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

Purification, crystallization and preliminary structural characterization of the N-terminal region of the human formin-homology protein FHOD1

Antje Schulte,* Alexey Rak, Olena Pylypenko, Diana Ludwig and Matthias Geyer

Max-Planck-Institut für Molekulare Physiologie, Abteilung Physikalische Biochemie, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany

Correspondence e-mail: antje.schulte@mpi-dortmund.mpg.de

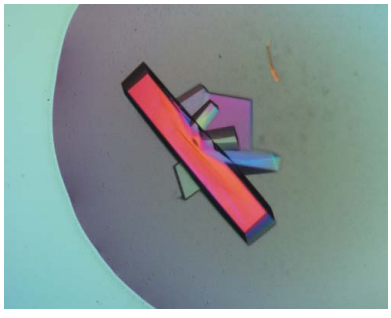
Received 12 July 2007
Accepted 4 September 2007

Formins are key regulators of actin cytoskeletal dynamics that constitute a diverse protein family that is present in all eukaryotes examined. They typically consist of more than 1000 amino acids and are defined by the presence of two conserved regions, namely the formin homology 1 and 2 domains. Additional conserved domains comprise a GTPase-binding domain for activation, a C-terminal autoregulation motif and an N-terminal recognition domain. In this study, the N-terminal region (residues 1–339) of the human formin homology domain-containing protein 1 (FHOD1) was purified and crystallized from 20%(*w*/*v*) PEG 4000, 10%(*v*/*v*) glycerol, 0.3 *M* magnesium chloride and 0.1 *M* Tris–HCl pH 8.0. Native crystals belong to space group *P*1, with unit-cell parameters $a = 35.4$, $b = 73.9$, $c = 78.7$ Å, $\alpha = 78.2$, $\beta = 86.2$, $\gamma = 89.7^\circ$. They contain two monomers of FHOD1 in the asymmetric unit and diffract to a resolution of 2.3 Å using a synchrotron-radiation source.

1. Introduction

Formin proteins are involved in the regulation of actin cytoskeletal rearrangements including cytokinesis, actin-cable and stress-fibre formation, polarity establishment, neurite outgrowth and intracellular trafficking (Faix & Grosse, 2006; Goode & Eck, 2007). Most eukaryotes have multiple formin isoforms, in line with their diverse cellular roles. Formins mediate actin polymerization by their ability to promote F-actin assembly at the barbed end of the filament and to move processively with the barbed end as it elongates (Otomo, Tomchick *et al.*, 2005). They are defined by a unique and highly conserved C-terminal formin homology 2 (FH2) domain of about 400 amino-acid residues that is preceded by an N-terminal proline-rich FH1 domain. The FH2 domain is necessary, and in some formins sufficient, to nucleate actin polymerization from G-actin *in vitro*, but binding of profilin to isolated FH1-FH2 fragments increases the elongation rate (Romero *et al.*, 2004; Vavylonis *et al.*, 2006). The presence of a C-terminal diaphanous-autoregulation domain (DAD) was found to define a formin subfamily, the diaphanous-related formins (DRFs), that are activated by Rho-family GTPases (Watanabe *et al.*, 1999; Alberts, 2001). Here, direct intramolecular contacts between the N- and C-terminal domains led to an autoinhibited state of the DRF that is believed to be released by structural rearrangements upon binding to the activated GTPase. The regulation mechanism of formin autoinhibition and activation is as yet only partially understood.

The DRF FHOD1 (formin homology 2 domain-containing protein 1) was identified as an interaction partner of the acute myeloid leukaemia transcription factor that is ubiquitously expressed and facilitates transcription from the serum response element (Westendorf, 2001). An activated form of FHOD1 in which autoinhibition is constitutively released induces the formation of and association with actin stress fibres *in vivo* (Gasteier *et al.*, 2003). Activated FHOD1 acts on Rho signalling downstream of the GTPase at the level of the Rho effector kinase ROCK, indicating that FHOD1 might serve as a switch between the Rho and Rac signalling cascades (Madrid *et al.*, 2005). Similarly to other DRFs, FHOD1 contains a C-terminal auto-



regulation domain that consists of the highly conserved MD*x*LL motif and a polybasic cluster (Schönichen *et al.*, 2006). This domain was found to interact with the N-terminal region of FHOD1 (1–377) to form an inactive closed conformation of the DRF. The autoinhibition is relieved upon binding of the GTPase Rac1 to a GTPase-binding domain (GBD), the domain boundaries of which have not been unambiguously identified in FHOD1 (Westendorf, 2001; Schönichen *et al.*, 2006). The modular domain architecture of the 60–70 kDa N-terminal regulatory region is not defined by sequence homology and is believed to vary significantly among the DRF subfamilies.

Previous work on the Rho-interacting formin mDia1 has established that the biological activity of DRFs is regulated by the autoinhibitory interaction of the DAD with the DRF N-terminus. Crystal structures of the N-terminal region of mDia1 in complex with the DAD showed that the MD*x*LL signal sequence of the DAD is bound in a helical conformation to its recognition domain (Lammers *et al.*, 2005; Nezami *et al.*, 2006). This domain is composed of five armadillo repeats followed by an intertwined dimerization region. The preceding GBD subdomain contains only three helices and binds to activated RhoA in the nanomolar affinity range (Rose *et al.*, 2005; Otomo, Otomo *et al.*, 2005). Although binding of Rho and DAD at the N-terminal region of mDia1 is mutually exclusive, the binding sites only partially overlap. It has therefore been suggested that activation of the DRF is accomplished by a restructuring of the N-terminal domains (Faix & Grosse, 2006).

Here, we report the large-scale expression of the N-terminal region of recombinant human FHOD1 and the first diffraction data collected from native and seleno-L-methionine-labelled protein crystals. Elucidation of this domain structure is an initial step towards understanding the autoregulation and activation mechanisms of FHOD1. Moreover, comparisons to mDia1 will reveal similarities and differences between these two formin subfamilies that may also hold for other DRFs.

2. FHOD1$_N$ cloning, expression and purification

The coding sequences for the N-terminal region of FHOD1 (residues 1–339), referred to as FHOD1$_N$, were amplified by PCR from eukaryotic expression vectors encoding the full-length human *fhod1* gene (accession code AF113615; Schönichen *et al.*, 2006). The forward and reverse oligonucleotide primers (5′-CATGCCATGG-CGGGCGGGGAAGACCGC-3′ and 5′-CGGAATTCTCATCCAT-CCTCCAATTTCAGGGCGT-3′, respectively) were used to generate a 1020 bp DNA fragment with *Nco*I and *Eco*R1 restriction sites at the 5′ and 3′ ends. The fragment was cloned in the prokaryotic expression vector pProEx-HTa (Invitrogen) containing a hexahistidine (His$_6$) tag followed by a TEV protease cleavage site N-terminal to the multiple cloning site. Site-directed mutagenesis of the cysteines at positions 31, 43, 71, 164 and 211 to serines was performed using the megaprimer method with both sense and antisense oligonucleotides in a similar way to that described in Schulte *et al.* (2005). All constructs were single-pass sequenced prior to expression to confirm the integrity of the plasmid.

The FHOD1 constructs were expressed in *Escherichia coli* BL21(DE3)RIL cells (Novagen). Cells were grown at 310 K in Luria–Bertani medium containing 100 mg l^{-1} ampicillin; expression was induced with 0.5 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) upon reaching an optical density (OD_{600}) of 0.6 and was followed by another 4 h growth at 303 K. Unless stated otherwise, all chemicals were purchased from Sigma–Aldrich. Cells were harvested, washed with PBS, resuspended in lysis buffer *A* (20 m*M* Tris–HCl pH 7.6, 500 m*M* NaCl, 5 m*M* *β*-mercaptoethanol) with 20 m*M* imidazole and disrupted using a microfluidizer. After centrifugation at 30 000*g*, the filtered supernatant was loaded onto a 5 ml Ni–NTA column (Qiagen) equilibrated with lysis buffer *A*. Elution was performed with a linear imidazole gradient (0–500 m*M* imidazole) and the protein-containing fractions were pooled. The N-terminal His$_6$ tag was cleaved off with His$_6$-tagged TEV protease (1:20 molar ratio of protease:protein) during overnight dialysis against buffer *A* at 277 K. TEV protease and uncleaved protein were removed by passing the solution over an Ni–NTA column and the concentrated sample was subjected to size-exclusion chromatography on Superdex75 column (buffer: 20 m*M* Tris–HCl pH 7.6, 150 m*M* NaCl). Fractions were analyzed by SDS–PAGE and samples containing FHOD1$_N$ protein were pooled, concentrated to 20 mg ml^{-1}, flash-frozen in liquid nitrogen and stored at 193 K. The purity of the protein was about 95% (Fig. 1). Protein concentrations were determined by absorbance measurements at 280 nm based on a calculated extinction coefficient of 29 130 M^{-1} cm^{-1}. The elution profile of the size-exclusion chromatography of FHOD1$_N$ suggested a monomer–dimer distribution.

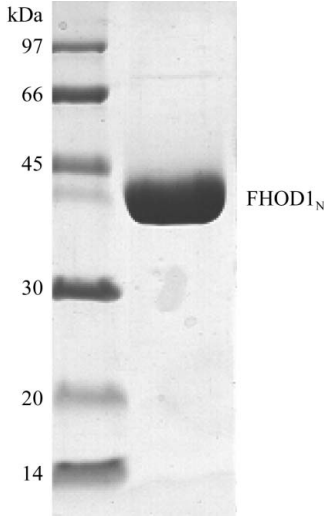


Figure 1
SDS–PAGE (15%, Coomassie blue staining) analysis of the purified N-terminal FHOD1 (1–339) protein after TEV protease cleavage of the His$_6$ tag.

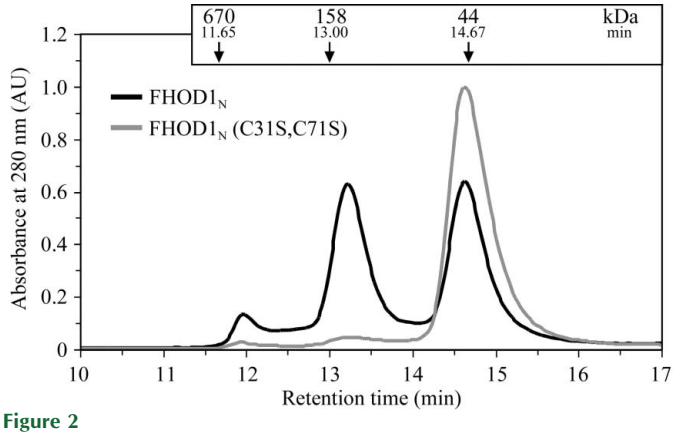
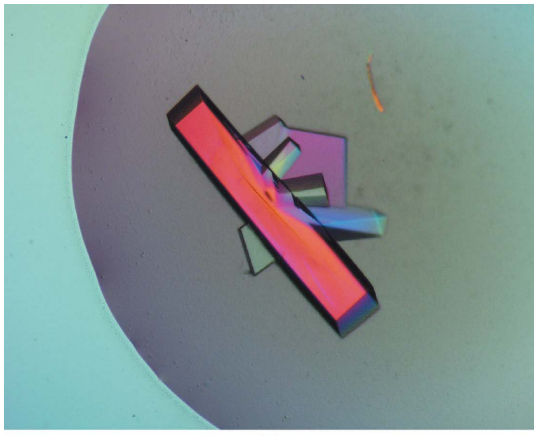


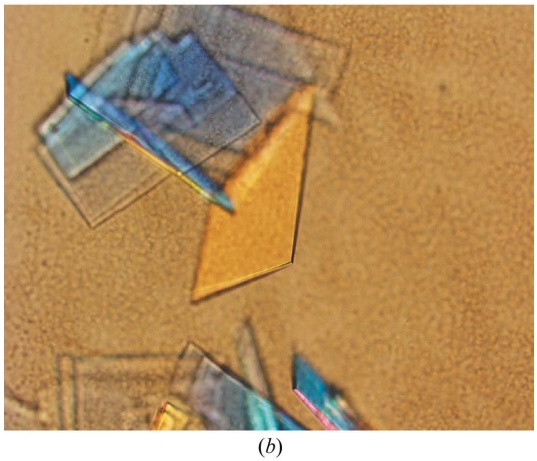
Figure 2
Analytical size-exclusion chromatography of FHOD1$_N$ proteins on a Biosep-SEC-S2000 column (300 × 1.8 mm). While the native protein (1–339) resized in both monomeric and dimeric forms in a buffer solution of 20 m*M* Tris–HCl pH 7.6, mutation of cysteines Cys31 and Cys71 to serines resulted in the monomeric form of the protein.

Upon addition of 10 m*M* dithioerythritol (DTE) to the sample prior to application onto the column, a single monomeric elution peak was obtained. This suggests that intermolecular disulfide-bond formation was occurring in the absence of DTE. A variety of cysteine-to-serine mutations were therefore generated. While individual C43S, C164S and C211S mutations resulted in unstable proteins, C31S and C71S mutant proteins were soluble and the double mutant FHOD1$_N$ (C31S,C71S) remained monomeric over a long time period at pH 8.0 (Fig. 2).

To facilitate experimental phasing using anomalous dispersion methods, seleno-L-methionine (SeMet) labelling of FHOD1$_N$ (C31S,C71S) was performed. Cells were grown in Luria–Bertani medium supplemented with 100 mg l^{-1} ampicillin, pelleted and resuspended in LeMaster media containing 100 mg l^{-1} ampicillin. This culture was grown at 303 K to mid-log phase (an OD_{600} of 0.8) and SeMet was added to 60 mg l^{-1} before induction with 0.5 m*M* IPTG. The purification protocol for the labelled protein was identical to that of the native protein except for the addition of 5 m*M* DTE to all buffers. The substitution of the methionine residues by SeMet was analysed by MALDI mass spectrometry. The molecular weight of 37 789.5 Da measured for (SeMet)-FHOD$_N$ (C31S,C71S) after TEV cleavage at the linker site corresponded to the incorporation of five Se atoms, indicating full labelling of the methionine residues at positions 1, 94, 188, 194 and 306 of the protein sample.



(*a*)



(*b*)

Figure 3
Photographs of human FHOD1$_N$ (1–339) crystals. (*a*) Crystals obtained using 4 *M* sodium chloride pH 9 that diffracted to 3.8 Å (approximate dimensions 0.7 × 0.2 × 0.2 mm). (*b*) Crystals of FHOD1$_N$ (C31S,C71S) grown from PEG/magnesium chloride conditions at pH 8. The anisotropic thin plates grew to dimensions of 0.4 × 0.3 × 0.05 mm.

Table 1
X-ray diffraction data and processing statistics of FHOD1$_N$ (C31S,C71S).

Data were collected at X10SA of the Swiss Light Source (Villigen, Switzerland).

Values in parentheses are for the highest resolution shell.

	SeMet-FHOD1$_N$		
	Inflection point	Peak	FHOD1$_N$
Space group	*P*1		*P*1
Unit-cell parameters (Å)			
a (Å)	35.53		35.44
b (Å)	73.38		73.93
c (Å)	79.26		78.67
α (°)	78.34		78.24
β (°)	86.11		86.17
γ (°)	89.91		89.67
No. of molecules in the ASU	2		2
Matthews coefficient (Å^3 Da^{-1})	2.66		2.66
Solvent content (%)	53.8		53.8
Wavelength (λ)	0.979746	0.979508	0.97883
Resolution (Å)	19.9–2.9	19.9–2.9	19.6–2.3
Total observations	71355	71527	109066
Unique reflections	33615	33714	32912
Average redundancy	2.12	2.12	3.31
Completeness (%)	96.8 (96.5)	97.1 (96.5)	94.8 (77.8)
R_{sym}†	10.0 (30.0)	8.7 (26.6)	13.9 (35)
$I/\sigma(I)$	11.8 (3.7)	13.1 (4.2)	8.4 (3.8)

† $R_{sym} = \sum |I(h)_j - \langle I(h)\rangle| / \sum I(h)_j$, where $I(h)_j$ is the scaled observed intensity of the *j*th symmetry-related observation of reflection *h* and $\langle I(h)\rangle$ is the mean value.

3. FHOD1$_N$ crystallization

Crystallization was initiated at 293 K using the sitting-drop method and a Mosquito robot (TTP Labtech, UK). A concentration of 20 mg ml^{-1} FHOD1$_N$ was used for initial screening. 0.1 µl volumes of protein solution were mixed with 0.1 µl reservoir solution in 96-well Greiner crystallization plates containing 100 µl reservoir solution. Small crystals were obtained with Grid Screen Sodium Chloride from Hampton Research under conditions consisting of 4 *M* NaCl and 0.1 *M* Bicine pH 9.0. After optimization, large single crystals (0.7 × 0.2 × 0.2 mm; Fig. 3*a*) grew in 2 µl hanging drops within 1 d in 4 *M* sodium chloride, 0.2 *M* sodium acetate and 0.1 *M* Bicine pH 9.0. FHOD1$_N$ crystals were flash-cooled in liquid nitrogen without addition of cryoprotectant. Despite many optimization efforts, these crystals, which belonged to space group *P*222, did not diffract to better than to 3.8 Å. Thus, a second screening was performed with the mutant FHOD1$_N$ (C31S,C71S). A new crystal form appeared in condition Nos. 42 and 53 of the MbClass I Nextal (Qiagen) screen. Small thin plates grew in 1 d. In optimized conditions consisting of 20%(*w*/*v*) PEG 4000, 10%(*v*/*v*) glycerol, 0.3 *M* magnesium chloride and 0.1 *M* Tris–HCl pH 8.0, crystal plates appeared after 1 d in 2 µl drops using the hanging-drop technique. The crystal quality was further improved by microseeding techniques, which yielded crystal dimensions of 0.4 × 0.3 × 0.05 mm (Fig. 3*b*). The best crystal of FHOD1$_N$ (C31S,C71S) diffracted to 2.3 Å and an SeMet-derivatized crystal diffracted to 2.9 Å at a synchrotron-radiation source.

4. Data collection and processing

Diffraction data were collected to 2.3 Å resolution from the FHOD1$_N$ (C31S,C71S) crystal at 100 K on beamline X10SA (PXII) of the Swiss Light Source (SLS, Villigen, Switzerland) equipped with a MAR 225 CCD detector (oscillation width per frame 1°; 360 frames collected). The *XDS* package (Kabsch, 1993) was used to process, integrate and scale the collected data. The crystals belong to space group *P*1, with unit-cell parameters *a* = 35.4, *b* = 73.9, *c* = 78.7 Å. Assuming the presence of two molecules in the asymmetric unit, the solvent content

of the crystals is 53.8%, corresponding to a Matthews coefficient of 2.66 Å^3 Da^{-1}. Data-collection statistics are given in Table 1. As molecular replacement with structures that were thought to be related was not successful, experimental phasing with anomalous data has been initiated. The SeMet-labelled crystals diffracted more weakly than the unlabelled mutant crystals. MAD data for wavelengths corresponding to both the peak and inflection point were collected to 2.9 Å for SeMet-labelled crystals at the same beamline. The oscillation width per frame was 1° and 400 images were collected per data set. We are currently refining the selenium sites for subsequent structure determination and refinement of FHOD1$_N$ (1–339).

We thank Michael Weyand and Nils Schrader for data collection and Roger S. Goody for continuous support. Georg Holtermann is acknowledged for excellent technical assistance. Access to beamline X10SA (PXII) at the Swiss Light Source, Villigen, Switzerland and technical support is gratefully acknowledged. This work was supported by a grant from the Deutsche Forschungsgemeinschaft to MG (GE-976/4). OP and AR are supported by funds from the European Young Investigator (EURYI) award, DFG grant RA 1364/1-1 and the EC Sixth Framework Programme.

References

Alberts, A. S. (2001). *J. Biol. Chem.* **276**, 2824–2830.
Faix, J. & Grosse, R. (2006). *Dev. Cell*, **10**, 693–706.
Gasteier, J. E., Madrid, R., Krautkramer, E., Schroder, S., Muranyi, W., Benichou, S. & Fackler, O. T. (2003). *J. Biol. Chem.* **278**, 38902–38912.
Goode, B. L. & Eck, M. J. (2007). *Annu. Rev. Biochem.* **76**, 593–627.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Lammers, M., Rose, R., Scrima, A. & Wittinghofer, A. (2005). *EMBO J.* **24**, 4176–4187.
Madrid, R., Gasteier, J. E., Bouchet, J., Schröder, S., Geyer, M., Benichou, S. & Fackler, O. T. (2005). *FEBS Lett.* **579**, 441–448.
Nezami, A. G., Poy, F. & Eck, M. J. (2006). *Structure*, **14**, 257–263.
Otomo, T., Otomo, C., Tomchick, D. R., Machius, M. & Rosen, M. K. (2005). *Mol. Cell*, **18**, 273–281.
Otomo, T., Tomchick, D. R., Otomo, C., Panchal, S. C., Machius, M. & Rosen, M. K. (2005). *Nature (London)*, **433**, 488–494.
Romero, S., Le Clainche, C., Didry, D., Egile, C., Pantaloni, D. & Carlier, M. F. (2004). *Cell*, **119**, 419–429.
Rose, R., Weyand, M., Lammers, M., Ishizaki, T., Ahmadian, M. R. & Wittinghofer, A. (2005). *Nature (London)*, **435**, 513–518.
Schönichen, A., Alexander, M., Gasteier, J. E., Cuesta, F. E., Fackler, O. T. & Geyer, M. (2006). *J. Biol. Chem.* **281**, 5084–5093.
Schulte, A., Czudnochowski, N., Barboric, M., Schonichen, A., Blazek, D., Peterlin, B. M. & Geyer, M. (2005). *J. Biol. Chem.* **280**, 24968–24977.
Vavylonis, D., Kovar, D. R., O'Shaughnessy, B. & Pollard, T. D. (2006). *Mol. Cell*, **21**, 455–466.
Watanabe, N., Kato, T., Fujita, A., Ishizaki, T. & Narumiya, S. (1999). *Nature Cell Biol.* **1**, 136–143.
Westendorf, J. J. (2001). *J. Biol. Chem.* **276**, 46453–46459.