Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the coiled-coil domain of dystrophia myotonica kinase

Pilar Garcia, Marco Marino and Olga Mayans*

Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Correspondence e-mail: olga.mayans@unibas.ch

Received 6 August 2004
Accepted 22 October 2004

The coiled-coil domain of dystrophia myotonica protein kinase (DMPK) has been cloned, overexpressed, purified and crystallized. Two crystal forms have been obtained that belong to space groups $P3$ and $P2_12_12_1$ and diffract to 2.4 and 1.6 Å resolution, respectively. Experimental phases were obtained by MAD from an SeMet derivative. The location of selenium sites used molecular-replacement phases obtained from search models lacking sequence similarity with the coiled-coil under study. Both crystal forms contain three polypeptide chains in the asymmetric unit.

1. Introduction

Regulation by oligomerization is a key feature of the dystrophia myotonica protein kinase (DMPK) related family of Ser/Thr kinases. This group of complex multi-domain kinases includes important members such as human $p160^{ROCK}$ (ROCKI), Rho-kinase (ROCKII), citron Rho-interacting kinase and Cdc42-binding kinase (MRCK) among others (Riento & Ridley, 2003). They are functionally related to reorganization events in the cytoskeleton during cellular processes such as cytokinesis, neurite outgrowth and smooth-muscle contraction. These kinases are characterized by the high conservation of their catalytic domains and by the unusual presence of one or more coiled-coil domains (CC) that are central to regulation of their activity. Based on sequence data, the CC domains from members of this family are predicted to form dimeric arrangements. However, it is remarkable that little or no overall sequence homology can be detected across these domains and that different biochemical roles have been attributed to them. At their C-termini, the CC moieties of ROCKI, ROCKII and citron kinase (Di Cunto *et al.*, 1998; Riento & Ridley, 2003) house binding motifs for small Rho GTPases that are known to regulate cytoskeletal dynamics. The CC domains of MRCK and DMPK are not known to mediate protein interactions, but have instead been proposed to act as an inhibitor and an enhancer of kinase activity, respectively (Tan *et al.*, 2001; Zhang & Epstein, 2003). Furthermore, it is generally accepted that these CC sections mediate dimerization in these kinases (Zhang & Epstein, 2003). However, it has been shown that other domains within these complex systems can also lead to oligomerization (Chen *et al.*, 2002; Tan *et al.*, 2001). In summary, the functionality of CC domains from this kinase family remains unclear in the absence of structural data.

To date, the structures of only two closely related CC fragments from this kinase family have been reported. These correspond to the C-termini of ROCKII-CC and ROCKI-CC, both of which are dimeric and span the Rho-binding motif. Their structures have been elucidated in the free form (PDB code 1uix; Shimizu *et al.*, 2003) and in complex with RhoA (PDB code 1s1c; Dvorsky *et al.*, 2004), respectively. In order to gain further understanding of the contribution of CC moieties to kinase regulation in this family, we have undertaken the study of the CC domain from DMPK (DMPK-CC). DMPK is speculated to be related to a progressive neuromuscular disorder known as myotonic dystrophy (DM), which is the most common form of muscular dystrophy in adults (Larkin & Fardaei, 2001). Its involvement in disease, however, remains polemic. Understanding the activation mechanism of DMPK will contribute to further clarifying its functionality.

2. Construction of the plasmid

Based on structure-prediction results from *MULTI-COIL* (Wolf *et al.*, 1997), DMPK-CC was identified to span residues 461–537. An expression construct comprising these residues was derived from human DMPK cDNA (GenBank L08835) using primers 5′-CA-TG**CCATGG**CAGAGGCTGAGGCCGAG-3′ and 5′-GG**GGTACC**TTACCCCGTGACAGC TGTGG-3′ carrying *Nco*I and *Kpn*I restriction sites (in bold). The PCR product was purified (Wizard, Promega) and cloned into the similarly digested expression vector pET-M11 (EMBL). This modified version of pET-24d (Novagen) includes a His_6 tag plus a TEV (tobacco etch virus) protease-cleavage site

N-terminal to the expression construct. The plasmid was constructed in *Escherichia coli* strain DH5α.

3. Protein overexpression and purification

DMPK-CC was expressed in *E. coli* BL21(DE3) grown in LB medium supplemented with 25 µg ml^{-1} kanamycin and 34 µg ml^{-1} chloramphenicol at 310 K. Induction took place at an $OD_{600\,nm}$ of 0.6 by the addition of 1 m*M* isopropyl-β-D-thiogalactopyranoside (IPTG) and expression took place overnight at 298 K. Cells were harvested by centrifugation at 5000*g* for 30 min at 277 K. The bacterial pellet was resuspended in lysis buffer (50 m*M* Tris pH 7.5, 50 m*M* NaCl) and sonicated in the presence of 1 mg ml^{-1} lysozyme. The homogenate was clarified by centrifugation at 20 000*g* for 1 h and 277 K. For affinity purification, the supernatant was applied onto a HiTrap Ni^{2+}-Chelating HP column (Amersham Biosciences) equilibrated in lysis buffer containing 40 m*M* imidazole. Protein was eluted at 200 m*M* imidazole. Tag removal took place by incubation with TEV protease (0.4 mg ml^{-1}) for 2 h at room temperature during dialysis against 50 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl. Subtractive purification of protease (His_6-tagged at the C-terminus) and non-digested protein was carried out on a subsequent chelating step. The cleaved protein was then concentrated to 5–10 mg ml^{-1} according to a Bradford assay (Bradford, 1976) and used in crystallization trials. Higher purity DMPK-CC could be obtained by further size-exclusion chromatography using a Hi-Load 16/60 Superdex 75 prep-grade column (Amersham Biosciences) equilibrated in 50 m*M* Tris pH 7.5 and 200 m*M* NaCl. However, this additional procedure did not result in detectable variation of crystallization results and was discontinued.

Table 1
X-ray data and phasing statistics.

Data processing and reduction used the *MOSFLM*/*SCALA* software (Kabsch, 1988; Leslie, 1992). Phase calculation and refinement was performed in *SHARP* (de La Fortelle and Bricogne, 1997) using data to 3.0 Å resolution. Phasing statistics are given for acentric reflections.

	*P*3	$P2_12_12_1$			
	Native	Native	SeMet peak	SeMet remote	SeMet edge
Unit-cell parameters (Å)	$a = b = 38.05$, $c = 114.37$	$a = 39.09$, $b = 46.17$, $c = 143.48$			
No. chains in AU	3	3			
V_M (Å^3 Da^{-1})	1.3	2.4			
Solvent content (%)	32	48			
Wavelength (Å)	0.886	0.98	0.9793	0.9392	0.9795
Resolution (Å)	20–2.4 (2.53–2.40)	20–1.6 (1.69–1.6)	20–2.3 (2.42–2.3)	20–2.5 (2.64–2.5)	20–2.9 (3.06–2.9)
Unique reflections	5395 (728)	33876 (3951)	11724 (1440)	9565 (1379)	6105 (873)
Completeness (%)	74.1 (68.2)	96.3 (79.2)	97.2 (83.9)	99.5 (99.5)	99.2 (100.0)
$R_{sym}(I)$	0.123 (0.228)	0.059 (0.330)	0.079 (0.204)	0.091 (0.316)	0.104 (0.304)
Redundancy	2.0 (1.3)	4.4 (2.3)	12.7 (8.7)	6.8 (7.0)	6.7 (7.0)
$\langle I/\sigma(I)\rangle$	6.5 (1.4)	15.0 (2.5)	25.6 (8.2)	17.8 (6.2)	16.5 (6.1)
Phasing power (iso/ano)			0.15/1.83	−/0.96	0.58/0.55
R_{Cullis}			1.0/0.64	−/0.84	0.91/0.93
FOM (before/after solvent treatment)				0.45/0.87	

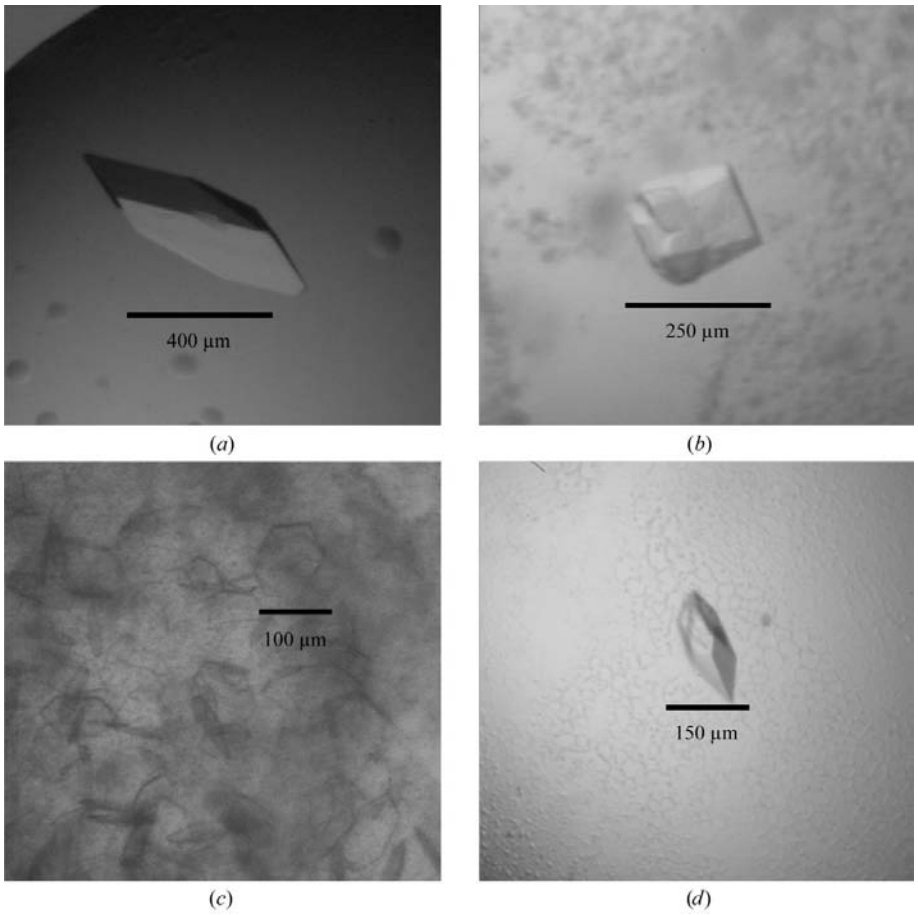

Figure 1
Crystal habits. (*a*) Crystals grown under condition (i) supplemented with 3% 2-propanol. (*b*) Crystals grown under condition (i) supplemented with 5% dioxane. (*c*) Thin crystal plates grown from heavy precipitate under condition (ii). (*d*) Best diffracting crystals obtained from condition (iii).

4. Crystallization

Crystallization was carried out by the hanging-drop vapour-diffusion method at 291 K. Each drop consisted of equal volumes (1.5 µl) of protein solution and reservoir solution. The components of the protein solution are described above and the protein concentration was 10 mg ml^{-1}. Initial screening used a wide range of commercially available crystallization kits. Crystals were obtained from Structure Screen 1 (Molecular Dimensions) under two sets of conditions: (i) 2.0 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.5 and (ii) 100 m*M* sodium acetate pH 4.6, 8% PEG 4000. Crystals obtained under condition (i) grew as stacks of multiple plates. Single crystals grew from optimized media containing 2.5 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 7.5 and either 3% 2-propanol or 5% dioxane (Figs. 1*a* and 1*b*). Despite the good visual appearance of these crystals, their poor diffraction quality made

them unsuitable for structure elucidation. As a consequence, this crystal form was abandoned. Crystals grown under condition (ii) appeared one month after drop mixing. They grew from heavy precipitate as extremely thin hexagonal plates of small dimensions (Fig. 1*c*). Efforts to increase the crystal size, including seeding protocols, yielded no improvement.

Finally, a customized screen resulted in an additional crystal type grown from 100 m*M* Tris–HCl pH 7.5, 2.0 *M* sodium/potassium tartrate (condition iii). Crystals exhibited a polyhedral habit and grew to final dimensions within two to three weeks of drop mixing (Fig. 1*d*). The diffraction properties of these showed them to be suitable for structure elucidation.

5. Data collection and NCS analysis

Native diffraction data have been recorded from crystals grown under conditions (ii) and (iii). For condition (ii), data were collected from a crystal extracted from the drop shown in Fig. 1(*c*). Diffraction was recorded to 2.4 Å resolution at beamline XS0A6 (SLS, Villingen) using a MAR CCD 165 detector and under cryoconditions in 27.5%(*v*/*v*) glycerol. Reflections were recorded in a single sweep of 70°, using 1.0° non-overlapping rotation steps. Crystals belong to space group *P*3. Unit-cell dimensions and data statistics are given in Table 1. The diffraction from these thin plate crystals suffered strongly from radiation damage, impeding the collection of complete data of high quality. Analysis of the crystal lattice using the self-rotation function in *AMoRe* (Navaza, 2001) revealed NCS relations at Eulerian angles $\alpha = 60$ and 180°, $\beta = \gamma = 0°$. This is consistent with the presence of three α-chains in the asymmetric unit aligned with the crystallographic *c* axis, which is in agreement with the fact that the dimension of this axis closely resembles the theoretical length calculated for this coiled-coil.

X-ray data optimal for structure elucidation were recorded to 1.6 Å resolution from crystals grown under condition (iii). Data collection took place at beamline ID14-2 (ESRF, Grenoble) on an ADSC Quantum 4 CCD detector and under cryoconditions using 20%(*v*/*v*) glycerol as a cryoprotectant. Reflections were recorded over a total rotation range of 105°, using 0.70° non-overlapping steps and two sweeps at high and low exposure doses (data statistics are given in Table 1). Crystals belong to space group $P2_12_12_1$ as confidently estimated from the pattern of systematic absences along the crystallographic axes. The calculated V_M value (Matthews, 1968) (Table 1) suggested the presence of more than one chain in the asymmetric unit of this crystal form. However, inspection of self-rotation functions and native Patterson maps did not reveal NCS relations.

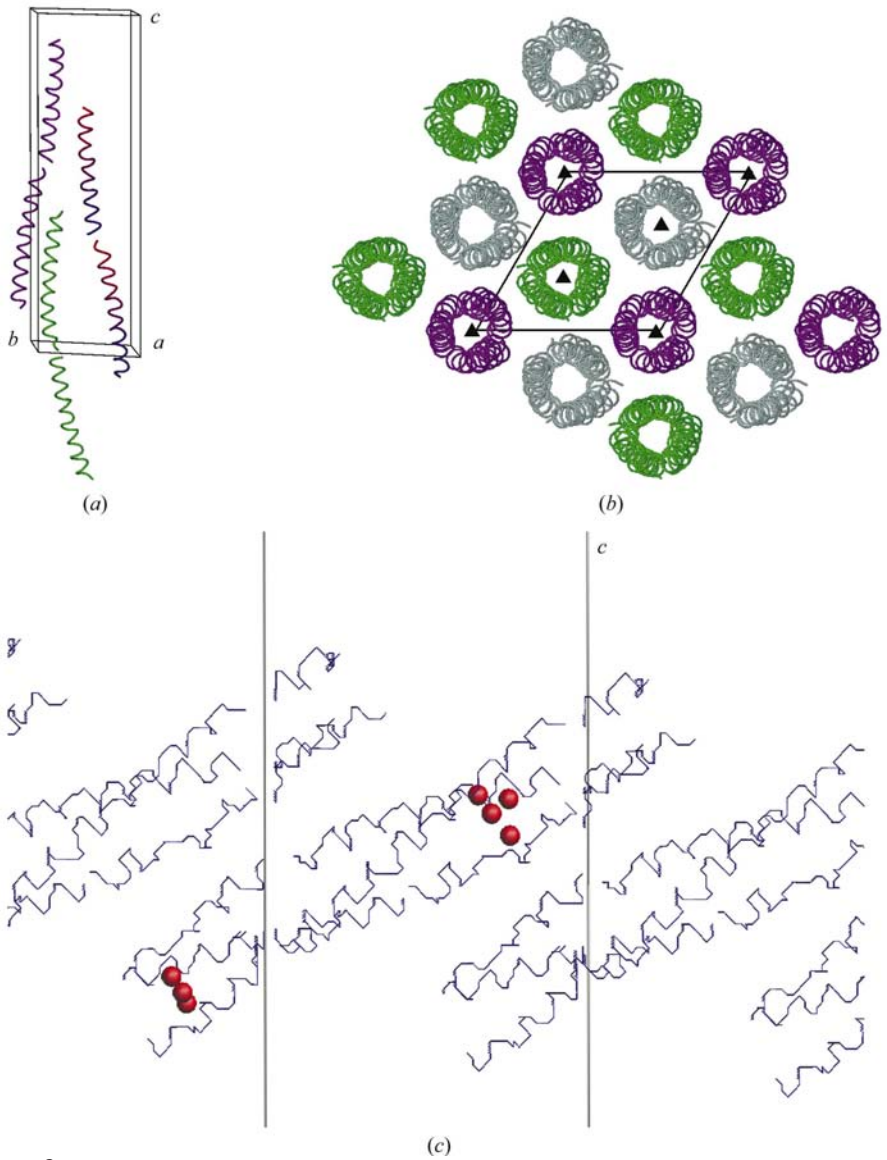

Figure 2
(*a*) Crystal lattices in the *P*3 and $P2_12_12_1$ space groups. The asymmetric unit of the *P*3 cell contains three α-helical chains aligned with the *c* crystallographic axis. Each chain originates from the fusion of two consecutive short helices as positioned by MR. To indicate this, one of the chains has its two halves coloured in a blue-to-red gradient, where blue indicates the N-termini and red the C-termini. (*b*) View of the *P*3 lattice perpendicular to the crystallographic threefold axis. Each NCS chain generates a full coiled-coil by applying crystal symmetry (reflected by different colouring). Crystallographic threefolds are indicated. (*c*) Experimental skeletons (after solvent treatment in *SHARP*) obtained by MAD phasing in $P2_12_12_1$ symmetry. SeMet sites located using MR phases and employed in phase calculation are shown as full circles. Five sites are unique, while one additional site has two related peaks, possibly owing to a double Met conformation.

6. Phasing

Phasing was initially attempted by molecular replacement (MR) in *AMoRe* (Navaza, 2001). Although DMPK-CC shares no sequence similarity with other coiled-coils for which structures are available, the structural similarity within this fold was expected to be sufficiently high to yield successful results in MR. This has been the case for other folds with a highly conserved architecture, such as the parallel β-helix, where correct solutions were obtained from search models with sequence identity values as low as 17% (Mayans *et al.*, 1997). In the current study, coiled-coil structures as well as their constituent single chains have been assayed for PDB entries 1avy, 1bb1, 1gzl, 1ij0 (trimeric models), 1ic2 (dimeric model),

1fe6 and 1gcl (tetrameric models). Only CC fragments without discontinuities were considered. All models were used as poly-alanine variants to account for the lack of sequence similarity. Because of the elongated shape of the models, a wide range of Patterson integration radii were tested. In general, short chains performed better than long ones, probably owing to long-range deviation along the coiled-coil axis between the search and target structures.

In space group $P3$, the best search model consisted of a single α-helical chain from 1ij0 (a trimeric variant of GCN4 leucine zipper 31 amino acids long). This yielded two best MR solutions (correlation coefficient 60.1% and R factor 52.2% after rigid-body fitting in *AMoRe* in the resolution range 12–3.5 Å), which corresponded to two different halves of the same α-helical chain of DMPK-CC (Fig. 2*a*). Both halves were fused to create an optimized search model 62 residues long. A new MR search resulted in three solutions corresponding to three NCS copies (correlation coefficient 68.2%, R factor 47.1%; Fig. 2*a*). Each chain generated a full trimeric coiled-coil by applying crystal symmetry. The lattice thus calculated (Fig. 2*b*) was in good agreement with NCS data for this crystal form.

In space group $P2_12_12_1$, trimeric 1ij0 was the best search model. Two solutions were obtained (correlation coefficient 53.3%, R factor 54.2% after rigid-body fitting in *AMoRe* in the resolution range 12–3.5 Å), each corresponding to half of DMPK-CC with a slight overlap along the helical axis. An equivalent MR solution was obtained when using the trimeric DMPK-CC model obtained by MR in the $P3$ space group as a search model.

Despite good crystal packing, the resultant MR phases were of insufficient quality to allow model building and refinement. Thus, experimental phasing was pursued by MAD on an SeMet derivative crystallized in space group $P2_12_12_1$. Diffraction was recorded at peak, inflection and high-energy remote points at ID14-2 (ESRF, Grenoble) (Table 1). Six selenium sites out of a total of nine (it should be noted that the three N-terminal Met residues are derived from the plasmid and are expected to be structurally disordered) have been located by anomalous difference Fourier ($>5\sigma$) using phases calculated from MR solutions (Fig. 2*c*). The sites within the core of the MR model are in good agreement with anomalous difference Patterson maps and are at the expected separation distance along the coil axis. Visual inspection of high-resolution electron-density maps calculated from the experimental phases using these selenium positions showed the MR solutions to be essentially correct and confirmed the trimeric state of DMPK-CC in these crystals. However, they also revealed an inversion of the α-helix termini in the MR model, which was probably a consequence of the lack of definition in the helix when removing side chains and at the resolution of the MR studies. This explains the low quality of these initial phases. Structure elucidation will now proceed making use of experimental phases. The fact that both crystal forms in this study, which were obtained at high and low salt conditions as well as at pH 4.6 and pH 7.5, contain a trimer in their asymmetric units leads us to speculate that the association state of this coiled-coil must be stable. However, further biophysical studies must be undertaken to establish the oligomeric state of this coiled-coil in solution and to resolve whether the currently observed arrangement might be the result of the crystallization media or of packing forces.

Very special thanks go to Professor Benjamin Perryman for facilitating the initial DNA coding for DMPK. We would like to acknowledge the support to this project from the Swiss National Foundation (grant 3100A0-100852) and the Association Française contre les Myopathies (grant No. 8421). Also, a most special mention should go to the staff of beamlines ID14-2 at ESRF and XS06A at SLS for providing exceptional assistance during data collection.

References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Chen, X. Q., Tan, I., Ng, C. H., Hall, C., Lim, L. & Leung, T. (2002). *J. Biol. Chem.* **277**, 12680–12688.
Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E. & Dotto, G. P. (1998). *J. Biol. Chem.* **273**, 29706–29711.
Dvorsky, R., Blumenstein, L., Vetter, I. R. & Ahmadian, M. R. (2004). *J. Biol. Chem.* **279**, 7098–7104.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Larkin, K. & Fardaei, M. (2001). *Brain Res. Bull.* **56**, 389–395.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., Pickersgill, R. & Jenkins, J. (1997). *Structure*, **5**, 677–689.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Riento, K. & Ridley, A. J. (2003). *Nature Rev. Mol. Cell Biol.* **4**, 446–456.
Shimizu, T., Ihara, K., Maesaki, R., Amano, M., Kaibuchi, K. & Hakoshima, T. (2003). *J. Biol. Chem.* **278**, 46046–46051.
Tan, I., Seow, K. T., Lim, L. & Leung, T. (2001). *Mol. Cell. Biol.* **21**, 2767–2778.
Wolf, E., Kim, P. S. & Berger, B. (1997). *Protein Sci.* **6**, 1179–1189.
Zhang, R. & Epstein, H. F. (2003). *FEBS Lett.* **546**, 281–287.