Structures of Complete RNA Polymerase II and Its Subcomplex, Rpb4/7*

Karim-Jean Armache, Simone Mitterweger, Anton Meinhart‡, and Patrick Cramer§
From the Gene Center, University of Munich (LMU), Department of Chemistry and Biochemistry, Feodor-Lynen-Strasse 25, 81377 Munich, Germany

We determined the x-ray structure of the RNA polymerase (Pol) II subcomplex Rpb4/7 at 2.3 Å resolution, combined it with a previous structure of the 10-subunit polymerase core, and refined an atomic model of the complete 12-subunit Pol II at 3.8-Å resolution. Comparison of the complete Pol II structure with structures of the Pol II core and free Rpb4/7 shows that the core-Rpb4/7 interaction goes along with formation of an α-helix in the linker region of the largest Pol II subunit and with folding of the conserved Rpb7 tip loop. Details of the core-Rpb4/7 interface explain facilitated Rpb4/7 dissociation in a temperature-sensitive Pol II mutant and specific assembly of Pol I with its Rpb4/7 counterpart, A43/14. The refined atomic model of Pol II serves as the new reference structure for analysis of the transcription mechanism and enables structure solution of complexes of the complete enzyme with additional factors and nucleic acids by molecular replacement.

Synthesis of mRNA in eukaryotic nuclei is carried out by RNA polymerase (Pol) II. The mechanism of Pol II transcription has been investigated in many biochemical and detailed structural studies (for recent reviews, see Refs. 1–4). Pol II consists of a 10-polypeptide catalytic core and the heterodimeric Rpb4/7 complex, which can dissociate from the yeast enzyme (5). Refined structures were reported for the Pol II core in free form (6, 7), in form of a minimal elongation yeast enzyme (5). Refined structures were reported for the Polerodimeric Rpb4/7 complex, which can dissociate from the structural studies (for recent reviews, see Refs. 1–4). Pol II transcription has been investigated in many biochemical and detailed structural studies of scarce complexes of the complete Pol II, such as a recently described complex with DNA, RNA, and the elongation factor TFIIS (16).}

EXPERIMENTAL PROCEDURES

Rpb4/7 Preparation and Crystallization—Rpb4 and Rpb7 were co-expressed in Escherichia coli and purified as described previously (12). The Rpb4/7 variant lacking 34 amino-terminal residues of Rpb4 and with loop residues 82–100 deleted was amplified by PCR from the full-length Rpb4/7 construct and subcloned into vector pET21b (Novagen). The Rpb4/7 variant was purified essentially as described above and concentrated to 3 mg/ml. Crystals were grown at 20 °C with the hanging drop method, using a reservoir solution of 0.2 M di-ammonium hydrogen phosphate, pH 7.9, 20–22% polyethylene glycol 3350, and 5 mM dithiotreitol. Crystals grew to a maximum size of 0.3 × 0.2 × 0.1 mm. Crystals were transferred stepwise to reservoir solutions additionally containing 5%, 10%, 18%, and 22% glycerol and plunged in liquid nitrogen.

X-ray Structure Determination of Rpb4/7—A multiwavelength anomalous diffraction experiment was performed with selenomethionine-containing crystals at the Swiss Light Source (Table I). Data were processed with DENZO and SCALEPACK (17). The program SOLVE (18) detected 14 selenium sites and was used for multiwavelength anomalous diffraction phasing (Z-score = 42.8; figure of merit = 0.37). A subsequent anomalous difference Fourier identified four additional peaks. After phasing with SHARP (19), six additional selenium peaks were detected, giving a total of 24 selenium peaks. The peaks stemmed from four copies of the Rpb4/7 complex in the asymmetric unit and could be assigned to methionines Met1, Met24, Met106, and Met115 in Rpb7 and Met145 in Rpb4 (double peak). Phasing with all selenium peaks resulted in a figure of merit of 0.40 and in an electron density map that enabled building of an initial model with program O (20) that was partially refined with crystallography NMR software (21). The native crystal belonged to the same space group as the selenomethionine crystals (P21212), but the crystallographic c-axis is half the length (Table I), resulting in two Rpb4/7 complexes per asymmetric unit instead of four complexes. Thus, a dimer of two Rpb4/7 complexes was used as a search model in AMORE (22) to solve the native structure by molecular replacement (Table I, correlation coefficient = 42.7, next...
highest peak = 36.9). An atomic Rpb4/7 model was refined, which shows 99% of the residues in allowed and additionally allowed regions of the Ramachandran plot, and none of the residues in disallowed regions (Table I).

X-ray Structure Analysis of the Complete Pol II—The complete 12-subunit Pol II was reconstituted as described previously (12, 14). Complete diffraction data were collected in 0.5° increments at the Swiss Light Source (Table I). Molecular replacement with AMORE and the 10-subunit Pol II core (Protein Data Bank code 1I6H) as search model resulted in a unique solution (correlation coefficient = 52.4, second best solution = 30.5). The Pol II structure was adjusted with program O and refined with crystallography NMR software (Table I). Only a few rounds of restrained rigid body, positional, and B-factor refinement were carried out, with strict monitoring of the free R-factor. The refined structure has 98.5% of the residues in allowed and additionally allowed regions of the Ramachandran plot.

FIG. 1. Structure of the Rpb4/7 complex. A, primary and secondary structure. Structure-based alignments of amino acid sequences of Saccharomyces cerevisiae Rpb7 with Methanococcus jannaschii RpoE (top) and of S. cerevisiae Rpb4 with M. jannaschii RpoF (bottom). Secondary structure elements are shown above and below the respective sequences (cylinders, α-helices; arrows, β-strands; lines, loops; dashed lines, disordered; gray shading, absent from the crystallized variant) and labeled according to Ref. 15. Conserved residues are highlighted according to decreasing conservation from dark green, through light green and orange, to yellow. Cleavage sites revealed by limited proteolysis are indicated with arrows. Residues that are within 4 Å of residues of the Pol II core in the complete polymerase structure are indicated with a blue dot. This figure was prepared with ALSCRIPT (24). B, comparison of the structures of yeast Rpb4/7 (this study; left) and archaeal RpoF/E (Ref. 15; right). Rpb7/RpoE are in blue, and Rpb4/RpoF are in red. Disordered in the Rpb4/7 structure are the Rpb7 tip loop (residues 57–68) and Rpb4 residues 35–46, 77–81, and 101–118. Rpb4 residues 1–34 and 82–100 were not present in the crystallized variant. These figures were prepared with MOLSCRIPT (25) and PYMOL (DeLano Scientific).
RESULTS AND DISCUSSION

Structure Determinations—The complete 12-subunit yeast Pol II was prepared as described previously (12) and crystallized under altered conditions (16). We could extend the resolution of the crystals from 4.2 to 3.8 Å (Table I), but no further, likely because of the high solvent content of around 80%. An initial electron density map was phased with the core Pol II structure and improved by solvent flipping. The high quality of the map allowed us to adjust the structure of the polymerase core (7, 8) and to build parts of the Rpb4/7 complex as an atomic model.

To complete the model, however, we had to solve the crystal structure of free Rpb4/7 at high resolution (see “Experimental Procedures”). Full-length recombinant Rpb4/7 failed to crystallize, but crystals were obtained after 34 residues were removed from the proteolytically sensitive amino-terminal tail of Rpb4 (Fig. 1). Additional deletion of a proteolytically sensitive loop in Rpb4 resulted in larger crystals that diffracted synchrotron radiation to 2-Å resolution. The Rpb4/7 structure was solved by multiwavelength anomalous diffraction from a selenomethionine-substituted crystal and refined to a free R-factor of 27.4% at 2.3-Å resolution (Table I; Figs. 1 and 2). Due to the high resolution, a total of 154 water molecules could be added to the refined model of free Rpb4/7.

\[
\text{FIG. 2. Electron density maps. } 2F_\text{o} - F_\text{c} \text{ electron density maps around the final model of Rpb4 residues 176–196 are shown for free Rpb4/7 at 2.3-Å resolution (left) and for the complete Pol II at 3.8-Å resolution (right). The maps are contoured at 1σ. The figure was prepared with BOBSCRIPT (26).}
\]

\[
\text{FIG. 3. Complete RNA polymerase II structure. Ribbon diagram. Two standard views (front and top) are shown (6, 7). The 12 subunits Rpb1-Rpb12 are colored according to the key below the views. Dashed lines represent disordered loops. Eight zinc ions and the active site magnesium ion are depicted as cyan spheres and a pink sphere, respectively. Secondary structure assignments for the Pol II core were taken from Ref. 7.}
\]

\[
\text{TABLE I}
\]

| Crystal | Rpb4/7 SeMet | Rpb4/7 native | Complete Pol II |
|---------|--------------|---------------|----------------|
| Space group | P2₁,2,2 | P2₁,2,2 | C222, |
| Wavelength (Å) | 0.97946 peak | 0.93900 remote | 0.97973 inflection |
| Unit cell axis (Å) | 103.39, 114.69, 161.42 | 103.39, 114.69, 161.42 | 103.65, 114.81, 80.48 |
| Resolution (Å) | 20–2.6 (2.69–2.60) | 20–2.6 (2.69–2.60) | 20–2.6 (2.69–2.60) |
| Completeness (%) | 99.6 (99.5) | 99.7 (99.6) | 99.7 (99.6) |
| Unique reflections | 59,077 (5,811) | 59,144 (5,816) | 59,129 (5,815) |
| Redundancy | 4.5 | 4.5 | 4.5 |
| R<sub>sym</sub> (%) | 7.9 (21.6) | 7.4 (29.4) | 8.3 (40.3) |
| f<sub>pos</sub> | 15.0 | 13.3 | 11.7 |
| f<sub>pos</sub> | −7.8 | −2.28 | −9.6 |
| Refinement | 3.5 | 2.5 |
| Residues | 583 | 3902 | 81 |
| Zn<sup>2+</sup>/Mg<sup>2+</sup> | −/− | −/− | −/− |
| Water molecules | 154 | 0.007 | 0.009 |
| r.m.s.d. bonds (Å) | 1.36 | 1.53 | 22.8 |
| r.m.s.d. angles (°) | 22.8 | 25.7 | 27.4 |
| R<sub>cryst</sub> (%) | 22.8 | 25.7 | 28.5 |

\[
\text{a All data were collected at beamline X06SA at the Swiss Light Source (Villigen, Switzerland).}
\]

\[
\text{b r.m.s.d., root mean square deviation.}
\]
The refined Rpb4/7 structure served as a guide for completion of the Rpb4/7 model within Pol II. After removal of the water molecules, restricted refinement of atomic positions and B-factors, minor adjustments of the complete Pol II model led to a structure with high stereochemical quality (Table I; Fig. 3). Two factors apparently enabled refinement of an atomic Pol II model with data to 3.8-Å resolution. First, combination of two refined structures, the core Pol II structure and the Rpb4/7 structure, resulted in a very good starting model. Second, the high solvent content allowed for powerful electron density modification and resulted in a favorable ratio of observed reflections to refinement parameters. The 3.8-Å electron density map has high quality, and side chains are generally visible (Fig. 2).

Rpb4/7 Structure—The overall structure of the Rpb4/7 complex is similar to that of its archaeal counterpart (15). Rpb4 binds between two Rpb7 domains, the amino-terminal “tip” domain and a carboxyl-terminal domain (Fig. 1B). Compared with the archaeal protein, Rpb7 lacks the carboxyl-terminal helix K4 and has an additional helix inserted into strand A3 (Fig. 1, K*). About half of the Rpb4 residues (residues 48–57 and 138–221) fold as in the archaeal counterpart. Rpb4 additionally contains a non-conserved amino-terminal extension (residues 1–46), a longer helix H1, and an insertion between helices H1 and H2 (H1-H2 insertion), which comprises a long disordered loop and an additional helix (Fig. 2, H*). The Rpb4 amino-terminal extension is flexible in free Rpb4/7 but forms a partially ordered loop in the Pol II structure. This Rpb4 amino-terminal extension is apparently associated weakly with Rpb2 in the complete Pol II structure (Fig. 3). However, residues involved in Rpb2-Rpb4 contacts are not conserved, and the 34 amino-terminal residues of Rpb4 are not required for binding to the Pol II core in vitro (data not shown). Three amino-terminal residues of Rpb4 associate with the sheet of the Rpb7 tip, similar to the amino-terminal half of strand A1 in the archaeal counterpart (Figs. 2 and 3).

Folding Transitions upon Pol II Core-Rpb4/7 Interaction—In addition to the Rpb4 amino-terminal extension, other protein elements fold upon Rpb4/7-core interaction. The Rpb7 tip domain binds to the Pol II core with the helical turn K1 and with the highly conserved “tip loop,” which is disordered in the free Rpb4/7 structure but folds upon core binding (Fig. 4). There are no obvious changes in the conformation of the H1-H2 insertion upon core binding. Rpb4/7 binding, however, induces formation of an α-helix in the previously disordered residues 1445–1455 in the linker region of the largest Pol II subunit (Fig. 4, Rpb1 linker helix α50). Folding transitions upon interaction with Pol II are also observed in the elongation factor TFIIIS (14) and the initiation factor TFIIIB (11) and appear to be a common theme in interactions of the Pol II core with accessory factors.

Specificity of the Pol II Core-Rpb4/7 Interaction—The complete Pol II structure reveals details of the Pol II core-Rpb4/7 interaction, which refine our previous description of the involved protein elements (12). For example, the structure reveals a hydrogen bond between residue Gly696 of Rpb7 and residue Gln100 of Rpb6 in the Pol II core. This hydrogen bond explains why mutation of residue Gln100 in Rpb6 leads to facilitated Rpb4/7 dissociation and reduced Pol II activity upon a shift to non-permissive temperatures (23). The detailed information on the core-Rpb4/7 interaction also helps to rationalize why another nuclear RNA polymerase, Pol I, binds specifically to A43/14, the Rpb4/7 counterpart. Specificity of A43/14 for the Pol I core may result from an ionic contact between a lysine that replaces the Rpb7 residue Pro63 in A43 and a patch of three glutamates that replace the Rpb1 residues Thr13, Met1433, and Gly1439 in the largest Pol I subunit. However, we could not find a straightforward explanation for specific interaction of the core of the third nuclear RNA polymerase, Pol III, with its Rpb4/7 counterpart, C25/17. It is possible that additional Pol III-specific subunits are involved in defining the specificity in subunit assembly.

Acknowledgments—We thank C. Schulze-Briese and the beamline X06SA of the Swiss Light Source for help. We thank H. Kettenberger and T. Kamenski (Gene Center Munich) for help. We thank members of the Cramer laboratory and K. Strässer for comments on the manuscript. Diffraction data were collected at the Swiss Light Source, Paul Scherrer Institut (Villigen, Switzerland).

REFERENCES

1. Cramer, P. (2004) Adv. Protein Chem. 67, 1–42
2. Hahn, S. (2004) Nat. Struct. Mol. Biol. 11, 394–403
3. Asturias, F. J. (2004) Curr. Opin. Struct. Biol. 14, 121–129
4. Woychik, N. A., and Hampsey, M. (2002) Cell 108, 453–463
5. Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991) J. Biol. Chem. 266, 71–75
6. Cramer, P., Bushnell, D. A., Fu, J., Gnaat, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. H., and Kornberg, R. D. (2000) Science 288, 640–649
7. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Science 292, 1863–1876
8. Gnaat, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) Science 292, 1876–1881
9. Bushnell, D. A., Cramer, P., and Kornberg, R. D. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 1218–1222
10. Westover, K. D., Bushnell, D. A., and Kornberg, R. D. (2004) Science 303, 1014–1016
11. Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004) Science 303, 983–988
12. Armache, K.-J., Kettenberger, H., and Cramer, P. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 6964–6968
13. Bushnell, D. A., and Kornberg, R. D. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 6969–6972
14. Kettenberger, H., Armache, K.-J., and Cramer, P. (2003) Cell 114, 347–357
15. Todone, F., Brick, P., Werner, F., Weinzierl, R. O., and Onesti, S. (2001) Mol. Cell 8, 1137–1143
16. Kettenberger, H., Armache, K.-J., and Cramer, P. (2003) Mol. Cell 16, 955–965
17. Gwinnowski, Z., and Minor, W. (1996) Methods Enzymol. 276, 307–326
18. Terwilliger, T. C. (2002) Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 1937–1940
19. de La Fortelle, E., and Bricegad, G. (1997) Methods Enzymol. 276, 472–494
20. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
21. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nicholls, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
22. Navaza, J. (1994) Acta Crystallogr. Sect. A 50, 157–163
23. Tan, Q., Prysak, M. H., and Woychik, N. A. (2003) Mol. Cell. Biol. 23, 3329–3338
24. Barton, G. J. (1993) Protein Eng. 6, 37–49
25. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
26. Ennor, R. M. (1997) J. Mol. Graph. 15, 132–134