Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Overproduction, purification, crystallization and preliminary X-ray diffraction studies of the human spliceosomal protein U5-15kD

Klaus Reuter and Ralf Ficner*

Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Emil-Mannkopff-Strasse 2, 35037 Marburg, Germany

Correspondence e-mail: ficner@imt.uni-marburg.de

Received 22 October 1998
Accepted 30 November 1998

The gene coding for the human spliceosomal U5 snRNP-specific 15 kDa protein (U5-15kD) was overexpressed in *Escherichia coli*, its product purified to homogeneity and crystallized. Well diffracting single crystals were obtained by the vapour-diffusion method in hanging drops and subsequent macroseeding. The crystals belong to the orthorhombic space group $P2_12_12$ with a = 62.3, b = 65.7, c = 37.1 Å. They diffract to at least 3.0 Å and contain one molecule in the asymmetric unit. A selenomethionine derivative of the protein was prepared and crystallized for multiwavelength anomalous diffraction (MAD) data collection.

1. Introduction

Catalysis of nuclear pre-mRNA splicing requires a large assembly of ribonucleoproteins known as the spliceosome. This macromolecular complex is a highly dynamic structure composed of four small nuclear ribonucleoprotein (snRNP) particles, U1, U2, U4/U6 and U5, and additional auxiliary factors (reviewed in Lamm & Lamond, 1993; Krämer, 1996). The spliceosomal snRNPs contain either one (U1, U2, U5) or two (U4/U6) snRNAs and more than 45 distinct proteins in total. Among the snRNP proteins there are the eight core (Sm) proteins B/B′, D1, D2, D3, E, F and G, which are common to all snRNPs, as well as proteins which are specific for a given snRNP species (reviewed in Will & Lührmann, 1997). In the case of the human U5 snRNP there are nine specific proteins with apparent molecular weights of 15, 40, 52, 100, 102, 110, 116, 200 and 205 kDa.

While catalysis of pre-mRNA splicing appears to be RNA based, the many conformational rearrangements of snRNAs that occur during the splicing process appear to be driven by spliceosomal proteins. Several RNA helicases belonging to the DEAD/H box superfamily (Lauber *et al.*, 1996; Laggerbauer *et al.*, 1998) as well as a cyclophilin with peptidyl–prolyl *cis*/*trans*-isomerase activity (Teigelkamp *et al.*, 1998) have been identified in the spliceosome. Sequence analysis of the gene coding for the U5-116kD protein revealed high homology to the ribosomal elongation factor EF-2, which hydrolyses GTP and acts as a translocase during translation (Fabrizio *et al.*, 1997). It is intriguing to speculate that U5-116kD perfomes a similar function in the spliceosome.

There is also structural information available on spliceosomal proteins. Much has been learned about protein–RNA interactions through the structure determination of several spliceosomal proteins complexed to snRNAs. The crystal and NMR structure of the U1A RNA-binding domain in complex with a hairpin within U1 snRNA (Oubridge *et al.*, 1994; Howe *et al.*, 1994) and the NMR structure of the same domain complexed to the polyadenylation-inhibition element of its own mRNA (Allain *et al.*, 1997) have been determined. The crystal structure of the U2B″–U2A′ complex bound to a fragment of U2 snRNA has only recently been published (Price *et al.*, 1998).

There is, however, still a considerable number of snRNP proteins which have no information available about their functions and interactions with snRNAs or other proteins. Among them is the U5 snRNP-specific 15 kDa protein (U5-15kD) recently isolated by Nottrott & Lührmann (personal communication). A point mutation in this protein was shown to have a temperature-dependent effect upon mitosis in *Schizosaccharomyces pombe*. Complete deletion of the U5-15kD gene produced a lethal phenotype, which could be complemented by its mammalian homolog (Berry & Gould, 1997). Because of the scarcity of detailed functional data on U5-15kD, we decided to initiate a structural study of this spliceosomal protein. Determination of its three-dimensional structure might reveal valuable insights into its function which are not revealed by its primary structure (Nottrott & Lührmann, personal communication). We describe here the details of the overexpression, purification and crystallization of

Table 1
X-ray data-collection statistics.

Number of crystals	1
Resolution (Å)	100–3.0
Wavelength (Å)	1.54
Space group	$P2_12_12$
Unit-cell parameters (Å)	
a	36.6
b	61.4
c	65.2
Data-collection temperature (K)	100
Number of observed reflections	29213
Number of unique reflections	2924
Completeness of all data (%)	88.2
R_{sym} for all data (%)	3.7
Completeness of outer shell (3.11–3.0 Å) (%)	91.9
R_{sym} in outer shell (%)	6.6

human U5-15kD as well as preliminary crystallographic data.

2. Methods, results and discussion

2.1. Cloning procedures, expression and purification

The coding region of the U5-15kD cDNA was PCR-amplified and cloned into the temperature-induceable overexpression vector pXC35 (Cheng & Patterson, 1992) *via* an *Nde*I and a *Hin*dIII restriction site introduced by the 5′ and 3′ PCR primers, respectively. The insert of the resulting plasmid p15KXC35 was completely resequenced from both directions to exclude mutations that might have occurred during PCR amplification.

Escherichia coli TAP106 transformed with p15KXC35 were grown in 0.5 l of LB medium (Miller, 1972) containing 100 mg l^{-1} ampicillin and 50 mg l^{-1} kanamycin at 303 K. At an OD_{600} of 1.0, the expression of U5-15kD was induced by addition of 0.5 l LB medium prewarmed to 336 K and subsequent growth of bacteria at 315 K for 4 h. Induction led to strong production of U5-15kD protein. After SDS–PAGE of a total cell lysate this protein was clearly visible as an intense band with an apparent molecular mass of 15 kDa, which is somewhat lower than the calculated molecular mass of 16.8 kDa (Nottrott & Lührmann, personal communication). For purification, cells were harvested by centrifugation and the cell pellet was resuspended in 10 ml of buffer *A* (100 m*M* Tris–HCl pH 7.8, 2 m*M* DTT) plus protease-inhibitor cocktail (one tablet Complete per 50 ml buffer; Boehringer Mannheim, Germany) plus 10 µg ml^{-1} bovine DNAaseI (Boehringer Mannheim, Germany). Cells were disrupted at 277 K by addition of hen egg-white lysozyme (Boehringer Mannheim, Germany) to a final concentration of 1 mg ml^{-1} and by sonication (15 cycles of 20 s on, 45 s off). 20 min after sonication, the cell extract was subjected to centrifugation at 30000*g* and 277 K for 1 h. The supernatant was further submitted to a second centrifugation at 100000*g* for 2 h. The S-100 was then loaded onto a Q-Sepharose (Fast Flow) column with a bed volume of 100 ml, which was equilibrated with lysis buffer. The column was run with a linear gradient of 0–0.4 *M* NaCl (five bed volumes) in the same buffer with a flow rate of 4 ml min^{-1}. The U5-15kD protein eluted at 0.2 *M* NaCl. $(NH_4)_2SO_4$ was added to the U5-15kD-containing fractions to a concentration of 1 *M*. This solution was loaded onto a Hi-Load phenyl Sepharose column (diameter 16 mm and height 10 cm). After being washed with two bed volumes of lysis buffer plus 1 *M* $(NH_4)_2SO_4$, the protein was eluted with a linear gradient of 1–0 *M* $(NH_4)_2SO_4$ (20 bed volumes) with a flow rate of 3 ml min^{-1}. U5-15kD eluted >95% pure at 0.2 *M* $(NH_4)_2SO_4$. To remove residual trace amounts of impurities, the U5-15kD-containing fractions were concentrated to 4 ml using a Centriprep-10 concentrator (Amicon, USA) and loaded onto a Superdex75 gel-filtration column (diameter 26 mm and height 60 cm). The column was eluted with buffer *A* (20 m*M* Tris–HCl pH 7.8, 120 m*M* NaCl, 2 m*M* DTT, 1 m*M* EDTA) at a flow rate of 2.5 ml min^{-1}. Pure U5-15kD eluted at 225 ml, proving that U5-15kD exists as a monomer, as assessed by comparison with gel-filtration runs of standard proteins. The pure protein was concentrated to 10 mg ml^{-1} using Centriprep-10 and Centricon-10 concentrators (Amicon, USA) for crystallization experiments. The chromatographic steps were carried out at room temperature using an Äkta-Explorer system (Pharmacia, Sweden). The amino-terminus of purified U5-15kD was sequenced. Apart from lacking the N-terminal methionine, the 14 N-terminal amino acids were identical to those predicted by the cDNA sequence.

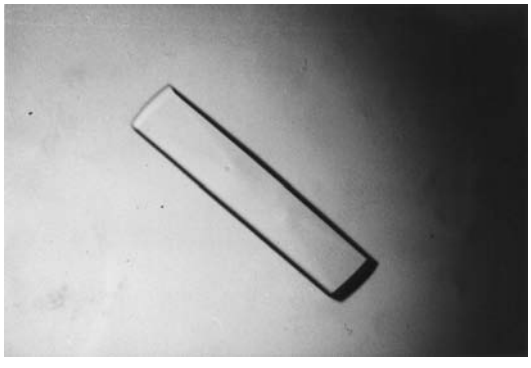

Figure 1
Crystal of U5-15kD obtained by macroseeding

2.2. Crystallization

Crystallization experiments were performed at 294 K in Linbro plates using the hanging-drop vapour-diffusion technique. A drop of 1.5 µl protein solution (100 mg protein per millilitre of buffer *A*) was mixed with an equal volume of reservoir solution and sealed against 1 ml reservoir solution. All reservoir solutions contained 2 m*M* DTT and 0.02%(*w*/*v*) sodium azide. In a systematic search for crystallization conditions, U5-15kD spontaneously formed numerous microcrystals in the presence of 1 *M* $(NH_4)_2SO_4$ as the precipitating agent buffered with 200 m*M* HEPES/NaOH in the pH interval 7.2–7.8. The first crystals appeared after four weeks at pH 7.6. No crystals were obtained at an $(NH_4)_2SO_4$ concentration of 0.8 or 1.2 *M*, or at 1 *M* $(NH_4)_2SO_4$ at 277 K. A U5-15kD derivative containing an N-terminal His-tag, which was kindly provided by Nottrott & Lührmann, showed no tendency to crystallize under these or various other tested conditions. To obtain X-ray quality crystals, some of the microcrystals were transferred to fresh droplets prepared identically to those leading to the microcrystals. After one week, the resulting single crystals, which had the dimensions 0.2 × 0.1 × 0.05 mm, were further passed on to new droplets. Within a week, the crystals grew to dimensions of approximately 0.7 × 0.3 × 0.2 mm (Fig. 1).

2.3. X-ray diffraction experiments and crystal characterization

In order to collect data under cryo-conditions, crystals were flash-frozen in a solution containing 62% saturated $(NH_4)_2SO_4$ and 20% sucrose as a cryoprotectant buffered with 100 m*M* HEPES/NaOH pH 7.6. Data were collected on an R-AXIS IV image-plate detector mounted on an RU300 rotating-anode generator (Rigaku/MSC) operating at 50 kV and 100 mA. The crystal-to-detector distance was 150 mm. A complete data set of 1° frames with a 15 min exposure time was collected. Crystals diffracted to 3.0 Å (Table 1). Processing the data with the program *DENZO* (Otwinowski & Minor, 1997) revealed an orthorhombic crystal system with unit-cell parameters a = 36.6 Å, b = 61.4 Å, c = 65.2 Å. From systematic absences in specific reflections in the

diffraction pattern the space group was determined to $P2_12_12$. Assuming one molecule in the asymmetric unit, the Matthews coefficient V_M (Matthews, 1968) was 2.2 Å^3 Da^{-1}, corresponding to a solvent content of 44%.

2.4. Purification and crystallization of selenomethionine U5-15kD

For multiwavelength anomalous diffraction (MAD) experiments, a derivative of U5-15kD in which the eight internal methionine residues of the protein are replaced by selenomethionines will be required (Hendrickson *et al.*, 1990). Temperature-induced overexpression of the U5-15kD gene by means of p15KXC35 requires the *E. coli* host strain TAP106 (Cheng & Patterson, 1992), which carries a temperature-sensitive λcI857 repressor allele but is *met*$^+$. The production of selenomethionine proteins is facilitated by using a methionine auxotrophic strain, which is grown in minimum medium supplemented with selenomethionine during induction. To our knowledge, no such *E. coli* strain exists with a λcI857 repressor allele in its genome. Therefore, the insert of p15KXC35 was recloned into the vector pET-21a(+) (Novagen, Madison, WI; Studier *et al.*, 1990) resulting in plasmid p15KET21. This plasmid allowed IPTG-induceable overexpression of the U5-15kD gene in the *met*$^-$ *E. coli* strain B834(DE3) (Novagen, Madison, WI).

A 1 l culture of B834(DE3)(pLysS/p15KET21) was grown at 310 K in minimum medium M9 (Sambrook *et al.*, 1989) with glucose (4 g l^{-1}) as the carbon source. The medium contained 100 mg l^{-1} ampicillin and 10 mg l^{-1} chloramphenicol. It was supplemented with biotin and thiamine (2 mg l^{-1} each) and L-methionine (50 mg l^{-1}). After the culture had reached an OD_{600} of 0.8, cells were spun down by centrifugation and resuspended in 1 l of the above medium devoid of methionine. After 1 h, 50 mg of DL-selenomethionine were added. After an additional 30 min, induction of selenomethionine U5-15kD expression was achieved through addition of IPTG to a final concentration of 1 m*M*. Expression was allowed to continue for 4 h. Purification of the selenomethionine U5-15kD was identical to that of the native protein, except that oxidation of selenomethionine was prevented by adding 5 m*M* DTT and 1 m*M* EDTA to all buffers. The success of selenomethionine incorporation was verified by electrospray mass spectrometry. A difference in molecular weight of 379.5 Da between the natural and selenomethionine protein was measured, corresponding to full incorporation of eight Se atoms (data not shown). To crystallize selenomethionine U5-15kD, natural U5-15kD crystals were used as a source for streak microseeding experiments (see Stura & Wilson, 1992). However, selenomethionine U5-15kD formed amorphous precipitates at 1 *M* $(NH_4)_2SO_4$ at pH 7.6. At a concentration of 0.8 *M* $(NH_4)_2SO_4$, crystals were obtained which were morphologically similar to the natural U5-15kD crystals. A decrease in solubility of selenomethionine proteins compared with their natural counterparts, which necessitates a reduced precipitant concentration in crystallization experiments, is commonly observed (Doublié, 1997). The resulting crystals were used for a second microseeding procedure to exclude non-selenomethionine protein from the crystals. Through subsequent macroseeding, selenomethionine U5-15kD crystals were obtained which were morphologically identical and grew to the same size as their natural counterparts. Selenomethionine U5-15kD crystals diffracted equally well and had the same unit-cell parameters as crystals of natural U5-15kD.

We thank Elke Penka for excellent technical assistance. We also thank Stephanie Nottrott and Reinhard Lührmann for communicating results prior to publication and for providing the human U5-15kD cDNA. Overexpression vector pXC35 was a generous gift from Tom A. Patterson. We greatly acknowledge the help of Thorsten Selmer, who performed the mass spectrometry of U5-15kD and selenomethionine U5-15kD at the MPI für terrestrische Mikrobiologie, Marburg, and of Gerhard Klebe (Institut für Pharmazeutische Chemie, Universität Marburg), whose X-ray diffractometer we used. We also thank D. Linder for N-terminal sequencing of the recombinant U5-15kD protein. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 286, TP A11).

References

Allain, F. H.-T., Howe, P. W. A., Neuhaus, D. & Varani, G. (1997). *EMBO J.* **16**, 5764–5772.
Berry, L. D. & Gould, K. L. (1997). *J. Cell Biol.* **137**, 1337–1354.
Cheng, X. & Patterson, T. A. (1992). *Nucleic Acids Res.* **20**, 4591–4598.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Fabrizio, P., Laggerbauer, B., Lauber, J., Lane, W. S. & Lührmann, R. (1997). *EMBO J.* **16**, 4092–4106.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Howe, P. W. A., Nagai, K., Neuhaus, D. & Varani, G. (1994). *EMBO J.* **13**, 3873–3881.
Krämer, A. (1996). *Annu. Rev. Biochem.* **65**, 367–410.
Laggerbauer, B., Achsel, T. & Lührmann, R. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4188–4192.
Lamm, G. M. & Lamond, A. I. (1993). *Biochim. Biophys. Acta*, **1173**, 247–265.
Lauber, J., Fabrizio, P., Teigelkamp, S., Lane, W. S., Hartmann, E. & Lührmann, R. (1996). *EMBO J.* **15**, 4001–4015.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miller, J. H. (1972). *Experiments in Molecular Genetics.* New York: Cold Spring Harbor Laboratory Press.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994). *Nature (London)*, **372**, 432–438.
Price, S. R., Evans, P. R. & Nagai, K. (1998). *Nature (London)*, **394**, 645–650.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd ed. New York: Cold Spring Harbor Laboratory Press.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
Stura, E. A. & Wilson, I. A. (1992). *Crystallization of Nucleic Acids and Proteins: a Practical Approach*, edited by A. Ducruix & R. Giegé, pp. 99–126. Oxford University Press.
Teigelkamp, S., Achsel, T., Mundt, C., Göthel, S. F., Cronshagen, U., Lane, W. S., Marahiel, M. & Lührmann, R. (1998). *RNA*, **4**, 1–15.
Will, C. L. & Lührmann, R. (1997). *Curr. Opin. Cell Biol.* **9**, 320–328.