Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary diffraction studies of the RNA polymerase subunit RPB5 from *Saccharomyces cerevisiae*

Meredith Hodach,[a] Flavia Todone,[b] Jyrki J. Eloranta,[a] Silvia Onesti[b] and Robert O. J. Weinzierl[a]*

[a]Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7, England, and [b]Blackett Laboratory, Imperial College of Science, Technology and Medicine, London SW7, England

Correspondence e-mail: r.weinzierl@ic.ac.uk

Received 10 March 1999
Accepted 26 May 1999

Crystals of the RNA polymerase subunit RPB5 from *Saccharomyces cerevisiae* have been obtained by vapour-diffusion techniques. The protein has been overexpressed in bacterial cells as a fusion with glutathione S-transferase. Two monoclinic crystal forms can be grown under different sets of conditions. In both cases, the diffraction is consistent with space group $P2_1$, with unit-cell parameters a = 45.3, b = 135.3, c = 47.3 Å, β = 118.6° for crystal form I and a = 48.4, b = 137.1, c = 47.1 Å, β = 118.6° for crystal form II.

1. Introduction

Transcription of DNA into RNA is carried out by a very complex and tightly regulated molecular apparatus. A variety of general and specific transcription factors regulate gene expression, allowing cells to adapt, differentiate and respond to various environmental stimuli. The enzymes at the core of these transcription processes are the RNA polymerases (RNAPs). In order to obtain detailed insights into the mechanisms of gene expression operating on the molecular level, a comprehensive knowledge of the structure and function of RNAPs is essential.

Although substantial progress has been made in the identification and characterization of various transcription factors, comparatively little is known about RNAPs in general (Thuriaux & Sentenac, 1993; Young, 1991). The three different RNAPs present in eukaryotic nuclei (RNAPI, RNAPII and RNAPIII) are complex enzymes, each consisting of 12–18 distinct subunits. Of these, the largest subunits are homologous to the β and β' subunits of bacterial RNAPs and are thought to contain the catalytic centre (Allison *et al.*, 1989; Sweeter *et al.*, 1987; Zaychikov *et al.*, 1996). The majority of the other subunits are relatively small (ranging in size from 8 to 40 kDa) and do not show any extensive sequence homologies to proteins of known function. Five of these 'small' subunits (RPB5, RPB6, RPB8, RPB10 and RPB12) are likely to play a central role in RNA transcription, as they are present in all three eukaryotic RNAPs (Woychik *et al.*, 1990). The overall importance of the small subunits is underscored by the fact that their primary sequences are highly conserved across the eukaryotic/archaeal evolutionary domains and by the observation that several human RNAP subunits can substitute for their yeast counterparts in genetic complementation studies (Shpakovski *et al.*, 1995). Furthermore, recent studies have shown that the quaternary arrangement of at least some of the RNAP subunits is very similar in the archaeal and eukaryotic enzymes, suggesting that many of the fundamental structural motifs have remained essentially invariant during the last 2–3 billion years of evolution (Eloranta *et al.*, 1998).

In order to learn more about RNAP architecture and to gain further insights into the structural motifs present in multi-subunit RNA polymerases, we have started to investigate the three-dimensional structures of the communal eukaryotic RNAP subunits. A high-resolution NMR study of RPB8 (Krapp *et al.*, 1998) revealed a structure with unexpected pseudo-twofold symmetry and the presence of a potential nucleic acid binding region encoded by a variant OB domain. RPB8 is thought to be present as a single copy in eukaryotic RNAPs, which prompts speculation as to whether it might participate in the assembly of a symmetrical protein–nucleic acid subcomplex. One of the other communal subunits, RPB5, is known to be present as two copies in RNAPII and has been shown to homodimerize under *in vitro* conditions (Lin *et al.*, 1997). These observations, together with indications of a direct interaction between RPB5 and RPB8 (Miyao *et al.*, 1996; Ishiguro *et al.*, 1998), point to a possible role of RPB5/RPB8 in establishing a symmetrical structure within RNAPII, which may also encompass the large subunits (Fu *et al.*, 1998). Other intriguing observations, including the fact that the transcriptional transactivator 'protein X' encoded by the human hepatitis virus may specifically interact with RPB5 to stimulate transcription in virally infected cells (Cheong *et al.*, 1995; Lin *et al.*, 1997), provide further evidence for the importance of RPB5 in eukaryotic RNAP structure and function.

The RPB5 subunit of *S. cerevisiae* has a molecular weight of 25.1 kDa and consists of two domains: a eukaryote-specific N-terminal 16 kDa domain and a smaller C-terminal domain (9 kDa), which is also found in archaea. We have used recombinant techni-

ques to express yeast RPB5 in quantities sufficient for structural studies.

2. Protein expression and purification

Recombinant full-length RPB5 was overexpressed as a GST fusion protein in *Escherichia coli* BL21 cells. The complete open reading frame of RPB5 was amplified by PCR from *S. cerevisiae* genomic DNA (Promega) in the presence of two oligonucleotide primers which are homologous to the N- and C-terminal ends of the published sequence of RPB5 (Woychik *et al.*, 1990) and are flanked by *Bam*HI linker sequences. The resulting PCR product was cloned into pGEM-T (Promega) and sequenced to eliminate the possibility of PCR-induced mutations and artefacts. The complete RPB5 open reading frame was then recovered as a single *Bam*HI fragment and ligated into the *Bam*HI site of the pGEX2TK expression vector (Pharmacia). This vector encodes a thrombin cleavage site upstream of the polylinker sequence, which allows the recombinant RPB5 protein to be cleaved from the GST fusion partner. After transfection into competent BL21 cells, production of recombinant GST–RPB5 fusion protein by several independent transformants was monitored by SDS–PAGE and a representative clone containing the RPB5 insert in the 'sense' orientation was selected for large-scale recombinant protein purification.

In a typical preparative-scale experiment, 2 l of bacteria harbouring the expression plasmid were grown in phosphate-buffered tryptone/yeast-extract medium in the presence of 0.2%(*w*/*v*) glucose and 100 µg ml^{-1} ampicillin at 310 K. At midlog phase ($A_{600} \simeq 0.6$–0.8), the incubation temperature was lowered to 289 K and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M*. The lower temperature considerably increased the yield of soluble fusion protein and reduced the extent of inclusion-body formation. After induction for 4 h, the cells were harvested by centrifugation and stored at 193 K. For the extract preparation, the cell pellets were thawed on ice and resuspended in 40–50 ml P300 buffer (300 m*M* potassium acetate, 20 m*M* Tris acetate buffer pH 7.9, 7 m*M* magnesium acetate, 10 m*M* DTT, 1 m*M* PMSF) containing lysozyme (at a concentration of 1 mg ml^{-1}) and the non-specific nuclease 'benzonase' (Merck; 1 unit ml^{-1}) to aid the breakup of cellular material and to reduce viscosity during the subsequent chromatography steps. After extensive sonication of the cell suspension, the soluble protein fraction was recovered by high-speed centrifugation (10 000*g*, 30 min at 277 K) and immediately applied to a chromatography column containing 10 ml glutathione agarose beads. Unbound protein was washed off with P300 buffer before the specifically bound fusion protein was eluted with buffer containing 5 m*M* reduced glutathione. The fractions containing fusion protein were combined and digested with thrombin (Sigma; 1 unit ml^{-1}) at 277 K overnight. After concentrating the sample by ammonium sulfate precipitation (80% saturation), RPB5 was purified from the GST fusion partner by size-exclusion chromatography on a Sephacryl-100 column (Pharmacia) equilibrated with 'crystallization buffer' (100 m*M* NaCl; 20 m*M* Tris–HCl pH 7.9; 10 m*M* β-mercaptoethanol; 10% glycerol). After assessing the purity of RBP5 in the various fractions by gel electrophoresis, the fractions containing electrophoretically pure RPB5 protein were concentrated with a Centricon-10 microconcentrator to a final protein concentration of 10–20 mg ml^{-1} and stored on ice prior to the crystallization trials.

3. Crystallization and preliminary diffraction studies

Crystals were grown by vapour diffusion in hanging drops at 277 K. Crystals suitable for diffraction studies were obtained under two different sets of conditions (I and II). In the first case (crystal form I), the drop was equilibrated against a well solution containing 20–24%(*w*/*v*) PEG 8K, 0.2 *M* ammonium sulfate, 0.1 *M* sodium cacodylate pH 6.5. In the second case (crystal form II), the well solution consisted of 30% PEG 4K, 0.2 *M* ammonium acetate, 0.1 *M* sodium citrate buffer at pH 5.6.

Both sets of conditions produced crystals with a comparable maximum size, having a cross-section of about 150 mm, a length of about 300 mm and similar diffraction properties. The diffraction patterns are both consistent with the space group $P2_1$, with unit-cell parameters $a = 45.3$, $b = 135.3$, $c = 47.3$ Å, $\beta = 118.6^\circ$ for crystal form I and $a = 48.4$, $b = 137.1$, $c = 47.1$ Å, $\beta = 118.6^\circ$ for crystal form II. The first set of conditions were chosen to routinely grow crystals for the structure determination. The crystals are very sensitive to changes in temperature. More stable crystals can be obtained by adding 15–20% glycerol to the well solution, with the additional advantage that crystals grown in the presence of glycerol can also be used directly for data collection at cryogenic temperatures.

A few images have been collected using a CCD detector at the SRS Daresbury and showed reflections up to 2.3 Å resolution. A complete data set to 3.0 Å resolution was collected on our in-house rotating-anode X-ray generator equipped with a 30 cm MAR Research imaging-plate system.

The calculation of the possible V_M (Matthews, 1968) is compatible with the presence of either one or two molecules in the asymmetric unit, with solvent contents of 77 and 53%, respectively. Biochemical data indicate that RPB5 is present as a dimer in the RNAP core complex (Kolodziej *et al.*, 1990). A self-rotation function calculation performed using the program *POLARRFN* (Collaborative Computational Project, Number 4, 1994) suggests the presence of a non-crystallographic twofold axis parallel to the crystallographic *a* direction. Heavy-atom screening is in progress.

We would like to thank Dr Peter Brick for critical reading of the manuscript. ROJW would like to acknowledge financial support from the Medical Research Council (Project Grant G9517777).

References

Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. (1989). *Cell*, **42**, 599–610.
Cheong, J. H., Yi, M., Lin, Y. & Murakami, S. (1995). *EMBO J.* **14**, 143–150.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Eloranta, J. J., Kato, A., Teng, M. S. & Weinzierl, R. O. J. (1998). *Nucleic Acids Res.* **26**, 5562–5567.
Fu, J., Gerstein, M., David, P. R., Gnatt, A. L., Bushnell, D. A., Edwards, A. M. & Kornberg, R. D. (1998). *J. Mol. Biol.* **280**, 317–322.
Ishiguro, A., Kimura, M., Yasui, K., Iwata, A., Ueda, S. & Ishihama, A. (1998). *J. Mol. Biol.* **279**, 703–712.
Kolodziej, P. A., Woychik, N., Liao, S. M. & Young, R. A. (1990). *Mol. Cell Biol.* **10**, 1915–1920.
Krapp, S., Kelly, G., Reischl, J., Weinzierl, R. O. J. & Matthews, S. (1998). *Nature Struct. Biol.* **5**, 110–114.
Lin, Y., Nomura, T., Cheong, J., Dorjsuren, D., Iida, K. & Murakami, S. (1997). *J. Biol. Chem.* **272**, 7132–7139.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miyao, T., Yasui, K., Sakurai, H., Yamagishi, M. & Ishihama, A. (1996). *Genes Cells*, **1**, 843–854.
Shpakovski, G. V., Acker, J., Wintzerith, M., Lacroix, J. F., Thuriaux, P. & Vigneron, M. (1995). *Mol. Cell. Biol.* **15**, 4702–4710.
Sweeter, D., Nonet, M. & Young, R. A. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 1192–1196.
Thuriaux, P. & Sentenac, A. (1993). In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, edited by E. W. Jones, J. R. Pringle & J. R. Broach. New York: Cold Spring Harbor Press.
Woychik, N. A., Liao, S. M., Kolodziej, P. A. & Young, R. A. (1990). *Genes Dev.* **4**, 313–323.
Young, R. A. (1991). *Annu. Rev. Biochem.* **60**, 689–715.
Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Goldfarb, A. & Mustaev, A. (1996). *Science*, **273**, 107–109.