Crystal structure of RNA polymerase II from *Komagataella pastoris*

Haruhiko Ehara [a, b, c], Takashi Umehara [b, c], Shun-ichi Sekine [a, b, c], Shigeyuki Yokoyama [a, b, d, *]

[a] *Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[b] *RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan*
[c] *RIKEN Center for Life Science Technologies, Division of Structural and Synthetic Biology, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan*
[d] *RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan*

ARTICLE INFO

ABSTRACT

RNA polymerase II (Pol II) is a 12-subunit protein complex that conducts the transcription of mRNA and some small RNAs. In this work, the crystal structure of Pol II from the methylotropic yeast *Komagataella pastoris* (*Pichia pastoris*) was determined. While the structure is highly homologous to that of Pol II from the budding yeast *Saccharomyces cerevisiae*, the stalk and clamp modules of the *K. pastoris* Pol II displayed large inward rotations, closing the central cleft to a greater extent than in the known *S. cerevisiae* Pol II structures. The conformational differences reflect the inherent flexibilities of the stalk and the clamp, as additional low-resolution structures of *K. pastoris* Pol II in different crystal forms revealed diverse stalk and clamp orientations. Comparisons with other eukaryotic/archaeal RNA polymerase structures in the Protein Data Bank revealed the distributions of the stalk and clamp orientations. The conformational plasticity should be essential for transcriptional functions and binding various regulatory factors.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

RNA polymerase II (Pol II) is a huge cellular factory that transcribes genes and synthesizes the precursors of mRNAs and small RNAs. Pol II is a protein complex consisting of 12 different subunits. Since the first crystal structure of Pol II from the budding yeast *Saccharomyces cerevisiae* [1], many structures of Pol II in complex with various nucleotides, nucleotide analogs, inhibitors, nucleic acids, and transcription factors have been reported [2–4]. Most of these high-resolution structural studies have been performed with *S. cerevisiae* Pol II, except for the crystal structure of Pol II from the fission yeast *Schizosaccharomyces pombe* [5] and the recent cryo-EM structures of mammalian Pol II [6–8]. The overall structures of Pol II are divided into five modules: the clamp, core, shelf, jaw-lobe, and stalk, and their relative positions are quite flexible [1,3]. The stalk and the clamp are especially mobile, and their dynamics are implicated in fundamental transcription functions and interactions with transcription factors [7–10]. However, since the

crystallographic studies of *S. cerevisiae* Pol II have been performed with only a few different crystal forms, the inherent ranges of motion of the modules have remained elusive.

We previously reported the tandem-affinity purification of Pol II from *Komagataella pastoris* (formerly called *Pichia pastoris*), a methylotrophic yeast well-known for high-density cultivation, and demonstrated its suitability for crystallization [11]. In this study, we determined the crystal structure of the tandem-affinity purified *K. pastoris* Pol II. In the *K. pastoris* Pol II structure, the stalk and the clamp are both extremely closed, indicating that the inherent ranges of motion of these two modules are much larger than those observed in previous studies.

2. Materials and methods

2.1. Protein preparation

TAP-tagged *K. pastoris* Pol II was prepared as previously described [11]. Unlike *S. cerevisiae* Pol II, *K. pastoris* Pol II existed as a 12 subunit complex containing the Rpb4/7 subunits, and fractions lacking the Rpb4/7 subunits were not detected during the chromatographic steps. The genomic fragment encoding *K. pastoris* TFIIEβ (a.a. 73–281) was cloned into the pET-47b vector (Novagen),

* Corresponding author. RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
E-mail address: yokoyama@riken.jp (S. Yokoyama).

and the protein was expressed in *E. coli* Rosetta 2 (DE3) by induction with 0.2 mM isopropyl-β-thiogalactopyranoside at 288K. The cells were resuspended in buffer A [50 mM Tris-HCl (pH 8.0), 700 mM NaCl, and 10 mM β-mercaptoethanol] and disrupted by sonication. The lysate was cleared by centrifugation and then loaded on a Ni-Sepharose FF column (GE Healthcare). The column was extensively washed with buffer A containing 30 mM imidazole, and then the on-column cleavage of the His-tag was performed with HRV-3C protease (Novagen) overnight. The protein was eluted with buffer A containing 30 mM imidazole, and then dialyzed against buffer B [50 mM Tris-HCl (pH 8.0), 100 mM potassium acetate, and 10 mM β-mercaptoethanol]. The protein was further purified by passage through a Resource Q column (GE Healthcare), and then dialyzed against buffer C [20 mM Tris-acetate (pH 8.0), 150 mM potassium acetate, 4 mM magnesium acetate, 2 μM zinc acetate, and 4 mM TCEP-HCl]. The protein was concentrated to 5.8 mg/ml, and stored at −20 °C.

2.2. Crystallization

Crystallization was performed by the sitting-drop vapor diffusion method. The type-1 crystals were obtained at 293K, with a reservoir solution containing 0.2 M sodium citrate, 8% PGA-LM (poly-γ-glutamic acid low molecular weight, Molecular Dimensions), and 10% glycerol. The type-2 crystals were grown at 285K with a reservoir solution containing 0.2 M sodium citrate, 8% PGA-LM, and 1% dextran sulfate sodium salt (Hampton Research). The type-3 crystals were obtained in the presence of TFIIEβ (73−281). When Pol II was mixed with TFIIEβ (73−281) at a molar ratio of 1:2, white precipitates were formed. The sample was cleared by centrifugation, and the supernatant was subjected to crystallization. The type-3 crystals were obtained at 293K with a reservoir solution containing 0.15 M sodium citrate, 10% PGA-LM, 0.5% dextran sulfate sodium salt, 15 mM Tris, and 35 mM Hepes.

2.3. Data collection

The type-1 crystals were cryoprotected with a solution containing 20 mM Tris-acetate (pH 8.0), 100 mM potassium acetate, 2 mM magnesium acetate, 200 mM sodium citrate, 8% PGA-LM, and 20% glycerol, and then flash-cooled in liquid nitrogen. For the type-2 crystals, a cryoprotectant containing 20 mM Tris-acetate (pH 8.0), 2 mM magnesium acetate, 100 mM sodium citrate, 12% PGA-LM, and 20% glycerol was used. The type-3 crystals were cryoprotected with a solution containing 20 mM Tris-acetate (pH 8.0), 15 mM Tris, 35 mM Hepes, 50 mM potassium acetate, 2 mM magnesium acetate, 150 mM sodium citrate, 0.5% dextran sulfate, 10% PGA-LM, and 20% glycerol. Diffraction data were collected at SPring-8 BL32XU (Harima, Japan) for the type-1 crystals, and at SPring-8 BL41XU for the type-2 and type-3 crystals.

2.4. Structure determination

Diffraction data were processed with HKL2000 for the type-1 crystal, and with XDS for the type-2 and type-3 crystals [12,13]. The initial model for the molecular replacement was created with phenix.sculptor [14], using the coordinates of the *S. cerevisiae* Pol II (PDB: 2VUM) [4]. After molecular replacement with phenix.automr, rigid body refinement of the Pol II structural modules was performed with phenix.refine for the type-1 and type-3 structures [15]. For the type-2 structure, the stalk was manually fitted to the electron density, and then the clashing residues and the regions with missing electron density were deleted from the structure, using Coot [16]. Further rigid body refinement and morph refinement were subsequently performed with phenix.refine and phenix.morph_models. Structural figures were created with PyMOL [17].

2.5. Structure comparison

Structural superimpositions were performed with the matchmaker script in UCSF Chimera [18,19], using the Rpb1 subunit excluding the clamp region (residues 466 to 1390) from the *S. cerevisiae* Pol II structure (PDB: 2VUM) as the reference. The centroids of the Cα positions of the stalk OB fold domain were used to represent the stalk positions. For the clamp coiled-coil, residues 281 to 287 in *S. cerevisiae* Pol II and their corresponding residues were used for the centroid calculation except in the Pol I structures with a longer coiled-coil domain, where residues 401, 402, 419, and 420 were employed instead. The calculation of the centroid positions was also performed with USCF Chimera.

3. Results

3.1. Crystallization and structure determination of K. pastoris pol II

We previously reported the purification and preliminary crystallization of Pol II from *K. pastoris* [11]. The protein was crystallized under conditions containing PGA-LM and sodium citrate at 293 K. In the present study, we optimized the crystallization conditions, and obtained three different crystal forms. Optimization of the reagent concentrations yielded plate-shaped single crystals (hereafter designated as the type-1 crystals; Supplementary Fig. S1). Bulky and differently shaped crystals (the type-2 crystals) appeared by performing the crystallization at a lower temperature (285 K) and/or in the presence of dextran sulfate as an additive. Furthermore, by adding a fragment of the TFIIEβ subunit to the type-2 conditions, small rectangular crystals (the type-3 crystals) were obtained.

X-ray diffraction data were collected for all three crystal types by using synchrotron radiation at SPring-8, Japan (Supplementary Table S1). Phases were determined by molecular replacement, using the coordinates of *S. cerevisiae* Pol II (PDB: 2VUM) as the search model. The type-1 and type-3 crystals were found to belong to the space group $P2_1$ with different unit cell constants, and diffracted X-rays to ~7 Å. For these crystal types, the asymmetric unit contained two Pol II molecules. The type-2 crystals belonged to the space group $P3_1$, and diffracted X-rays to ~4 Å. The molecular replacement revealed one Pol II molecule in the asymmetric unit. Given the higher resolution, we hereafter describe the *K. pastoris* Pol II structure based on the type-2 crystal, unless otherwise specified.

3.2. Highly-closed structure of K. pastoris pol II

K. pastoris Pol II exhibited the typical "crab-claw" shape characteristic of multisubunit RNA polymerases (Fig. 1A; PDB: 5X50). The structures of the individual subunits and domains are mostly identical between *K. pastoris* Pol II and *S. cerevisiae* Pol II, as expected from their high sequence conservation (see Supplementary Fig. S2 for the alignment). However, *K. pastoris* Pol II exhibited significant differences in the orientations of the Rpb4/7 stalk and the clamp, relative to the rest of the molecule, as compared with the reported structure of *S. cerevisiae* Pol II (Fig. 1B; PDB: 2VUM). The stalk of *K. pastoris* Pol II is tilted toward the Pol II body by approximately 20°, which moved the distal tip of the stalk by more than 20 Å (Fig. 1C). The clamp displayed an inward rotation of 5°, which caused its coiled-coil domain to be displaced by around 6 Å. This indicates that the stalk and the clamp of Pol II can adopt more closed orientations than those observed in the *S. cerevisiae* Pol II elongation complex.