-------THGSFANGINWKSGKGYNYSYKVS 306
H-ficolin           NSNCAVIVHGAWWYASCYRSNLNGRYAVS-------EAAAHKYGIDWASGRGVGHPYRRV 294
Tachylectin-5A      ETHCAQTYKGGWWYDRCHESNLNGLYLNG-------EHNSYADGIEWRAWKGYHYSLPQV 255
fibrinogen γ        EGNCAEQDGSGWWMNKCHAGHLNGVYYQGGTYSKASTPNGYDNGIIWATWKTRWYSMKKT 382
                     :**     **   *: :*** *           ** * : :   :

M-ficolin           EMKVRPA- 326
L-ficolin           EMKVRPA- 313
H-ficolin           RMMLR--- 299
Tachylectin-5A      EMKIRPVE 263
fibrinogen γ        TMKIIPFN 390
                     * :
```

Figure 1
Alignment of fibrinogen-like domains of three human ficolins, *T. tridentatus* tachylectin-5A and human fibrinogen γ fragment. Identical and conservatively substituted amino acids are marked with a star and a colon, respectively. The alignment was produced using *ClustalW* (http://www.ebi.ac.uk/clustalw).

crystallization communications

was adjusted to pH 7.8–8.2 and was loaded onto Ni–NTA agarose (3–5 ml; Qiagen) previously equilibrated with a wash buffer containing 50 mM Tris–HCl pH 8.0 and 300 mM NaCl. After washing, FD1 was eluted with the wash buffer containing 300 mM imidazole. The eluted protein was diluted threefold with 50 mM Tris–HCl pH 8.0 and was further purified by anion-exchange chromatography (Resource-Q; Amersham Biosciences) in 50 mM Tris–HCl pH 8.0. FD1 was eluted in a single peak using an NaCl gradient from 25 to 500 mM in the same buffer. The pooled fractions containing FD1 were concentrated and loaded onto a gel-filtration column (Superdex 200HR; Amersham Biosciences) equilibrated with 10 mM Tris–HCl pH 8.0, 100 mM NaCl and 5 mM CaCl$_2$. The purified protein (Fig. 2) was concentrated to 10 mg ml^{-1} and was stored at 193 K. About 5–8 mg of FD1 were purified from 1 l of culture by this method. The
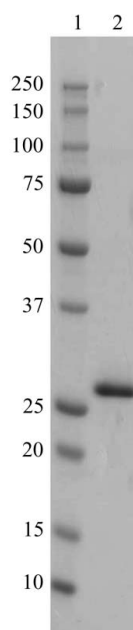
Figure 2
SDS–PAGE of purified FD1 on a 10–20% gradient gel. Lanes 1 and 2 contain molecular-weight markers (labelled in kDa) and FD1 (26.8 kDa), respectively. Proteins were stained with Coomassie blue.
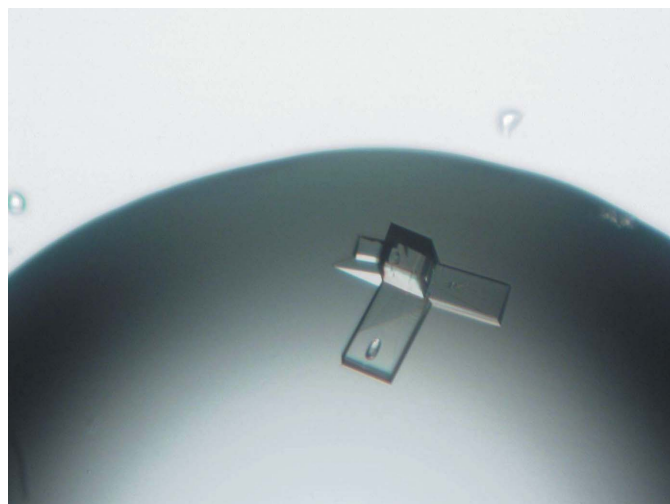

Figure 3
Crystal of FD1. The size of the single quadratic prismatic crystal is approximately 0.1 × 0.1 × 0.15 mm.

Table 1
Data-collection statistics for FD1.

Values in parentheses are for the highest resolution shell.

X-ray source	SPring-8, BL24XU
Detector	Rigaku R-AXIS V IP
Wavelength (Å)	0.82656
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 55.16$, $b = 117.45$, $c = 55.19$, $\beta = 99.88$
Resolution range (Å)	30–1.9 (1.97–1.9)
Measured reflections	197686
Unique reflections	53715
Redundancy	3.7 (3.7)
Completeness (%)	98.8 (98.3)
R_{sym}† (%)	6.2 (25.6)
$\langle I/\sigma(I)\rangle$	34.8 (6.8)

† $R_{sym} = \sum_h \sum_i |I(h)_i - \langle I\rangle| / \sum_h \sum_i I(h)_i$ where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection h.

purified FD1 shows GlcNAc-binding activity (data not shown), indicating that the recombinant protein is correctly folded.

3. Crystallization

To form the ligand–protein complex, a protein solution containing 8 mg ml^{-1} FD1, 8 mM Tris–HCl pH 8.0, 80 mM NaCl, 4 mM CaCl$_2$ and 20 mM GlcNAc was prepared by addition of GlcNAc to the purified protein solution. Screening for crystallization conditions was performed using the sitting-drop vapour-diffusion method in 96-well plates of CrystalClear strips (Hampton Research) at 278 and 293 K by mixing 0.5 µl protein solution and 0.5 µl reservoir solution and equilibrating against 50 µl reservoir solution. Initial screens included Crystal Screens I and II (Hampton Research), Wizard Screens I and II (Emerald Biostructures) and Cryo Screens I and II (Emerald Biostructures). Crystallization conditions were further optimized in a 96-well plate that was turned upside down to form the conditions for the hanging-drop method. Diffraction-quality FD1 crystals were obtained within a week at 293 K as clustered quadratic prisms (Fig. 3) using a reservoir solution containing 100 mM MES pH 5.6, 320 mM Li$_2$SO$_4$ and 17%(w/v) PEG 4000.

4. X-ray data collection

A single quadratic prismatic crystal (approximately 0.1 × 0.1 × 0.15 mm) was separated from the cluster, soaked in the reservoir solution with 15%(v/v) glycerol as a cryoprotectant for a few seconds and then flash-frozen in a nitrogen stream. X-ray diffraction images were collected from the crystal at 100 K under a nitrogen stream using a Rigaku R-AXIS V image plate with 0.82656 Å synchrotron radiation at SPring-8 beamline BL24XU at Hyogo, Japan. The distance between the crystal and the detector was 350 mm. 1° oscillation images were recorded with an exposure times of 48 s. The diffraction data were processed and scaled with the *HKL* program package (Otwinowski & Minor, 1997). The diffraction data were collected to 1.9 Å resolution with an R_{sym} of 0.062 and a completeness of 98.8% (Table 1). The crystals contain three molecules per asymmetric unit. Molecular replacement with *EPMR* (Kissinger *et al.*, 1999) using the crystal structure of tachylectin-5A (PDB code 1jc9; Kairies *et al.*, 2001), which shows 49% identity with FD1, was successful (PDB code 2d39; Tanio *et al.*, 2006). However, no obvious density for GlcNAc, which was added to the protein solution for crystallization, was observed in the current model. Model fitting and

further refinement, as well as a search for other crystallization conditions for the GlcNAc-binding form, are in progress.

References

Barton, W. A., Tzvetkova, D. & Nikolov, D. B. (2005). *Structure*, **13**, 825–832.

Endo, Y., Sato, Y., Matsushita, M. & Fujita, T. (1996). *Genomics*, **36**, 515–521.

Frederiksen, P. D., Thiel, S., Larsen, C. B. & Jensenius, J. C. (2005). *Scand. J. Immunol.* **62**, 462–473.

Fujita, T., Matsushita, M. & Endo, Y. (2004). *Immunol. Rev.* **198**, 185–202.

Kairies, N., Beisel, H. G., Fuentes-Prior, P., Tsuda, R., Muta, T., Iwanaga, S., Bode, W., Huber, R. & Kawabata, S. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 13519–13524.

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.

Krarup, A., Thiel, S., Hansen, A., Fujita, T. & Jensenius, J. C. (2004). *J. Biol. Chem.* **279**, 47513–47519.

Le, Y., Lee, S. H., Kon, O. L. & Lu, J. (1998). *FEBS Lett.* **425**, 367–370.

Le, Y., Tan, S. M., Lee, S. H., Kon, O. L. & Lu, J. (1997). *J. Immunol. Methods*, **204**, 43–49.

Liu, Y., Endo, Y., Iwaki, D., Nakata, M., Matsushita, M., Wada, I., Inoue, K., Munakata, M. & Fujita, T. (2005). *J. Immunol.* **175**, 3150–3156.

Lu, J., Le, Y., Kon, O. L., Chan, J. & Lee, S. H. (1996). *Immunology*, **89**, 289–294.

Lynch, N. J., Roscher, S., Hartung, T., Morath, S., Matsushita, M., Maennel, D. N., Kuraya, M., Fujita, T. & Schwaeble, W. J. (2004). *J. Immunol.* **172**, 1198–1202.

Matsushita, M., Endo, Y., Taira, S., Sato, Y., Fujita, T., Ichikawa, N., Nakata, M. & Mizuochi, T. (1996). *J. Biol. Chem.* **271**, 2448–2454.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pratt, K. P., Cote, H. C., Chung, D. W., Stenkamp, R. E. & Davie, E. W. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 7176–7181.

Sugimoto, R., Yae, Y., Akaiwa, M., Kitajima, S., Shibata, Y., Sato, H., Hirata, J., Okochi, K., Izuhara, K. & Hamasaki, N. (1998). *J. Biol. Chem.* **273**, 20721–20727.

Tanio, M., Kondo, S., Sugio, S. & Kohno, T. (2006). Submitted.

Teh, C., Le, Y., Lee, S. H. & Lu, J. (2000). *Immunology*, **101**, 225–232.