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

Crystallization and preliminary X-ray diffraction analysis of the C-terminal domain of the human spliceosomal DExD/H-box protein hPrp22

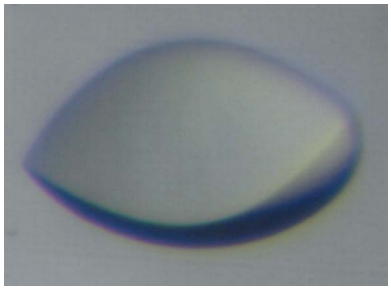
Denis Kudlinzki,‡ Christian Nagel and Ralf Ficner*

Institut für Mikrobiologie und Genetik, Abteilung Molekulare Strukturbiologie, Georg-August-Universität Göttingen, Justus-von-Liebig Weg 11, 37077 Göttingen, Germany

‡ Current address: Institut für Organische Chemie und Chemische Biologie, Abteilung Strukturelle Chemie und Biologie/NMR-Spektroskopie, Goethe-Universität Frankfurt-am-Main, Max-von-Laue Strasse 7, 60438 Frankfurt, Germany.

Correspondence e-mail: rficner@uni-goettingen.de

Received 18 June 2009
Accepted 12 August 2009



The *Homo sapiens* DExD/H-box protein hPrp22 is a crucial component of the eukaryotic pre-mRNA splicing machinery. Within the splicing cycle, it is involved in the ligation of exons and generation of the lariat and it additionally catalyzes the release of mature mRNA from the spliceosomal U5 snRNP. The yeast homologue of this protein, yPrp22, shows ATP-dependent RNA-helicase activity and is capable of unwinding RNA/RNA duplex molecules. A truncated construct coding for residues 950–1183 of human Prp22, comprising the structurally and functionally uncharacterized C-terminal domain, was cloned into an *Escherichia coli* expression vector. The protein was subsequently overproduced, purified and crystallized. The crystals obtained diffracted to 2.1 Å resolution, belonged to the tetragonal space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 78.2$, $c = 88.4$ Å, and contained one molecule in the asymmetric unit.

1. Introduction

Splicing of pre-mRNA is a highly dynamic process that occurs in all eukaryotic cells. It is catalyzed by the spliceosome, which excises the introns and ligates the exons of pre-mRNAs (Burge *et al.*, 1999; Will & Lührmann, 2006; Wahl *et al.*, 2009). Rearrangements of the RNA–RNA, RNA–protein and protein–protein interactions lead to conformational changes, which result in proper splice-site recognition, activation of the catalytically active spliceosome and the release of the products during spliceosome disassembly (Madhani & Guthrie, 1994; Burge *et al.*, 1999). In addition to the U snRNAs, a multitude of proteins are involved in the splicing cycle (Makarov *et al.*, 2002). Some of them are mobile factors, while others are integral parts of the spliceosomal U snRNPs. Most of these proteins are not directly involved in the catalytic process. Rather, they mediate the establishment of protein–protein and protein–RNA contacts or catalyze reactions which result in rearrangements of the spliceosome. One class of spliceosomal proteins are the ATPases of the DExD/H-box family. They drive structural rearrangements of the spliceosome (Staley & Guthrie, 1998). These DExD/H-box proteins are members of helicase superfamily II and are characterized by seven motifs which are not only important for ATP binding and hydrolysis but also for RNA binding (Gorbalenya & Koonin, 1993; Hall & Matson, 1999; Tanner & Linder, 2001). One member of this protein family is the DEAH-box protein Prp22. The yeast orthologue yPrp22 is essential for cell growth of *Saccharomyces cerevisiae* (Schwer & Meszaros, 2000; Schneider *et al.*, 2002). In fact, there is evidence that yPrp22 has an ATP-dependent RNA-helicase activity and it has been shown to unwind RNA duplexes (Tanaka & Schwer, 2005). Its main function is the release of mature mRNA from the U5 snRNP of the post-spliceosomal complex (Company *et al.*, 1991; Wagner *et al.*, 1998; Schneider *et al.*, 2004). Prp22 is also involved in the second transesterification step of splicing, *i.e.* the ligation of exons and generation of the lariat (Schwer & Gross, 1998). Additionally, it has exon proofreading activity (Mayas *et al.*, 2006). The functional human orthologue of yPrp22 (hPrp22, DHX8, HRH1; NCBI accession No. NP_004932) has been identified (Ono *et al.*, 1994) and further

investigated (Ohno & Shimura, 1996). Prp22 exhibits a modular domain composition that is comparable to those of other spliceosomal DExD/H-box proteins. Among the spliceosomal DExD/H-box proteins, Prp2, Prp16, Prp22 and Prp43, which are involved in splicing steps subsequent to the spliceosome activation, share a conserved C-terminal domain (CTD) proximal to the helicase core. However, this C-terminal domain has not yet been functionally and structurally characterized. Here, we report the preparation, crystallization and preliminary X-ray analysis of the hPrp22 CTD (residues 950–1183).

2. Materials and methods

2.1. Cloning and protein overproduction

The C-terminal hPrp22 truncation was designed using the programs *SMART* (Schultz *et al.*, 1998) to define the borders of the CTD and *PSIPRED* (Jones, 1999; Bryson *et al.*, 2005) to predict secondary-structure motifs. The truncation sites of hPrp22(950–1183) are close to the start and end of α-helices, respectively. The coding sequence of hPrp22(950–1183) containing the predicted C-terminal domains HA2 (helicase-associated domain 2) and DUF1605 (domain of unknown function 1605) was amplified by the polymerase chain reaction (PCR) with forward primer 5′-TAT**GGATCC**CCTATGGAAACTTTGATCAC-3′ and reverse primer 5′-ATA**CTCGAG**GACCTTGAAGAAGGCTG-3′ using Phusion DNA polymerase (New England Biolabs Inc., USA) and the MegaMan commercial human transcriptome library (Stratagene, USA) as a template. The PCR product was cloned into *Escherichia coli* expression vector pPR-IBA1 (IBA, Germany) using the *Bam*HI and *Xho*I restriction sites introduced by the PCR primers. The correct sequence of the construct was confirmed by DNA sequencing. *E. coli* strain BL21 (DE3) was transformed with the pPR-IBA1-hPrp22(950–1183) plasmid. The cells were grown in auto-inducing ZYM-5052 medium (Studier, 2005) containing 100 µg ml^{-1} ampicillin. To optimize cell growth, a cool-down protocol was established. A 5 ml pre-culture was diluted with 300 ml ZYM-5052 and incubated at 310 K until an OD_{600} of 1–1.5 was reached. At this point the incubation temperature was reduced to 291 K. Cells were harvested by centrifugation at 4800*g* for 15 min at 277 K.

2.2. Protein purification

The cells were resuspended in lysis buffer (100 m*M* Tris–HCl pH 8.5, 2 *M* LiCl, 150 m*M* NaCl, 2 m*M* DTT, 2 m*M* EDTA) and lysed by microfluidization. The lysate was centrifuged at 100 000*g* for 30 min at 277 K. The supernatant was applied onto a 15 ml Strep-Tactin Sepharose column (IBA GmbH, Germany) equilibrated with buffer *A* (50 m*M* Tris–HCl pH 8.5, 150 m*M* NaCl, 2 m*M* EDTA). After washing the column with buffer *A*, the protein was eluted with buffer *B* (buffer *A* with 4 m*M* D-desthiobiotin). HPrp22-CTD was further purified using a HiLoad XK26/60 Superdex 75 column (GE Healthcare, USA); 10 m*M* Tris–HCl pH 8.5, 150 m*M* NaCl, 1 m*M* EDTA was used as eluant. All purification steps were performed at 277 K using an ÄKTA Prime system (GE Healthcare, USA).

2.3. Crystallization

Freshly purified hPrp22(950–1183) containing a C-terminal Strep-tag was concentrated to 7 mg ml^{-1} by ultrafiltration using Vivaspin6 centrifugal concentrators (MWCO 10000 PES; Sartorius, Germany) and subjected to Crystal Screen I, II, Lite, Cryo, PEG/Ion (Hampton Research, USA), Magic Screen (BioGenova, Canada) and JBScreen 1–10 (Jena Bioscience GmbH, Germany) as initial screening kits. Crystallization trials were performed at 277 and 293 K using the sitting-drop vapour-diffusion method on 24-well Cryschem plates

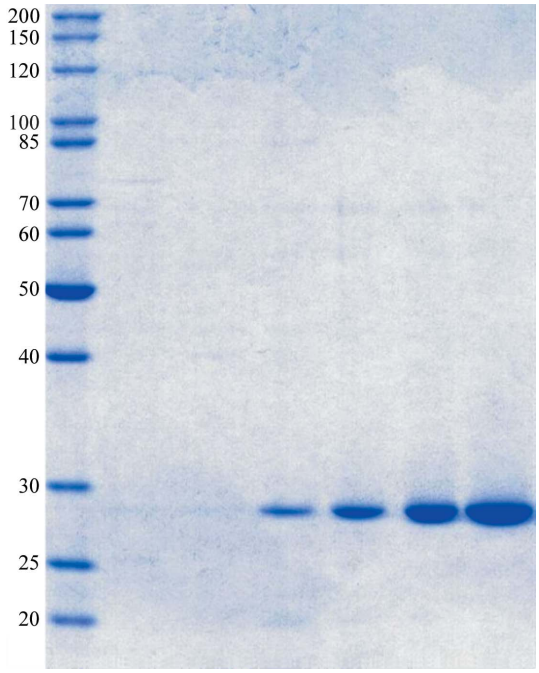


Figure 1
12.5% SDS–PAGE of hPrp22(950–1183) stained with Coomassie Brilliant Blue. The left lane contains electrophoresis standards with molecular masses in kDa. The other lanes contain samples of purified hPrp22(950–1183) after gel filtration on XK26/60 Superdex 75 with increasing concentrations.

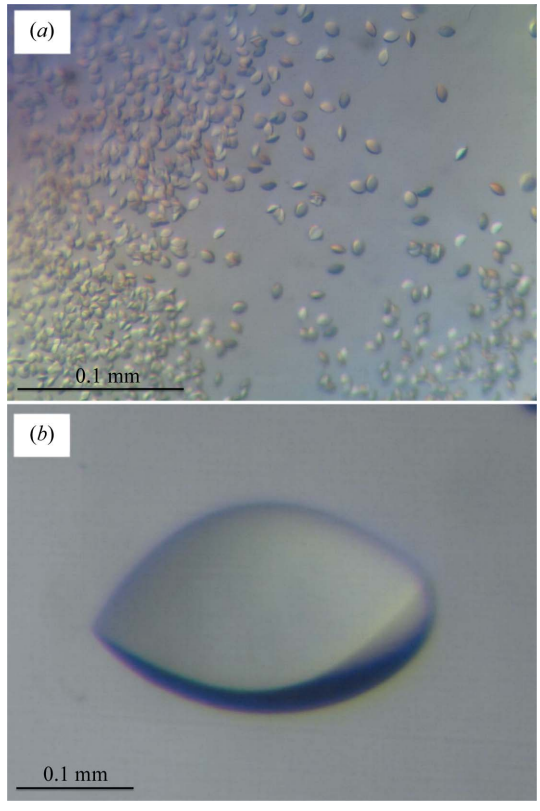


Figure 2
Lens-shaped crystals of hPrp22(950–1183) grown from a sitting drop on a 24-well plate. (*a*) Initial microcrystals (length 0.01 mm). (*b*) Optimized crystal. The crystal has approximate dimensions 0.3 × 0.15 × 0.15 mm.

Table 1
Data-collection statistics for hPrp22(950–1183) on beamline BW7B at EMBL/DESY, Hamburg.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	0.8423
Resolution (Å)	47–2.1 (2.16–2.1)
Completeness (%)	99.8 (100)
$R_{\rm merge}$† (%)	5.8 (35.9)
$R_{\rm p.i.m.}$‡ (%)	2.3 (14.2)
Mean $I/\sigma(I)$	22.1 (4.4)
Space group	$P4_12_12$ or $P4_32_12$
Unit-cell parameters (Å, °)	$a = b = 78.2$, $c = 88.4$, $\alpha = \beta = \gamma = 90$
No. of observed reflections	135241
No. of unique reflections	16819
Molecules per ASU	1
$V_{\rm M}$ (Å^3 Da^{-1})	2.17
Redundancy	7.2
Mosaicity (°)	0.5
Mean B value (Å^2)	31.8
Solvent content (%)	43.3

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections. ‡ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections. It is an indicator of the precision of the final merged and averaged data set (Weiss & Hilgenfeld, 1997). The R values were calculated using *XPREP*.

(Hampton Research, USA). A mixture of 1 µl protein solution and 1 µl reservoir solution was equilibrated against 500 µl reservoir solution.

2.4. Data collection

For data collection, crystals were transferred into a cryobuffer containing 1.5 *M* ammonium sulfate and 25%(*v*/*v*) glycerol. X-ray diffraction data were collected on a MAR 345 mm image-plate detector on beamline BW7B at EMBL/DESY, Hamburg. The crystals were flash-cooled and maintained in a 100 K nitrogen-gas stream (Oxford Instruments Inc., England) during data collection. The highest quality data set was collected using a rotation angle of 1° per image at a distance of 200 mm and a wavelength of 0.8423 Å. The data were integrated using *HKL*-2000 (HKL Research Inc., USA). Scaling was performed with *SADABS* (G. Sheldrick, Universität Göttingen) and the space group was determined with *XPREP* (Bruker AXS, Germany).

3. Results and discussion

The purified construct of the CTD of hPrp22 comprises 270 residues, of which 234 originate from the CTD. The bases coding for the residual 36 residues were introduced in the cloning procedure: 15 amino acids at the N-terminus (MGDRGPEFELGTRGS) and 21 amino acids at the C-terminus (LEVDLQGDHGLSAWSHPQFEK). The purified protein has a molecular mass of 31.2 kDa and a theoretical isoelectric point of 7.7. hPrp22(950–1183) overproduced in *E. coli* BL21 (DE3) harbouring the recombinant plasmid was purified to homogeneity in two steps. In the final purification step, the protein eluted at 170 ml in size-exclusion chromatography, corresponding to a monomeric state, and its purity was examined by 12.5% SDS–PAGE (Fig. 1). 1 l culture yielded about 5 mg hPrp22(950–1183) with a purity of >95% (Fig. 1). SDS–PAGE showed that the overproduced protein has an apparent molecular mass of 29 kDa. In an initial crystallization screen, numerous lens-shaped microcrystals (Fig. 2*a*) were obtained under several conditions containing high concentrations of ammonium sulfate (Crystal Screen I condition No. 39, JBScreen 6 condition Nos. 10 and 16 and Magic Screen condition No. 62). After optimizing the initial crystallization conditions, single clear lens-shaped crystals with dimensions of 0.3 × 0.15 × 0.15 mm were obtained in 1.5 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8 at 277 K after 2 d (Fig. 2*b*) and 1.7 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.5 at 293 K after 3 d. After equilibrating the crystals in cryobuffer (1.5 *M* ammonium sulfate, 25% glycerol), diffraction data could be recorded to a resolution of 2.1 Å. The crystals belonged to the enantiomorphic space groups $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 78.2$, $c = 88.4$ Å. The collected diffraction data set was 99.8% complete, with a redundancy of 7.2 and an $I/\sigma(I)$ of 22.1. Data-collection statistics are summarized in Table 1. The Matthews coefficient ($V_{\rm M}$) of 2.17 Å^3 Da^{-1} and the solvent content of 43.3% (Matthews, 1968) suggested the presence of one molecule per asymmetric unit. Since a search for a homologous structure with 3*D-JIGSAW* (Bates *et al.*, 2001) failed and structure determination by means of molecular replacement is consequently not possible, phases will have to be obtained experimentally in the future.

We would like to thank Achim Dickmanns for helpful discussions and the staff of the EMBL BW7B beamline at the DORIS storage ring, DESY, Hamburg for support during data collection.

References

Bates, P. A., Kelley, L. A., MacCallum, R. M. & Sternberg, M. J. E. (2001). *Proteins*, **45**, *Suppl.* 5, 39–46.
Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S. & Jones, D. T. (2005). *Nucleic Acids Res.* **33**, W36–W38.
Burge, C. B., Tuschl, T. H. & Sharp, P. A. (1999). *The RNA World*, 2nd ed., edited by R. F. Gesteland, T. R. Cech & J. F. Atkins, pp. 525–560. New York: Cold Spring Harbor Laboratory Press.
Company, M., Arenas, J. & Abelson, J. (1991). *Nature (London)*, **349**, 487–493.
Gorbalenya, A. E. & Koonin, E. V. (1993). *Curr. Opin. Struct. Biol.* **3**, 419–429.
Hall, M. C. & Matson, S. W. (1999). *Mol. Microbiol.* **34**, 867–877.
Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.
Madhani, H. D. & Guthrie, C. (1994). *Annu. Rev. Genet.* **28**, 1–26.
Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M. & Lührmann, R. (2002). *Science*, **298**, 2205–2208.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mayas, R. M., Maita, H. & Staley, J. P. (2006). *Nature Struct. Mol. Biol.* **13**, 482–490.
Ohno, M. & Shimura, Y. (1996). *Genes Dev.* **10**, 997–1007.
Ono, Y., Ohno, M. & Shimura, Y. (1994). *Mol. Cell. Biol.* **14**, 7611–7620.
Schneider, S., Campodonico, E. & Schwer, B. (2004). *J. Biol. Chem.* **279**, 8617–8626.
Schneider, S., Hotz, H. R. & Schwer, B. (2002). *J. Biol. Chem.* **277**, 15452–15458.
Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5857–5864.
Schwer, B. & Gross, C. H. (1998). *EMBO J.* **17**, 2086–2094.
Schwer, B. & Meszaros, T. (2000). *EMBO J.* **19**, 6582–6591.
Staley, J. P. & Guthrie, C. (1998). *Cell*, **92**, 315–326.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Tanaka, N. & Schwer, B. (2005). *Biochemistry*, **44**, 9795–9803.
Tanner, N. K. & Linder, P. (2001). *Mol. Cell*, **8**, 251–262.
Wagner, J. D., Jankowsky, E., Company, M., Pyle, A. M. & Abelson, J. N. (1998). *EMBO J.* **17**, 2926–2937.
Wahl, M. C., Will, C. L. & Lührmann, R. (2009). *Cell*, **136**, 701–718.
Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
Will, C. L. & Lührmann, R. (2006). *The RNA World*, 3rd ed., edited by R. F. Gesteland, T. R. Cech & J. F. Atkins, pp. 369–400. New York: Cold Spring Harbor Laboratory Press.