Crystallographic studies of the structured core domain of Knr4 fromSaccharomyces cerevisiae

Sylviane Julien, Patrick Tondl, Fabien Durand, Adilia Dagkessamanskaia, Herman van Tilbeurgh, Jean Marie François, Lionel Mourey, Didier Zerbib, Hélène Martin-Yken, Laurent Maveyraud

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Crystallographic studies of the structured core domain of Knr4 from *Saccharomyces cerevisiae*

Sylviane Julien, Patrick Tondl, Fabien Durand, Adilia Dagkessamanskaia, Herman van Tilbeurgh, Jean Marie François, Lionel Mourey, Didier Zerbib, Hélène Martin-Yken and Laurent Maveyraud

Acta Cryst. (2015). F**71**, 1120–1124

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Received 28 April 2015
Accepted 30 June 2015

Edited by R. A. Pauptit, Macclesfield, England

Keywords: intrinsically disordered proteins; hub proteins; *Saccharomyces cerevisiae*; cell-wall biogenesis regulation.

Crystallographic studies of the structured core domain of Knr4 from *Saccharomyces cerevisiae*

Sylviane Julien,[a,b] Patrick Tondl,[a,b] Fabien Durand,[c,d,e] Adilia Dagkessamanskaia,[c,d,e] Herman van Tilbeurgh,[f] Jean Marie François,[c,d,e] Lionel Mourey,[a,b] Didier Zerbib,[a,b,c,d,e] Hélène Martin-Yken[c,d,e] and Laurent Maveyraud[a,b]*

[a]Institut de Pharmacologie et de Biologie Structurale (IPBS), Centre National de la Recherche Scientifique (CNRS), 205 Route de Narbonne, BP 64182, 31077 Toulouse, France, [b]Université de Toulouse, Université Paul Sabatier, IPBS, 31077 Toulouse, France, [c]Université de Toulouse, INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, 31077 Toulouse, France, [d]INRA, UMR792, Ingénierie des Systèmes Biologiques et des Procédés, 31400 Toulouse, France, [e]CNRS, UMR5504, 31400 Toulouse, France, and [f]Institut de Biologie Intégrative de la Cellule, UMR9198, CNRS–Université Paris Sud, Bâtiment 430, 91400 Orsay, France. *Correspondence e-mail: laurent.maveyraud@ipbs.fr

The potentially structured core domain of the intrinsically disordered protein Knr4 from *Saccharomyces cerevisiae*, comprising residues 80–340, was expressed in *Escherichia coli* and crystallized using the hanging-drop vapour-diffusion method. Selenomethionine-containing (SeMet) protein was also purified and crystallized. Crystals of both proteins belonged to space group $P6_522$, with unit-cell parameters $a = b = 112.44$, $c = 265.21$ Å for the native protein and $a = b = 112.49$, $c = 262.21$ Å for the SeMet protein, and diffracted to 3.50 and 3.60 Å resolution, respectively. There are two molecules in the asymmetric unit related by a twofold axis. The anomalous signal of selenium was recorded and yielded an electron-density map of sufficient quality to allow the identification of secondary-structure elements.

1. Introduction

Intrinsically disordered proteins (IDPs) are involved in numerous essential biological processes (Dunker *et al.*, 2002, 2008). Their functions often rely on versatile interactions, with specific binding partners contributing to their regulation or to the assembly of supramolecular complexes. The conformational flexibility of IDPs gives them the ability to be involved in one-to-many binding (where a single disordered domain is able to bind several structurally diverse partners). Highly connected proteins located at nodes in interaction networks are termed 'hubs', and they are significantly enriched in disordered domains (Dunker *et al.*, 2005; Uversky, 2013). Hence, hub proteins might actually fulfil their multiple interactions and signalling functions thanks to their intrinsically disordered and adaptive nature.

The Knr4/Smi1 protein is specifically found in the fungal kingdom, and evidence from *Saccharomyces cerevisiae* (Basmaji *et al.*, 2006; Dagkessamanskaia, Durand *et al.*, 2010; Dagkessamanskaia, El Azzouzi *et al.*, 2010; Durand *et al.*, 2008; Martin *et al.*, 1999; Martin-Yken *et al.*, 2003) indicates that the *S. cerevisiae KNR4* gene and its counterparts from the pathogenic yeast *Candida albicans* (*CaSMI1*) and the model fungus *Neurospora crassa* (*NcGS1*) are involved in cell-wall biogenesis and morphogenesis. *CaSMI1* is involved in

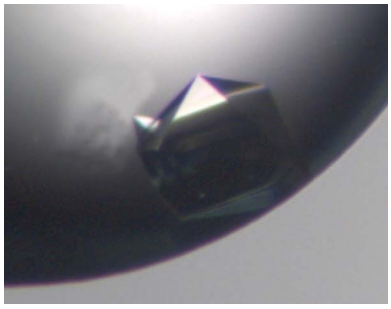


Table 1
Cloning and expression of the recombinant protein.

Source organism	*S. cerevisiae* strain W303-1A *MAT***a** *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* (ATCC 208352)
DNA source	Genomic DNA
Forward primer†	GGATCCCATATGTCCACGGAGTCAAACGATGGTG-TC
Reverse primer†	AAGCTTCTCGAGTTGATACTTGATCCACGTTCTT-CT
Expression vector	pGEX-6P-3
Expression host	*E. coli* BL21 (DE3) Codon Plus, *E. coli* B834 (DE3)
Complete amino-acid sequence of the construct produced‡	GPLGSHMSTESNDGVSETLLAWRHIDFWTSEHNP-DLNATLSDPCTQNDITHAEEDLEVSFPNPVKA-SFKIHDGQEDLESMTGTSGLFYGFQLMTLDQV-VAMTQAWRNVAKNLNKRSQQGLSHVTSTGSSS-SMERLNGNKFKLPNIPDQKSIPPNAVQPVYAH-PAWIPLITDNAGNHIGVDLAPGPNGKYAQIIT-FGRDFDTKFVIAENWGEFLLSFANDLEAGNWY-LVDDNDDYFSGDGELVFRDKKSNGPIQDYFEV-LKRRTWIKYQLERPHRD

† BamHI and XhoI restriction sites are underlined. ‡ Non-Knr4 amino acids are underlined.

resistance phenotypes toward drugs of medical importance and in biofilm β-glucan matrix production (Nett *et al.*, 2011) and is induced in hyphal cells (Candida Genome Database; http://www.candidagenome.org/), which highlights the interest in this protein as a possible drug target. Recent analysis of sequence similarity indicates that Knr4/Smi1 proteins might be structurally related to even more distant gene products from the bacterial kingdom, including a newly described family of contact-dependent inhibitory toxin systems, and suggests that their genes may even have reached the eukaryote kingdom through viruses (Zhang *et al.*, 2011). Thus, the origins and evolutionary pattern of these proteins establish their great interest based on their structural specificities/characteristics and the advantages conferred by their ability to associate with different protein partners.

Knr4 is the best studied member of the Knr4/Smi1 proteins. Data available on its cellular function show that it influences gene transcription (Martin *et al.*, 1999; Martin-Yken *et al.*, 2003), cell-cycle progression (Martin-Yken *et al.*, 2002) and morphogenesis (Riquelme *et al.*, 2011) and that it potentially interacts with numerous partners (Basmaji *et al.*, 2006). In most *S. cerevisiae* genetic backgrounds, including the widely used reference strain BY4741, *KNR4* deletion leads to growth defects under stress conditions such as elevated temperature and the presence of SDS, caffeine, various antibiotics or cell wall-affecting drugs such as calcofluor white and Congo red. Even if the *KNR4* gene is not, *per se*, an essential gene in optimal laboratory conditions (303 K, YPD medium), many synthetic lethal mutants have been isolated, revealing the central role of Knr4 at a node position connecting several essential pathways (Basmaji *et al.*, 2006; Cherry *et al.*, 2012).

The three-dimensional structure of fungal Knr4 is unknown to date. *In silico* analysis, combined with biophysical and biochemical methods, has shown that it contains large disordered domains. The protein displays unstructured regions, mainly in both the N-terminal (1–80) and the C-terminal (341–505) parts, while the central core is supposed to be structured and globular (Durand *et al.*, 2008). The central core (80–340) contains the essential biological functions of the protein, as attested by its ability to fully complement *S. cerevisiae* null mutant phenotypes. The N- and C-terminal parts are involved in ensuring and controlling the interactions of Knr4 with its protein partners. The physical interaction of the N-terminal domain (1–80) with protein partners is necessary for correct localization of the protein (Dagkessamanskaia, El Azzouzi *et al.*, 2010; Liu *et al.*, 2015). On the contrary, the disordered C-terminal domain (341–505) inhibits these interactions. Hence, *KNR4*/*SMI1* codes for hub proteins, which are unique to the fungal kingdom, the particular structural characteristics of which provide the ability to interact with several partners and to ensure specific functions. Elucidating the structure of these IDPs is a real challenge and of great relevance owing to the fact that Knr4, as a hub protein of the yeast interactome involved in several central processes and being specific to fungi, may be an attractive and efficient antifungal target. Given the unstructured nature of both the N-terminal and C-terminal domains of the protein, deciphering the structure of the central functional core (80–340) represents the first step towards the achievement of this challenge.

2. Materials and methods

2.1. Expression and purification

The gene encoding residues 80–340 of Knr4 was amplified from yeast genomic DNA and cloned into pGEX-6P-3 as a C-terminal fusion with glutathione *S*-transferase (GST). The primers, relevant restriction sites and amino-acid sequence of the recombinant protein are indicated in Table 1.

2.1.1. Native Knr4(80–340) protein. The pGEX-6P-3::*KNR4*(80–340) vector was transformed into *Escherichia coli* BL21 (DE3) Codon Plus competent cells (Novagen). The cells were grown in LB medium supplemented with ampicillin (150 mg l^{-1}) at 310 K. The production of GST-Knr4(80–340) was induced by the addition of 0.1 m*M* isopropyl β-D-1-thio-galactopyranoside (IPTG) at an $OD_{600\,nm}$ of 0.8. The induction was continued overnight at 298 K. The cells were collected and resuspended in lysis buffer (10 m*M* Tris–HCl, 1 m*M* EDTA, 100 m*M* NaCl pH 8.0) in the presence of 100 mg l^{-1} lysozyme and lysed by sonication. The lysate was clarified by centrifugation (10 000*g*, 12 min, 277 K), and Triton X-100 (2% final concentration) was added before incubation with 2 ml glutathione Sepharose 6B beads (GE Healthcare) equilibrated in lysis buffer. After 30 min, the beads were washed three times (50 m*M* Tris–HCl, 150 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT pH 7.0) before the addition of PreScission protease (1 unit per 100 µg protein, GE Healthcare). The Knr4(80–340) protein was eluted by three successive washes (1 ml) with the washing buffer. The protein was applied onto a HiLoad 16/60 Superdex 75 column equilibrated in 50 m*M* Tris-HCl, 150 m*M* NaCl, 2 m*M* TCEP, 0.2 m*M* AEBSF pH 7.5 and was eluted with the same buffer. The peak was collected and

Table 2
Crystallization.

Protein	Native Knr4(80–340)	SeMet-Knr4(80–340)
Method	Hanging drop	Hanging drop
Plate type	Linbro, 24-well	Linbro, 24-well
Temperature (K)	277	277
Protein concentration (mg ml^{-1})	10.5	6.3
Buffer composition of protein solution	0.15 *M* NaCl, 0.1 *M* Tris–HCl pH 7.5	0.15 *M* NaCl, 0.1 *M* Tris–HCl pH 7.5
Composition of reservoir solution	1.2 *M* ammonium sulfate, 0.1 *M* MES–NaOH pH 5.5	1.4 *M* ammonium sulfate, 0.1 *M* MES–NaOH pH 6.0
Volume and ratio of drop	2 µl + 2 µl	2 µl protein solution + 1 µl reservoir solution
Volume of reservoir (µl)	500	500

the protein was concentrated using Vivaspin ultracentrifugation devices (Sartorius; molecular-weight cutoff 10 kDa).

2.1.2. Selenomethionylated Knr4(80–340). The pGEX-6P-3::*KNR4*(80–340) vector was transformed into *E. coli* B834 (DE3) competent cells (Novagen). The strain was adapted to M9 medium containing L-selenomethionine (SeMet) by performing five successive cultures at 310 K containing an increasing proportion of M9-SeMet medium and a decreasing proportion of LB medium. The final culture was performed at 310 K in 100% M9 medium with SeMet (44 mg l^{-1}) until the optical density reached 0.5, and was induced with 1 m*M* IPTG. The temperature was then lowered to 303 K in order to allow protein production for 12 h. The purification of SeMet-Knr4(80–340) was achieved using same protocol as used for the native protein.

2.2. Crystallization

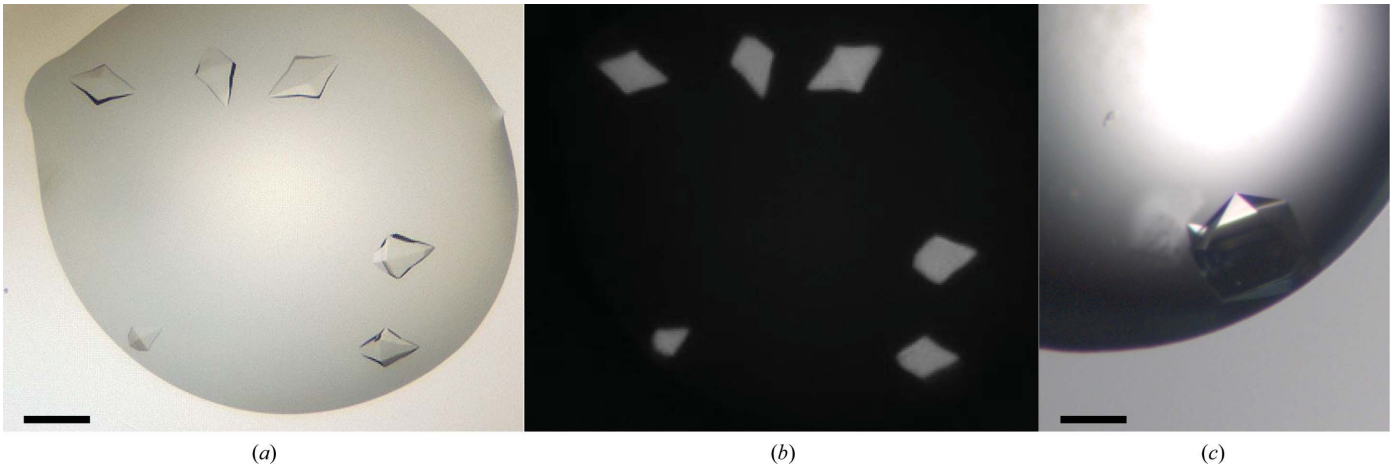
Initial crystallization conditions were identified using commercial screening kits from Qiagen (The Classics, AmSO4, Anions, Cations, JCSG Core I–IV, PEGs, PEGs II, MPD, pHClear and pHClear II Suites) with native Knr4(80–340). Sitting drops were formed by mixing 100–200 nl native protein at 6–12 mg ml^{-1}, as evaluated by absorbance at 280 nm (using a sequence-derived extinction coefficient of 48 930 l mol^{-1} cm^{-1}), with the same volume of reservoir solution using a NanoDrop ExtY crystallization robot (Innovadyne) at 277, 285 and 293 K. The reservoir volume was 80 µl. Drops were automatically imaged with normal and UV light using a Rock Imager RI-1000 (Formulatrix). Optimization of the crystallization conditions was performed manually using the hanging-drop method by incubating 2–4 µl drops on siliconized glass slides with a 500 µl reservoir.

2.3. Data collection and processing

Cryoprotection of crystals was achieved by brief immersion into the crystallization solution supplemented with 20% ethylene glycol prior to transfer into a gaseous nitrogen flux at 100 K. Diffraction data for the native protein were collected to a 3.5 Å resolution limit on the ID14-EH2 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble at 0.933 Å wavelength using an ADSC Q4 detector. 90 oscillations of 1° were collected with an exposure time of 10 s and a crystal-to-detector distance of 337 mm.

Crystals of selenomethionylated protein were cryoprotected similarly to crystals of native protein. Data were collected on the ID29 beamline at ESRF. A fluorescence scan was performed in order to determine the most appropriate wavelength for Se-SAD data collection. 290 oscillations of 1° were collected at a wavelength of 0.97930 Å (12 661 eV) with an exposure time of 1 s. The incident beam was attenuated to 15%. The crystal-to-detector distance was set to 455 mm. The detector was an ADSC Q315r.

Data processing was initially performed automatically with *autoPROC* (Vonrhein *et al.*, 2011) and was further optimized with *XDS* (Kabsch, 2010). All subsequent operations were



(*a*) (*b*) (*c*)

Figure 1
Crystals of native Knr4(80–340) (*a*) under visible light and (*b*) under UV light. (*c*) Crystal of SeMet-Knr4(80–340). The bar corresponds to 200 µm.

performed with *CCP*4 (Winn *et al.*, 2011), *PHENIX* (Adams *et al.*, 2011) and *SHELX* (Sheldrick, 2010).

3. Results and discussion

The production and purification of native Knr4(80–340) yielded about 2 mg pure protein per litre of culture. In the case of SeMet-Knr4(80–340), 200 µg pure protein was purified from 1 l culture.

Native Knr4(80–340) crystals were observed after 2 d at 285 K in solution No. 73 from the JSCG Core II Suite screening kit (2.0 *M* ammonium sulfate, 5% 2-propanol), as well as in some solutions from The PEGs and PEGs II Suites (25–30% PEG 3000, PEG 4000 and PEG 6000, pH 8.0–9.0). Only crystals obtained in ammonium sulfate could be optimized to give diffraction-quality crystals. The best diffracting crystals were obtained with a protein concentration of 10.5 mg ml^{-1} using 1.2 *M* ammonium sulfate in 0.1 *M* MES buffer pH 5.5 at 277 K (Figs. 1*a* and 1*b*, Table 2). Crystals of the selenomethionylated protein were obtained at 277 K by mixing 2 µl protein solution at a concentration of 6.3 mg ml^{-1} with 1 µl 1.0–1.4 *M* ammonium sulfate, 0.1 *M* MES buffer pH 5.0–6.0. Crystals grew within a week (Fig. 1*c*, Table 2).

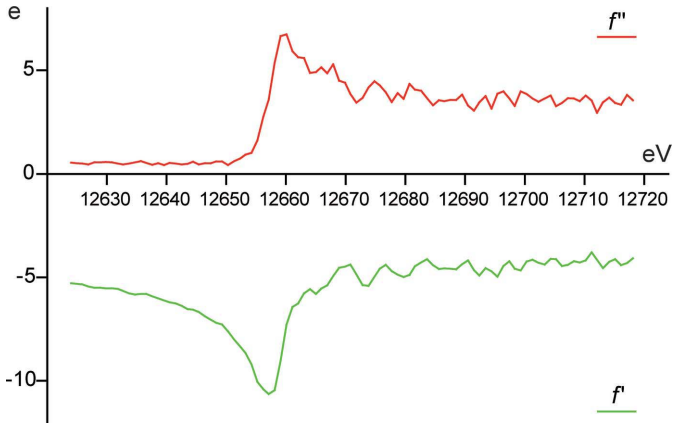


Figure 2
Fluorescence spectrum collected from SeMet-Knr4(80–340) crystals on beamline ID29 at the ESRF (f'', red). The Kramers–Kronig transformation (f') is represented in green.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell. The statistics for the SeMet-Knr4 crystals were computed with unmerged Friedel pairs.

Protein	Native Knr4(80–340)	SeMet-Knr4(80–340)
Diffraction source	ID14-EH2, ESRF	ID29, ESRF
Wavelength (Å)	0.933	0.9793
Temperature (K)	100	100
Detector	ADSC Q4	ADSC Q315r
Crystal-to-detector distance (mm)	337	455
Rotation range per image (°)	1	1
Total rotation range (°)	90	290
Exposure time per image (s)	10	1
Space group	$P6_522$	$P6_522$
a, b, c (Å)	112.44, 112.44, 265.21	112.49, 112.49, 262.21
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution range (Å)	50.0–3.50 (3.55–3.50)	50.0–3.60 (3.65–3.60)
Total No. of reflections	139639 (5713)	343114 (4531)
No. of unique reflections	13180 (522)	21456 (868)
Completeness (%)	99.8 (100.0)	99.9 (99.8)
Multiplicity	10.6 (10.9)	16.0 (5.2)
$\langle I/\sigma(I)\rangle$	32.7 (2.6)	23.1 (1.9)
R_{meas}	0.047 (1.148)	0.082 (1.061)
$CC_{1/2}$	100.0 (82.6)	99.9 (58.3)
Overall *B* factor from Wilson plot (Å^2)	122.6	128.8

Data-processing statistics are given in Table 3. Native Knr4(80–340) crystals belonged to space group $P6_122$ or $P6_522$

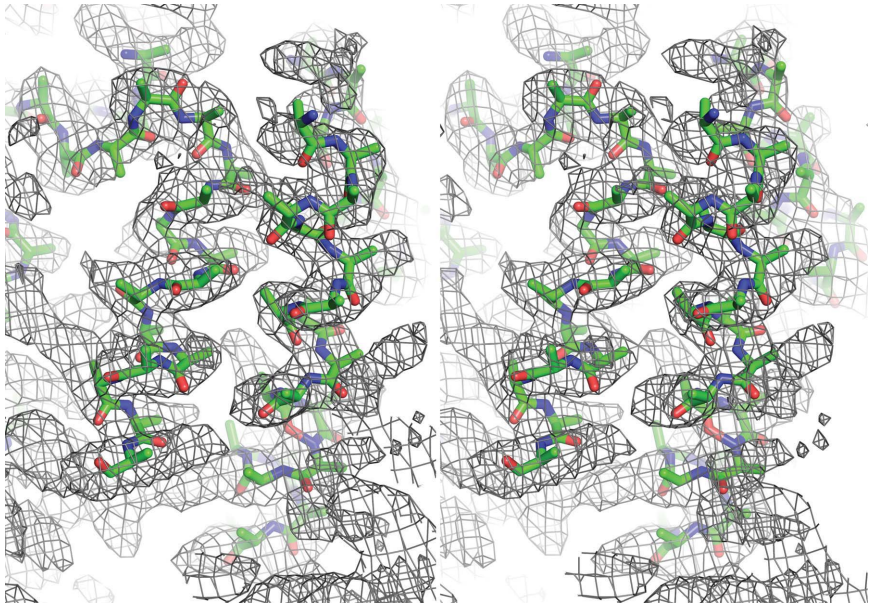


Figure 3
Stereoview of the electron-density map obtained after SAD phasing with *SHELXC*/*D*/*E* and further density modification with *DM*. The electron-density map is displayed as a grey mesh contoured at 1σ. Constructed α-helices and β-strands are displayed as polyalanines.

according to scaling statistics and to systematic absences. Analysis of the asymmetric unit contents indicated that three molecules were likely to be present, with a Matthews coefficient of 2.62 $Å^3$ Da^{-1} and a solvent content of 53% (the Matthews coefficients for one, three and four molecules in the asymmetric unit are 7.85, 3.92 and 1.96 $Å^3$ Da^{-1}, respectively, with solvent contents of 84.3, 68.7 and 37.4%, respectively). A self-rotation function and a self-Patterson map were computed, but did not allow the identification of rotational or translational NCS. SeMet-Knr4(80–340) crystals displayed similar unit-cell parameters to those of the native protein, with a 3 Å shorter *c* axis. They diffracted to 3.60 Å resolution. The fluorescence spectrum collected before data collection (Fig. 2) confirmed the presence of Se atoms in the crystals.

SHELXC/*D*/*E* was used to determine the selenium substructure for phasing as well as for density modification. According to the $\langle d''/\mathrm{sig}\rangle$ analysis performed by *SHELXC*, significant anomalous signal is present up to 4.4 Å resolution. *SHELXD* identified 11 selenium sites, while there are five expected seleniums per molecule, resulting in CC_{all}/CC_{weak} values of 51.57/32.94. Only the first eight sites were conserved for phasing with *SHELXE*, as the occupancies of the following sites were below 0.15. The resulting modified electron-density map suggested that the correct space group was $P6_522$, as indicated by the map correlation coefficients and the contrast values provided by *SHELXE* (56.0 and 0.798, respectively, for space group $P6_522$ and 30.6 and 0.193, respectively, for space group $P6_122$). After further density modification with *DM* (Cowtan, 2010), the electron-density map allowed the identification of several α-helices and β-strands (Fig. 3). According to the fragments of protein that were constructed in this electron-density map, there are two molecules in the asymmetric unit related by a twofold axis. Completion of the model and crystallographic refinement are currently in progress.

Acknowledgements

We acknowledge the ESRF for the provision of synchrotron-radiation facilities. The crystallization equipment used in this study is part of the Integrated Screening Platform of Toulouse (PICT, IBiSA). This work has been partially supported by grants from the EU (6th FP, Fungcellwall project) and Region Midi Pyrénées (grant No. 10051296) to JMF.

References

Adams, P. D. *et al.* (2011). *Methods*, **55**, 94–106.
Basmaji, F., Martin-Yken, H., Durand, F., Dagkessamanskaia, A., Pichereaux, C., Rossignol, M. & Francois, J. (2006). *Mol. Genet. Genomics*, **275**, 217–230.
Cherry, J. M. *et al.* (2012). *Nucleic Acids Res.* **40**, D700–D705.
Cowtan, K. (2010). *Acta Cryst.* D**66**, 470–478.
Dagkessamanskaia, A., Durand, F., Uversky, V. N., Binda, M., Lopez, F., El Azzouzi, K., Francois, J. M. & Martin-Yken, H. (2010). *Protein Sci.* **19**, 1376–1385.
Dagkessamanskaia, A., El Azzouzi, K., Kikuchi, Y., Timmers, T., Ohya, Y., François, J.-M. & Martin-Yken, H. (2010). *Yeast*, **27**, 563–574.
Dunker, A. K., Brown, C. J. & Obradovic, Z. (2002). *Adv. Protein Chem.* **62**, 25–49.
Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. (2005). *FEBS J.* **272**, 5129–5148.
Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. (2008). *Curr. Opin. Struct. Biol.* **18**, 756–764.
Durand, F., Dagkessamanskaia, A., Martin-Yken, H., Graille, M., Van Tilbeurgh, H., Uversky, V. N. & François, J. M. (2008). *Yeast*, **25**, 563–576.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Liu, J., Formosa, C., Schiavone, M., Dague, E., François, J. M. & Yken-Martin, H. (2015). In the press.
Martin, H., Dagkessamanskaia, A., Satchanska, G., Dallies, N. & François, J. (1999). *Microbiology*, **145**, 249–258.
Martin-Yken, H., Dagkessamanskaia, A., Basmaji, F., Lagorce, A. & Francois, J. (2003). *Mol. Microbiol.* **49**, 23–35.
Martin-Yken, H., Dagkessamanskaia, A., Talibi, D. & Francois, J. (2002). *Curr. Genet.* **41**, 323–332.
Nett, J. E., Sanchez, H., Cain, M. T., Ross, K. M. & Andes, D. R. (2011). *Eukaryot. Cell*, **10**, 1660–1669.
Riquelme, M. *et al.* (2011). *Fungal Biol.* **115**, 446–474.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Uversky, V. N. (2013). *Protein Sci.* **22**, 693–724.
Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011). *Acta Cryst.* D**67**, 293–302.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhang, D., Iyer, L. M. & Aravind, L. (2011). *Nucleic Acids Res.* **39**, 4532–4552.