DNA-binding domain of myelin-gene regulatory factor: purification, crystallization and X-ray analysis

WenYu Wu,[‡] Xiangkai Zhen[‡] and Ning Shi*

State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, 155 Yangqiao Road West, Fuzhou, Fujian 350002, People's Republic of China. *Correspondence e-mail: shining@fjirsm.ac.cn

Received 19 May 2017

Accepted 26 May 2017

Edited by R. L. Stanfield, The Scripps Research Institute, USA

[‡] These authors should be considered co-first authors.

Keywords: myelin-gene regulation; MRF; DNA-binding domain; trimeric transcription factor; X-ray crystallography.

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

The myelin sheath, which envelops axons in the vertebrate central nervous system, is crucial for the rapid conduction of action potentials. Myelin-gene regulatory factor (MRF) is a recently identified transcription factor that is required for myelin-sheath formation. Loss of MRF leads to demyelinating diseases and motor learning deficiency. MRF is a membrane-bound transcription factor that undergoes autocleavage from the endoplasmic reticulum membrane. The N-terminus of MRF contains a DNA-binding domain (DBD) that functions as a homotrimer. In this study, the MRF DBD was cloned, purified and crystallized in order to understand the molecular mechanism that regulates the transcription of myelin genes. Selenomethionine was subsequently introduced into the crystals to obtain the phases for the MRF DBD structure. The native and selenomethionine-labelled crystals exhibited diffraction to 2.50 and 2.51 Å resolution, respectively. The crystals belonged to space group *P*321 and the selenomethionine-labelled crystals had unit-cell parameters $a = 104.0$, $b = 104.0$, $c = 46.7$ Å, $\alpha = 90$, $\beta = 90$, $\gamma = 120^\circ$. The calculated Matthews coefficient was $3.04 \text{ Å}^3 \text{ Da}^{-1}$ and the solvent content was 59.5%, indicating the presence of one MRF DBD molecule in the asymmetric unit.

1. Introduction

Myelin, which envelops axons in the vertebrate central nervous system (CNS), is crucial for the rapid and efficient conduction of action potentials in axons. Myelin is a specialized structure that is generated by terminally differentiated oligodendrocytes (OLs) in the CNS (Küspert *et al.*, 2011; Li *et al.*, 2007). The production of myelin and the generation of mature OLs involve a fine-tuned process that is regulated by numerous extracellular and intercellular factors (Taveggia *et al.*, 2010). The loss of myelin causes many severe diseases, such as multiple sclerosis and leukodystrophies. Researchers have recently found that newly generated myelin is required for motor learning in adults (Xiao *et al.*, 2016; McKenzie *et al.*, 2014). The myelin-gene regulatory factor (MRF) is required for myelin-gene expression and myelin maintenance (Emery *et al.*, 2009; Koenning *et al.*, 2012). The expression of many myelin genes, including MBP, MAG and MOG, is absent when MRF is knocked down. MRF is specifically expressed in post-mitotic OLs and is a membrane-bound transcription factor that is located in the endoplasmic reticulum (ER) membrane (Li *et al.*, 2013; Bujalka *et al.*, 2013). After expression, it is autocleaved from the ER membrane, and its N-terminal fragment, which includes a DNA-binding domain (DBD), enters the nucleus to regulate myelin-gene expression. Previous studies showed that the MRF DBD is functional as a

homotrimer, which is a rare configuration for a transcription factor (Muth *et al.*, 2016; Kim *et al.*, 2017; Li *et al.*, 2013). To date, structural information on MRF remains lacking, which hinders the complete understanding of this protein and its specific functions. In the present study, we investigated the expression, purification, crystallization and X-ray diffraction of the MRF DBD.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Production and purification of native MRF DBD. The mouse MRF DNA fragment containing the DBD and ICA domains (residues 351–717) was amplified from a full-length gene clone (NP_001028653.1) by standard PCR. The amplified products were then cloned into a modified pRSFDuet-1 expression vector (Novagen) with an N-terminal 6×His tag. The sequence of the final recombined vector RSF_MRF-DBD was validated *via* DNA sequencing.

The recombinant MRF fragment was obtained by over-expressing RSF_MRF-DBD in *Escherichia coli* BL21 (DE3) cells (Novagen) that had been induced overnight with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 18°C. The bacterial cells were harvested by centrifugation at 7000 rev min⁻¹ for 10 min. 5 g of the cell pellet was then

resuspended in 50 ml lysis buffer consisting of 50 mM Tris-HCl pH 8.0, 150 mM NaCl. The cells were lysed on ice by ultrasonication and were then centrifuged at 12 000 rev min⁻¹ for 30 min. The clarified supernatant was incubated with nickel Sepharose affinity resin (GE Healthcare) and then washed with lysis buffer. The protein eluted with lysis buffer containing 300 mM imidazole was analyzed by limited proteolytic mapping as follows. Firstly, trypsin (Worthington) was dissolved in a buffer solution consisting of 50 mM Tris-HCl pH 8.0, 20 mM CaCl₂ to prepare a 1 mg ml⁻¹ trypsin stock solution. The aliquots were then frozen at -20°C. Trypsin was added to the purified MRF protein in a buffer solution consisting of 50 mM Tris-HCl pH 8.0, 150 mM NaCl. For trypsin limited proteolytic mapping analysis, a stable band was observed with a trypsin:protein ratio of 1:500 000 (*w:w*) at 4°C for 30 min. The eluted protein was subsequently digested with trypsin at 4°C for 30 min. The digestion reaction was terminated by adding a tenfold excess of soybean trypsin inhibitor (Sigma) to the reaction mixture. The stable fragments were further purified by size-exclusion chromatography on a Superdex 75 10/300 GL column (GE Healthcare) that had been pre-equilibrated in a buffer solution consisting of 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM DTT. The peak fraction was concentrated to 20 mg ml⁻¹ for crystallization (Figs. 1*b* and 1*c*). Macromolecule-production information is summarized in Table 1.

Figure 1

Schematic of MRF, trypsin limited proteolytic mapping and purification of the MRF DBD using size-exclusion chromatography. (a) Schematic of the MRF protein showing the fragment expressed and the location of the proteolytic sites (both the autocleavage and trypsin cut sites). (b) Trypsin limited proteolytic mapping was performed and checked by 15% SDS-PAGE. Lane 1 contains the 300 mM imidazole eluate from a nickel-affinity column containing the purified autocleaved product of the MRF DBD. Lane 2 contains trypsinized MRF DBD at a trypsin:protein ratio of 1:500 000 (*w:w*) after gel filtration. Lanes 4–10 contain trypsinized MRF DBD at trypsin:protein ratios of 1:1 000 000, 1:800 000, 1:600 000, 1:500 000, 1:400 000, 1:300 000 and 1:200 000 (*w:w*), respectively. (c) Elution profiles from gel-filtration chromatography of the MRF DBD using Superdex 75 10/300 GL column (GE Healthcare). A calibration curve of proteins with known molecular weights (bovine serum albumin, 66 kDa; chicken ovalbumin, 44 kDa; carbonic anhydrase, 29 kDa) is shown. The molecular weights of these proteins were plotted against their calculated V_e/V_o and fitted by exponential regression analysis. On the basis of this calculation, the V_e/V_o of the trypsinized MRF DBD results in a molecular weight of about 24 kDa.

2.1.2. Production and purification of selenomethionine-labelled MRF DBD. The RSF_MRF-DBD plasmid (see above) was transformed into *E. coli* BL21 (DE3) cells. Selenomethionine-labelled MRF DBD was obtained in *E. coli* BL21 (DE3) cells *via* methionine-pathway inhibition at 37°C in M9 medium containing 0.4% glucose, 2 mM MgSO₄, 0.1 mM CaCl₂ and 50 µg l⁻¹ kanamycin (Qoronfleh *et al.*, 1995; Hendrickson *et al.*, 1990). The cells were cultured to an OD₆₀₀ of 0.8. Prior to the addition of IPTG, selenomethionine was supplemented along with leucine, isoleucine and valine at final concentrations of 50 µg l⁻¹. Lysine, threonine and phenylalanine were supplemented to final concentrations of 100 µg l⁻¹. The cells were harvested *via* centrifugation. Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM PMSF, 1 mM β-mercaptoethanol). The cells were lysed on ice *via* ultrasonication and were then centrifuged at 12 000 rev min⁻¹ for 30 min at 4°C. The clarified supernatant was loaded onto nickel Sepharose affinity resin (GE Healthcare) that had been equilibrated in lysis buffer. The column was extensively washed and the target protein was eluted with lysis buffer containing 300 mM imidazole. Fractions containing pure protein were pooled and treated with trypsin at a trypsin:protein ratio of

Table 1

Macromolecule-production information.

Source organism	Mouse
DNA source	cDNA
Forward primer	CTGCAGGACAGTGACAGCCTCA
Reverse primer	CAAACTGGCCTGGTTAGAGGC
Cloning vector	pRSFDuet-1 vector
Expression vector	pRSFDuet-1 vector
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced†	MGSSHHHHHSQDPLEVLFGQPEIIKWQPHQQNK- WATLYDANYKELPMLTYRVDADKGFNFSGDD- AFVCQKKNHQVTVYIGMLGEPKYVKTPEGLK- PLDCFYLLKLGKLEALNQSINIEQSQSDRSK- RPFNPVTVNLPEQVTKVTVGRHLHFSETTANN- MRKKGKPNPDQRYFMLVVALQAHAQNQNYTLA- AQISERIIVRASNPQGFEESDVLWQRAQLPD- TVFHHGRVGINTRPDEALVHGNVVMGSLM- HPSDLRAKEHVQEVDTEQLKRISRMRLVHYR- YKPEFAASAGIEATAPETGVIAQEVKEILPEA- VKDTGDVVFANGKTIENFLVVKERIFMENVG- AVKELCKLTDNLETRIDELERWSHKLAKLRRL- DSLKS

† The underlined residues are autocleaved after the protein has been expressed.

1:500 000(*w:w*) at 4°C for 30 min. After the reaction had been terminated with trypsin inhibitor, the reaction mixture was concentrated and then further purified by size-exclusion

Figure 2

Photographs of MRF DBD crystals, a silver-stained gel of the crystals and an X-ray diffraction image. (a) Optimized native crystals of the MRF DBD obtained in 20% PEG 4000, 50 mM magnesium acetate, 0.1 M HEPES pH 7.5. The approximate dimensions of the crystals are 0.15 × 0.05 × 0.04 mm. (b) Selenomethionine-labelled MRF DBD crystals obtained in 20% PEG 4000, 50 mM magnesium acetate, 0.1 M HEPES pH 7.5. The approximate dimensions of the crystals are 0.05 × 0.02 × 0.02 mm. (c) SDS-PAGE gel of the crystals with silver staining. Lane 1 is a control containing the purified MRF DBD protein. Lane 2 contains crystals washed twice with buffer. Lanes 3 and 4 contain the crystal wash buffers. (d) X-ray diffraction image of a selenomethionine-labelled MRF DBD crystal.

Table 2
Data collection and processing.

	Selenomethionine-labelled crystal	Native crystal
Diffraction source	BL17U1, SSRF	BL17U1, SSRF
Wavelength (Å)	0.9792	0.9792
Temperature (K)	100	100
Detector	ADSC Q315r	ADSC Q315r
Crystal-to-detector distance (mm)	250	300
Rotation range per image (°)	1	1
Total rotation range (°)	180	180
Exposure time per image (s)	1	1
Space group	<i>P</i> 321	<i>P</i> 321
<i>a</i> , <i>b</i> , <i>c</i> (Å)	104.0, 104.0, 46.7	106.8, 106.8, 48.1
α , β , γ (°)	90, 90, 120	90, 90, 120
Mosaicity (°)	0.34	0.25
Resolution range (Å)	34.75–2.51 (2.55–2.51)	48.11–2.50 (2.54–2.50)
Total No. of reflections	107970 (4391)	119004 (5444)
No. of unique reflections	10187 (481)	11154 (552)
Completeness (%)	99.8 (97.4)	99.7 (97.4)
Multiplicity	10.6 (9.1)	10.7 (9.9)
$\langle I/\sigma(I) \rangle$	18.1 (2.0)	16.4 (2.3)
<i>R</i> _{meas}	0.100 (0.983)	0.093 (0.956)
Overall <i>B</i> factor from Wilson plot (Å ²)	49.7	49.7

chromatography using a Superdex 75 10/300 GL column (GE Healthcare) that had been equilibrated in 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 1 mM DTT. The fractions containing pure protein were concentrated to 18 mg ml^{−1} for crystallization screening.

2.2. Crystallization

Initial crystallization screening was conducted *via* sitting-drop vapour diffusion. In brief, 0.5 µl purified protein solution was mixed with 0.5 µl reservoir solution at 18°C. Although crystals appeared in PEG 4000, 1 M ammonium formate after 3 d, these crystals did not exhibit diffraction despite extensive optimization. After two weeks, another type of rod-like crystal appeared in buffer consisting of 0.1 M HEPES pH 7.5, 20% PEG 4000, 50 mM magnesium acetate. Selenomethionine-labelled crystals were obtained under similar conditions as for the native crystal. Protein identification was performed on a MALDI TOF-TOF 5800 Analyzer (AB SCIEX, Foster City, California, USA).

2.3. Data collection and processing

The crystals were transferred for 30 s to mother liquor containing 20% glycerol as a cryoprotectant. The crystals were then cooled in liquid nitrogen. Data were collected at the BL17U1 station of the Shanghai Synchrotron Radiation Facility (SSRF; Wang *et al.*, 2015) and were processed using *HKL*-2000 and *xia2* (Otwinowski & Minor, 1997; Winter, 2010; Kabsch, 2010; Winn *et al.*, 2011; Evans, 2006). One data set was collected from a selenomethionine-substituted crystal at the Se peak wavelength (0.9792 Å). The Se sites were identified using *SHELX* (Sheldrick, 2008; Schneider & Sheldrick, 2002). The structure was solved with *AutoSol* and an initial model was built with *AutoBuild*. Both programs are part of the

PHENIX suite of crystallographic programs (Adams *et al.*, 2010).

3. Results and discussion

The MRF fragment (351–717) was overexpressed in *E. coli* and then underwent self-cleavage. The fragment produced a protein with a molecular weight of 30 kDa, compared with the calculated molecular weight of 45 kDa (Fig. 1*a*). The protein was purified using a nickel-affinity column. Initial crystallization trials using this fragment failed to produce crystals. Previous studies showed that this fragment was the product of autoproteolysis by the ICA domain and includes the DBD and the bridge between the DBD and the ICA domain. However, the bridge may not be well structured and possibly hindered crystallization. Limited proteolytic analysis with trypsin confirmed this speculation and generated a stable band (about 24 kDa) for the MRF DBD (Fig. 1*b*). The size-exclusion chromatograph of the protein showed a single peak that corresponded to the monomer (Fig. 1*c*). Crystals appeared in a condition consisting of PEG 4000 and 1 M ammonium formate after 3 d. Unfortunately, few diffraction spots appeared in the low-resolution range for the initial crystals, and the diffraction of crystals obtained under this condition could not be improved by optimizing the pH, the protein concentration or the precipitant concentration. Finally, diffraction-quality crystals and those of the selenomethionine-labelled variant were obtained using a reservoir solution consisting of 0.1 M HEPES pH 7.5, 20% PEG 4000, 50 mM magnesium acetate (Figs. 2*a* and 2*b*). The SDS–PAGE gel showed that the crystals were protein crystals and that they had the same molecular weight as the purified MRF DBD protein (Fig. 2*c*). Western blotting with an anti-His monoclonal antibody showed that the N-terminal His tag is removed by trypsin (Supplementary Fig. S1). The crystals were confirmed to contain MRF protein by a *MASCOT* search (Supplementary Fig. S2) and had a mass from MALDI–TOF of 24 002 Da, which we believe to correspond to the fragment containing residues 351–559. X-ray diffraction data were collected on beamline BL17U1 of SSRF and were processed at 2.51 Å resolution (Fig. 2*d*). The crystals belonged to space group *P*321, with unit-cell parameters *a* = 104.0, *b* = 104.0, *c* = 46.7 Å, α = β = 90°, γ = 120°. Data-collection and processing statistics are summarized in Table 2. The rational Matthews coefficient and solvent content of the crystal are 3.04 Å³ Da^{−1} and 59.5%, respectively, which suggest the presence of one molecule per asymmetric unit (Matthews, 1968).

The MRF DBD is conserved in evolution, and its closest homologue is the yeast sporulation-specific transcription factor Ndt80. A molecular-replacement solution could not be found using Ndt80 (PDB entry 1m6u; Montano *et al.*, 2002) as a model. Structure analysis showed that Ndt80 is an Ig-fold transcription factor with an extra helix and a larger loop compared with other similar transcription factors (Lamoureux *et al.*, 2002). However, the size of MRF is closer to those of the general Ig-fold transcription factors. This difference may be the reason why a molecular-replacement solution could not be

Figure 3

Electron density of Se atoms as anomalous signal in an MRF DBD trimer. The red density represents an anomalous map contoured at 5.0σ . The green lines indicate the unit cell.

found. To determine the phases of the MRF DBD structure, single-wavelength anomalous diffraction data were collected from a selenomethionine-labelled crystal. Four Se sites were found in the asymmetric unit (of the expected four sites per molecule) using *SHELX*. The structure was solved with *AutoSol* and an initial model was built with *AutoBuild*. Analysis of the crystal packing revealed that the MRF DBD forms a trimer with a crystallographic threefold axis (Fig. 3). The detailed structural analysis and biochemical studies of the MRF DBD will be published elsewhere. The high-resolution structure of the MRF DBD may provide insights into the mechanism of trimeric transcription factors.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Bujalka, H., Koenning, M., Jackson, S., Perreau, V. M., Pope, B., Hay, C. M., Mitew, S., Hill, A. F., Lu, Q. R., Wegner, M., Srinivasan, R., Svaren, J., Willingham, M., Barres, B. A. & Emery, B. (2013). *PLoS Biol.* **11**, e1001625.
- Emery, B., Agalliu, D., Cahoy, J. D., Watkins, T. A., Dugas, J. C., Mulinyawe, S. B., Ibrahim, A., Ligon, K. L., Rowitch, D. H. & Barres, B. A. (2009). *Cell*, **138**, 172–185.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kim, D., Choi, J.-O., Fan, C., Shearer, R. S., Sharif, M., Busch, P. & Park, Y. (2017). *Nucleic Acids Res.* **45**, 5112–5125.
- Koenning, M., Jackson, S., Hay, C. M., Faux, C., Kilpatrick, T. J., Willingham, M. & Emery, B. (2012). *J. Neurosci.* **32**, 12528–12542.
- Küspert, M., Hammer, A., Bösl, M. R. & Wegner, M. (2011). *Nucleic Acids Res.* **39**, 1280–1293.
- Lamoureux, J. S., Stuart, D., Tsang, R., Wu, C. & Glover, J. N. M. (2002). *EMBO J.* **21**, 5721–5732.
- Li, H., Lu, Y., Smith, H. K. & Richardson, W. D. (2007). *J. Neurosci.* **27**, 14375–14382.
- Li, Z., Park, Y. & Marcotte, E. M. (2013). *PLoS Biol.* **11**, e1001624.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McKenzie, I. A., Ohayon, D., Li, H., Paes de Faria, J., Emery, B., Tohyama, K. & Richardson, W. D. (2014). *Science*, **346**, 318–322.
- Montano, S. P., Cote, M. L., Fingerhman, I., Pierce, M., Vershon, A. K. & Georgiadis, M. M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 14041–14046.
- Muth, K. N., Piefke, S., Weider, M., Sock, E., Hermans-Borgmeyer, I., Wegner, M. & Küspert, M. (2016). *Glia*, **64**, 2120–2132.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Qoronfleh, M. W., Ho, T. F., Brake, P. G., Banks, T. M., Pulvino, T. A., Wahl, R. C., Eshraghi, J., Chowdhury, S. K., Ciccirelli, R. B. & Jones, B. N. (1995). *J. Biotechnol.* **39**, 119–128.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772–1779.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Taveggia, C., Feltri, M. L. & Wrabetz, L. (2010). *Nature Rev. Neurol.* **6**, 276–287.
- Wang, Q. S. *et al.* (2015). *Nucl. Sci. Tech.* **26**, 010102.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
- Xiao, L., Ohayon, D., McKenzie, I. A., Sinclair-Wilson, A., Wright, J. L., Fudge, A. D., Emery, B., Li, H. & Richardson, W. D. (2016). *Nature Neurosci.* **19**, 1210–1217.