Structure of an Archaeal RNA Polymerase

Anselm G. Kusser1†, Michela G. Bertero1†, Souad Naji2
Thomas Becker1, Michael Thomm2, Roland Beckmann1
and Patrick Cramer1* 

1Gene Center Munich and Center for integrated Protein Science CiPSM
Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25
81377 Munich, Germany
2Institut für Biochemie Genetik und Mikrobiologie Universität Regensburg
Universitätsstr. 31
93053 Regensburg, Germany

Received 30 June 2007; received in revised form
24 August 2007; accepted 28 August 2007
Available online
5 September 2007

Edited by R. Huber

Unravelling the mechanisms of gene transcription and its regulation requires detailed structural information for the key enzymes of transcription, the multisubunit RNA polymerases (RNAPs). Bacteria and archaea contain a single RNA polymerase (RNAP), and eukaryotes contain three RNA polymerases: RNAP I, RNAP II, and RNAP III. Crystalllographic structures are available for bacterial RNAP and the eukaryotic RNAP II, which reveal a conserved core architecture and active centre cleft. However, no structural information is available for archaeal RNAPs, except for the small polymerase subcomplex F/E, the counterpart of the RNAP II subcomplex Rpb4/7. Sequence analysis suggested that archaeal RNAPs are related to eukaryotic RNAP II more closely than they are to bacterial RNAPs. Archaeal homologues are known for all RNA polymerase II subunits except Rpb8 and Rpb9 (Table 1). The similarity between the archaeal and eukaryotic RNAP II transcription machineries extends to the archaeal initiation factors TBP, TFB, and TFE, which are structural and functional counterparts of the eukaryotic factors TBP, TFIIB, and the large subunit of TFIIE, respectively. 

To tackle the archaeal RNAP structure, we established a large-scale purification protocol for the endogenous polymerase from Pyrococcus furiosus (P.fu) based on a published method (Figure 1(a)). The improved protocol enabled purification of 5 mg of RNAP from 60 g of P.fu cells. Pure RNAP comprised all subunits in an apparently stoichiometric

*Corresponding author. E-mail address: cramer@LMB.uni-muenchen.de.
† A.G.K. and M.G.B. contributed equally to this work.

Present address: M. G. Bertero, Centre for Genomic Regulation CRG, Systems Biology Programme,
Dr. Aiguader 88, 08003 Barcelona, Spain.

Abbreviations used: RNAP, RNA polymerase; Pfu, Pyrococcus furiosus; EM, electron microscopy; TBP, TATA box-binding protein; TF, transcription factor; FSC, Fourier shell correlation.

© 2007 Published by Elsevier Ltd.
**Table 1. P.fu RNAP subunits in comparison to RNAP II**

<table>
<thead>
<tr>
<th>Polymerase part</th>
<th>P.fu subunit</th>
<th>(M) (kDa)</th>
<th>Corresponding Pol II subunit</th>
<th>Sequence identity* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>A’</td>
<td>103.1</td>
<td>Rpb1 N-term. Part</td>
<td>42.2</td>
</tr>
<tr>
<td></td>
<td>A”</td>
<td>44.4</td>
<td>Rpb1 C-term. Part</td>
<td>37.0</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>127.0</td>
<td>Rpb2</td>
<td>43.1</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>29.8</td>
<td>Rpb3</td>
<td>26.8</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>9.2</td>
<td>Rpb5</td>
<td>40.2</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>6.2</td>
<td>Rpb6</td>
<td>42.1</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>7.8</td>
<td>Rpb9</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>5.8</td>
<td>Rpb10</td>
<td>52.3</td>
</tr>
<tr>
<td></td>
<td>E’</td>
<td>21.7</td>
<td>Rpb7</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>14.1</td>
<td>Rpb4</td>
<td>5.0</td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>380.2</td>
<td>–</td>
<td>38.2</td>
</tr>
</tbody>
</table>

* Number of amino acid residues in the P.fu RNAP subunit that are identical in the corresponding RNAP II subunit divided by the total number of residues in the P.fu RNAP subunit.

**Figure 1.** Cryo-EM structure of P.fu RNAP. (a) SDS-PAGE analysis of purified endogenous P.fu RNAP. P.fu cells were thawed by resuspending 60 g in 50 mM Tris–HCl (pH 8.0), 100 mM KCl, 10 mM MgCl₂, 20% v/v glycerol, 10 mM β-mercaptoethanol, 1% Protease Inhibitor Cocktail (0.28 µg Leupeptin, 1.37 µg pepstatin A, 0.17 mg Phenylmethylsulfonylfluoride, 0.33 mg benzamide per 1 ml of 50% ethanol) at 4 °C over night, while stirring. An Emulsiflex (Avestin Emulsiflex C5) was used for cell disruption, applying 2 pressure cycles at 1500 bar (1 bar = 10⁵ Pa). To ensure complete cell lysis, a small aliquot was checked by phase-contrast microscopy. The lysate was centrifuged (30 min at 11,532 g, Beckman Ti 45), recovering the supernatant after each step. The supernatant was loaded onto a column (GE Healthcare XK50) packed with (group II) that is observed for RNAP I. Each particle was aligned to both references, the two corresponding PDB file that corresponded to the initially used reference but contained the clamp domain in an entirely closed conformation.

For the first round of sorting, we used the current reconstructed volume (group I), and a volume obtained from a reference volumes for the next round of sorting. The process was repeated several times until convergence was reached for all the numbers of particles assigned to each group. This resulted in 22,240 particles with a closed clamp conformation, and 19,429 particles with other clamp conformations. The resolution for the group I volume was 16.8 Å based on a cut-off value of 0.5 for the Fourier shell correlation. The group I volume was further refined to a resolution of 13 Å (cf. (e)). (f) Gallery of regions of the P.fu RNAP structure deviating from yeast RNAP II (cf. Figure 2). Depicted domains in RNAP II are highlighted using the original RNAP II colour code.
Figure 1 (legend on previous page)
manner (Figure 1(a)). RNAP preparations were monodisperse and catalytically active (data not shown), and contained single particles according to electron microscopy (EM) with negative staining, and enabled collection of high-quality cryo-EM data (Figure 1(b)). The cryo-EM reconstruction with 22,240 conformations produced a map at 13 Å resolution (Figure 1(d)–(f)).

The detailed EM map for Pfu RNAP enabled a unique fit of the crystal structure of the complete 12 subunit Saccharomyces cerevisiae RNAP II (Figure 1(d)). Comparison of the EM map with the RNAP II structure confirmed the overall conservation of the enzyme architecture and active centre, including the polymerase bridge helix, the pore, and the clamp. Outside the active centre, at least one functional surface is conserved. The dock domain and its surrounding regions are highly similar in Pfu RNAP and RNAP II, reflecting a conserved interaction with the initiation factor TFB/TFIIB.15,16 No density was present at the locations for the RNAP II subunits Rpb8 and Rpb9 (Figure 1(d) and (f)), consistent with a lack of archaeal homologues for these subunits.

To explain the observed differences between the EM map and the RNAP II structure, we constructed an RNAP II-based homology model for Pfu RNAP (Figure 2; Supplementary Data Figure S1). In the model, the regions with the highest level of sequence conservation cluster around the active centre cleft. Many peripheral regions are also conserved, and only several surface domains are divergent. In particular, the archaeal subunit A’, which is homologous to the N-terminal part of the largest RNAP II subunit Rpb1, contains a deletion in the clamp head, an extended pore domain, and it largely lacks the foot domain (Figures 1(f) and 2). In subunit A”, which is homologous to the C-terminal part of Rpb1, several loops in the jaw domain are shorter, and the C-terminal repeat domain is missing. The C terminus of A’ and the N terminus of A” form residual structure in the region of the Rpb1 foot (Figures 1(f) and 2). Subunit B contains an eight-residue insertion in the Rpb2 protrusion domain (Figures 1(f) and 2).

Among the small archaeal subunits, D and L resemble the RNAP counterparts Rpb3 and Rpb11, respectively, except that several surface loops are missing, including the zinc loop in Rpb3 (Figure 2). Subunit H lacks the N-terminal jaw domain present in its eukaryotic counterpart Rpb5 (Figure 1(d)). Subunit K lacks the N-terminal region of the counterpart Rpb6, and the last two β-strands in the C-terminal assembly domain of Rpb6 (Figure 1(f)). Consequently, subunit K closely resembles the bacterial RNAP subunit homolog ω.17 The hetero-dimeric subcomplex F/E’ is situated at the expected surface position that is occupied by its counterpart

![Figure 2](image-url)
Rpb4/7 in RNAP II (Figure 1(d)). However, EM density is largely lacking for the outermost domains of the subcomplex F/E, the OB and HRDC domains (Figure 1(d)), indicating their mobility. Indeed, the OB domain is the most flexible region of RNAP II according to normal mode analysis (not shown), and the HRDC domain is mobile in RNAP III.\textsuperscript{16,19}

In conclusion, the first structure of an archaean RNAP identified similarities to the related structure of eukaryotic RNAP II, but also revealed unique deviations. The main finding is that the archaean enzyme can largely be regarded as a truncated version of RNAP II. Pfur RNAP mainly lacks parts that are peripheral in RNAP II, including two small subunits and several surface domains and loops. Finally, our results provide the basis for structural studies of archaean RNAP complexes that should provide significant mechanistic insights into eukaryotic transcription as well.

Acknowledgements

We thank Anass Jawhari and other members of the Cramer laboratory for help. We thank Michael Forchheim for help. We thank Thorsten Mielke for help with EM data collection. P.C. was supported by the Deutsche Forschungsgemeinschaft, the Sonderforschungsbereich SFB646, the Center for integrated Protein Science CiPS\textsuperscript{CM}, and the Fonds der chemischen Industrie. R.B. was supported by the Deutsche Forschungsgemeinschaft and the Sonderforschungs bereiche SFB646 and SFB894. EM data were collected at the USN (UltraStructureNetwork) supported by the European Union and the Senatsverwaltung für Wissenschaft, Forschung und Kultur Berlin.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.08.066

References