NIH Public Access
Author Manuscript
Nature. Author manuscript; available in PMC 2010 January 13.

Published in final edited form as:
Nature. 2008 February 14; 451(7180): 851–854. doi:10.1038/nature06530.

The X-ray Crystal Structure of RNA Polymerase from *Archaea*

Akira Hirata[1], **Brianna J. Klein**[1], and **Katsuhiko S. Murakami**[1]
[1]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802

Abstract

The transcription apparatus in *Archaea* can be described as a simplified version of its eukaryotic RNA polymerase (RNAP) II counterpart, comprising a RNAPII-like enzyme as well as two general transcription factors, the TATA-binding protein (TBP) and the eukaryotic TFIIB ortholog TFB[1,2]. It has been widely understood that precise comparisons among cellular RNAP crystal structures could reveal structural elements common to all enzymes and that these insights would be useful to analyze components of each enzyme that enable it to perform domain-specific gene expression. However, the structure of archaeal RNAP has been limited to individual subunits[3,4]. Here, we report the first crystal structure of the archaeal RNAP from *Sulfolobus solfataricus* at 3.4 Å resolution, completing the suite of multi-subunit RNAP structures from all three domains of life. We also report the high resolution (at 1.76 Å) crystal structure of the D/L subcomplex of archaeal RNAP and provide the first experimental evidence of any RNAP possessing an iron-sulfur (Fe-S) cluster, which may play a structural role in a key subunit of RNAP assembly. The striking structural similarity between archaeal RNAP and eukaryotic RNAPII highlights the simpler archaeal RNAP as an ideal model system for dissecting the molecular basis of eukaryotic transcription.

RNAP is the central enzyme of gene expression, and all life forms have RNAPs that function as multi-subunit protein complexes (multi-subunit RNAP) with subunit compositions that vary depending on the domain of life[5]. The common core of the multi-subunit RNAPs comprises 5 subunits that are conserved from bacteria to humans. Bacterial RNAP is the simplest form of this family (composed of the minimum $\beta'\beta\alpha_I\alpha_{II}\omega$ subunits) while in *Eukarya*, RNAPII possesses additional polypeptides to form a 12 subunit complex, which is responsible for synthesizing all mRNAs in the cell. Previous crystal structures of bacterial core[6] and holoenzymes[7], as well as RNAPII[8–14], have provided insight into structural and functional aspects of RNAPs, but the archaeal RNAP structure remains to be elucidated. Based on subunit composition, sequence similarity and the requirement of general transcription factors for promoter recognition and transcription, archaeal and eukaryotic RNAPs have been proposed to be structurally similar. A crystal structure of the archaeal RNAP, however, has been needed in order to allow a precise comparison of the transcriptional machineries from all three domains of life, and generate new insights about structural domains and motifs of cellular RNAPs that participate in the assembly of these multi-component molecular machines and in transcription processes.

Author information Correspondence and requests for materials should be addressed to K.S.M. (kum14@psu.edu)..
Author contributions A.H. crystallized and solved the structures of the *S. solfataricus* RNAP and D/L subcomplex. B.J.K. supported the RNAP purification and its structure determination. K.S.M. and A.H. wrote the manuscript and all authors discussed the results and commented on the manuscript.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Coordinates and structure factors have been deposited to the Protein Data Bank (accession codes 2PMZ and 2PA8 for the *S. solfataricus* RNAP and D/L subcomplex structures, respectively). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We have purified native RNAP from *S. solfataricus* for crystallization and structure determination (see Methods). Because of the high sequence conservation among RNAPs from all species in *Archaea*[15], insights derived from the *S. solfataricus* RNAP can be generalized to the transcription apparatus from this entire domain. The overall structure of the RNAP resembles a "crab claw" with a protruding "stalk" formed by the E'/F subcomplex (Fig. 1a). The relative positioning of the RNAP core and stalk are highly conserved between archaeal RNAP and the three classes of eukaryotic RNAPs (Fig. 2)[16]. In the crystal structure of archaeal RNAP, the clamp is in its closed conformation (Supplementary Fig. 1), most likely due to the presence of the E'/F subcomplex. Structural studies of RNAPII have proposed that the function of the stalk (Rpb7/Rpb4 subcomplex) is to modulate the clamp conformation: in the absence of the subcomplex, the RNAPII clamp is in the open conformation, while in the presence of the subcomplex, the clamp is closed[8,11,12]. In archaeal RNAP, another function of the E'/F subcomplex has been proposed from *in vitro* archaeal RNAP reconstitution studies: it facilitates transcription bubble formation of the RNAP-promoter DNA complex under certain conditions[17–19].

The archaeal RNAP structure enables a direct comparison of RNAPs from *Bacteria*, *Archaea* and *Eukarya* (Fig. 2). Conserved structures are found around the vicinity of the active center and DNA binding channel (Supplementary Figs. 2 and 3) that reflect the fact that the fundamental transcription mechanism is conserved among cellular RNAPs. The structures of the surfaces vary between the bacterial and the eukaryotic enzymes (Supplementary Fig. 2b); in contrast, there is striking structural conservation between archaeal and eukaryotic enzymes distributed throughout the entire structure (Supplementary Fig. 2a). For the largest subunit, the entire architecture of Rpb1 (RNAPII) can be superimposed on the A'/A" subunits (archaeal RNAP) with small deviations around the foot domain (Figs. 3a and 3b, Supplementary Fig. 4b; structural similarity is 68.9 %, Supplementary Table 1). Counterparts for the Rpb1 jaw and clamp head domains are disordered in the archaeal RNAP crystal structure, making their structural similarities uncertain. In contrast, the structural similarity between Rpb1 and the bacterial β' subunit is only 50 % (Supplementary Table 1); there are substantial differences in the foot (Fig. 3c), pore1, cleft and dock domains, and there are no Rpb1 jaw and clamp head counterparts in the β' subunit (Supplementary Fig. 4c). In addition, *T. aquaticus* β' contains a lineage-specific insertion (*Taq* β' NCD, 283 residues) between conserved regions A and B[6,20].

In *Archaea*, the largest subunit is split into two polypeptides, A' and A", which are encoded by separate genes in an operon[1,2,15]. Sequence alignments reveal that archaeal A' and A" correspond to the N-terminal two-thirds and the C-terminal one-third of the RNAPII Rpb1 subunit, respectively, and that the junction between A' and A" is positioned at the Rpb1 "foot" domain[8] (Fig. 3a and Supplementary Fig. 4). The archaeal RNAP foot domain consists of four α helices (Supplementary Fig. 5a), composed of the C-terminal of A' (15 residues) and the N-terminal of A" (62 residues). The corresponding RNAPII foot has a more complex architecture formed by a continuous polypeptide including 8 α-helices and 2 anti-parallel β-strands (Fig. 3b and Supplementary Fig. 5b), but the same four α-helix architecture, which is found in the archaeal RNAP, is conserved in the center of the RNAPII foot domain. In contrast, the bacterial foot domain has a completely different architecture (Fig. 3c).

Compared to the largest subunit, the structural similarity between RNAPII Rpb2 and the archaeal B subunit is higher (90 %, Supplementary Fig. 6b, Supplementary Table 1). In contrast, the structural similarity between RNAPII and bacterial β subunit is only 64.9 %, and there are deviations in the external domains 1 and 2, the hybrid binding, and clamp domains as well as the flap loop/flap-tip helix region (Supplementary Fig. 6c).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Although the two largest conserved subunits provide most of the catalytic functions, their assembly is dependent on the presence of two other conserved subunits which form a platform for assembly (the D/L[15], Rpb3/Rpb118 and α_I/α_{II}[6,21,22] in the archaeal, eukaryotic, and bacterial enzymes, respectively, Fig. 2). Interestingly, phylogenetic analyses[23] and UV-visible spectra (Supplementary Fig. 7a) suggested that *S. solfataricus* RNAPs may contain an Fe-S cluster in the D subunit. In order to obtain direct evidence of an Fe-S cluster within RNAP as well as determine its chemical structure and chelation to the protein, we solved the crystal structure of the *S. solfataricus* D/L subcomplex at 1.76 Å resolution. The D subunit forms a heterodimer with the L subunit (D/L subcomplex), which has an architecture that is almost identical to the eukaryotic (Rpb3/Rpb11 heterodimer) and bacterial (α homodimer) counterparts (Supplementary Fig. 8). The D subunit consists of three domains: 4Fe-4S cluster binding domain, domain 2, and dimerization domain (Fig. 1b). The folding of domain 2 in the D subunit is highly conserved in Rpb3 and α; however, the 4 Cys residues in the D subunit form two disulfide bonds whereas the 4 Cys residues in Rpb3 chelate a Zn^{2+} ion. The α subunit has neither a disulfide bond nor a Zn^{2+} ion binding motif in domain 2. The 4Fe-4S cluster binding domain is unique to *S. solfataricus* RNAP since the corresponding region in Rpb3 forms a loosely packed domain (called the "loop"8) and there is no corresponding domain in the bacterial α subunit. In the 4Fe-4S cluster binding domain, three Cys residues (C183, C203 and C209) provide ligands for the 3Fe-4S cluster (Fig. 1c) while one additional C206 is positioned nearby suggesting that the cluster may exist in a 4Fe-4S *in vivo* (Supplementary Fig. 7c). Two other C176 and C213 in the 4Fe-4S cluster binding domain form a disulfide bond. Moreover, a strong Fe anomalous signal was detected in the RNAP crystal within the 4Fe-4S cluster binding domain (Supplementary Fig. 9b), which further verifies that the RNAP does indeed carry an Fe-S cluster. Although there are many instances where Fe-S clusters provide redox potential for an enzymatic reaction, the Fe-S cluster in the RNAP is located 45 Å from its catalytic center (Fig. 1a) and therefore unlikely to be involved in RNAP catalysis. To examine the role of the Fe-S cluster, we constructed a mutant D subunit where the four Cys residues that coordinate the cluster were all replaced with Ser. The four-Ser variant protein does not chelate the Fe-S cluster, and it forms aggregates even when the D and L subunits are co-expressed in *E. coli* (Fig. 4). We also obtained the same results from D subunit variants having C183 or C203 substituted with Ala (data not shown). Interestingly, the 4Fe-4S cluster binding domain is not directly involved in heterodimer formation between the D and L subunits (Fig. 1b), but the absence of an Fe-S cluster causes the D subunit to aggregate and prevent the functional D/L subcomplex from being formed. These observations suggest that the cluster may play a role in supporting the structural integrity of the D subunit. When the Fe-S cluster is present within the *S. solfataricus* D subunit, it is not oxygen-sensitive since we observed a complete 3Fe-4S cluster in the D/L subcomplex crystal structure, which was purified and crystallized under aerobic conditions. Although the Fe-S cluster can be removed from the D/L subcomplex using the Fe chelator 2, 2'-dipyridyl, it is stable within the assembled RNAP and protected from 2, 2'-dipyridyl treatment (data not shown).

S. solfataricus is not the only archaeon possessing the Fe-S cluster in RNAP (Supplementary Fig. 10). The 4Fe-4S cluster binding motif has been identified in the RNAP sequence of 16 out of 28 sequenced archaeal genomes including almost all methanogens (except *Methanocaldococcus jannaschii* and *Methanosarcina maripaludis*). In addition, the 4Fe-4S cluster binding motif is not limited to archaeal RNAPs. Twelve eukaryotic genomes, including those of plants (*Arabidopsis thaliana*) and protozoa (*Tetrahymena thermophila*) encode the 4Fe-4S cluster binding motif within AC40, the D subunit ortholog of RNAPI and RNAPIII. Furthermore, UV-visible absorption spectra as well as iron and acid-labile sulfide analyses have shown that *Arabidopsis thaliana* AC40 contains an Fe-S cluster (A. H. and K.S.M., unpublished data). Our studies show that certain cellular RNAPs are Fe-S proteins. This result raises other intriguing questions including why do only certain cellular RNAPs posses Fe-S cluster and what is the relationship between the presence of Fe-S cluster in the RNAPs and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

their living environments. The archaeal RNAP structure provides a framework for addressing the interesting question of what functional roles do Fe-S clusters play within the transcription machinery of *Archaea* and *Eukarya*.

The archaeal K subunit is an ortholog of eukaryotic Rpb6 and bacterial ω subunits, and their folds are highly conserved in the central regions. In archaeal RNAP, subunit K, along with the E' subunit tip loop, participates in protecting the C-terminal tail of the largest subunit (Fig. 3a), and the arrangement of these subunits is conserved in the RNAPII (Fig. 3b). The C-terminal tail (residues 78–95) of ω subunit, however, has a different trajectory compared to K and Rpb6 (Fig. 3c). The ω tail wraps around and completely covers the C-terminal end of β'. Intriguingly, this bacteria-specific protein fold may have the similar function of the corresponding tip loops of E' of archaeal RNAP and Rpb7 of RNAPII.

Precise structural comparisons between archaeal and eukaryotic RNAPs reveal structural elements common to both systems; these insights will be useful to elucidate components unique to RNAPII that may enable it to perform highly regulated gene expression[24]. The distinct structural differences between these RNAPs are limited to the side of the face (Fig. 2, Supplementary Fig. 11) that faces downstream DNA in the transcription elongation complex. From a structural perspective, it is interesting that almost all these differences can be classified as simple additions of RNAPII-specific polypeptides, including class III subunits (Rpb8 and Rpb9) and domains (Rpb1 CTD and Rpb5 jaw) to the archaeal RNAP, rather than changes to the core structure (Supplementary Figs. 11 and 12).

The key to understanding the mechanism of an enzymatic reaction is to have a simple and robust system that can be probed *in vitro*. Studies on bacterial transcription are the most advanced because the RNAP can be assembled from recombinant[25] subunits with full activity. This system allows one to use straightforward molecular biology techniques, as well as site-specific incorporation of chemical probes[26,27]. The latter has led to a number of successful biophysical and biochemical studies on RNAP. Currently, these tools are not applicable to eukaryotic RNAPII since this enzyme cannot be reconstituted *in vitro*. Furthermore, because of substantial structural differences in key regions including downstream DNA configuration and RNA exit channel[28], bacterial RNAP may not serve as a complete and accurate model system for the transcription in eukaryotes. Many biochemical studies, as well as this current crystallographic study, have established that the structural and functional similarities between archaeal RNAP and eukaryotic RNAPII exist on many levels. Furthermore, active archaeal RNAP can be conveniently reconstituted from its individual subunits *in vitro*[18,29 17], and studies of promoter-dependent transcription can easily be established with fully purified general transcription factors and *in vitro* reconstituted RNAP. The crystal structure presented here reveals that archaeal RNAP is not only a simplified version of its eukaryotic RNAPII counterpart but also an excellent model system for dissecting the molecular basis of eukaryotic transcription.

Methods Summary

S. solfataricus RNA polymerase purification, crystallization and structure determination

Native RNAP was purified from *S. solfataricus* P2 cells for crystallization and structure determination studies. Primitive monoclinic $P2_1$ crystal contained two 377 kDa RNAP molecules per asymmetric unit. The final R-work and R-free factors are 27.0 % and 34.3 %, respectively. Figures are prepared by Pymol.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Purification, crystallization and structure determination of the *S. solfataricus* subcomplex D/L

An electron density map at 1.76 Å resolution (Supplementary Fig. 9c) was obtained by a combination of molecular replacement by using the yeast Rpb3/Rpb11 structure[8] as a search model and density modification. The final R-work and R-free factors are 21.0 % and 24.7 %, respectively.

Methods

Sulfolobus solfataricus P2 RNA polymerase purification

Native RNAP was purified from *S. solfataricus* P2 cells for crystallization. Cell paste was mixed with TGEMD buffer (10 mM Tris-HCl (pH 8), 5% glycerol, 1 mM EDTA, 10 mM 2-Mercaptoethanol and 2 mM DTT) in the presence of a protease inhibitor cocktail and lysed by sonication. Ammonium sulfate was added to the soluble fraction of the cell lysate (30 % saturation), and the precipitated proteins were removed by centrifugation. The soluble fraction was added to a Phenyl-Sepharose High Performance resin (GE Healthcare) equilibrated with TGEDM + 30 % ammonium sulfate. The RNAP was eluted from the resin with TGEMD buffer and further purified by three consecutive chromatography columns including Heparin-Sepharose, Superdex-200, and Mono-Q columns (GE Healthcare). The RNAP was concentrated to 5 mg/mL in buffer containing 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM EDTA and 2 mM DTT for crystallization. According to the *S. solfataricus* sequence in the database, the B subunit is also split into two separate polypeptides, B' and B”. However, the RNAP we purified contains a single polypeptide which represents the B subunit. In order to resolve this discrepancy, we cloned and sequenced the DNA covering the B subunit and found differences between our sequence results and the genomic database. Our sequencing data is consistent with a single polypeptide B subunit. According to other genomic studies, B subunits from closely related *Sulfolobus* genera also encode a single B subunit polypeptide. Therefore, we believe that the DNA sequence coding for the B' and B” genes in the *S. solfataricus* database may be in error.

Crystallization and structure determination of the RNAP

RNAP was crystallized by sitting drops at 22 °C against a reservoir containing 0.1 M Hepes (pH 7.5), 0.1 M K_2CO_3, 0.1 M sodium thiosulfate, 12 % (w/v) polyethylene glycol (PEG) 10,000 and 2 mM spermine tetrahydrochloride. For cryocrystallography, the crystals were presoaked in stabilization solution (same as the crystallization solution except with 15 % PEG 10,000), and then transferred to a cryo-solution that contained stabilization solution plus 25 % glycerol, followed by flash-freezing by immersion in liquid nitrogen. A complete data set was collected at 3.4 Å resolution (Supplementary Table 3), and processed by the HKL2000 software package[31]. Primitive monoclinic $P2_1$ crystals (a=125.8 Å, b=201.2 Å, c=196.1 Å, β=100.9°) contained two 377 kDa RNAP molecules per asymmetric unit, with a solvent content of 61.9 %. Electron density maps were calculated using a combination of molecular replacement (Phaser32) and density modification (Supplementary Fig. 5a). The search model was derived from 10 subunits of the yeast RNAPII transcribing complex[14] (PDB code: 1R9T) with the following regions removed that were specific for RNAPII: subunits Rpb8 and Rpb9; domains Rpb1 foot (residues 871–1058), Rpb1 jaw (residues 1141–1274), Rpb5 jaw (residues 1–144). The molecular replacement solution used included two RNAP molecules per asymmetric unit. The electron density map calculated using phases from the molecular replacement was further improved using the density modification program Resolve[33]. The resulting electron density map has several deviations from the molecular replacement solution indicating that model bias was effectively removed by density modification. Rigid body refinements were performed with four mobile modules8 and further adjustments to the model were carried out manually. The resulting model phases allowed us to position a *M. jannaschii* subcomplex E'/F structure[3] in

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the electron density map. The Rpb3/Rpb11 in the search model was replaced with the *S. solfataricus* subcomplex D/L that we determined in this study and the position of the Fe-S cluster in the D subunit was confirmed by Fe-anomalous difference Fourier map of the RNAP crystal (Supplementary Fig. 9b). Next, we replaced the amino acid sequence of the model with the *S. solfataricus* RNAP. Positional refinement incorporating tight geometry and NCS restraints was carried out by programs CNS[34] and Refmac5[35] by carefully monitoring the R-free factor. The final R-work and R-free factors are 27.0 % and 34.3 %, respectively. The resulting map allowed segments that were not present in the search model to be built manually including the foot domain of A′/A”, solvent exposed loops of the B subunit, and an N-terminal α-helix of K subunit.

Purification, crystallization and structure determination of the *S. solfataricus* subcomplex D/L

The genes encoding *S. solfataricus* D and L subunits were cloned from genomic DNA by PCR, and the D/L co-expression vector was constructed using a pET21a (Novagen). D/L subunits were over-expressed in BL21-CodonPlus(DE3)-RIPL (Stratagene) in TB media supplemented with 100 μg/mL of ferric ammonium citrate. After the cells were lysed by sonication, most of the *E. coli* proteins were removed by heat treatment (65 °C for 30 min). The recombinant subcomplex D/L was further purified by Q-sepharose anion-exchange and Superdex-75 gel filtration column chromatography (GE Healthcare). The protein was concentrated to 10 mg/mL with buffer containing 10 mM Tris-HCl (pH 8), 50 mM NaCl, 1 mM EDTA and 2 mM DTT for crystallization. Crystals were prepared by vapor diffusion in hanging drops at 22 °C against reservoir containing 0.1 M sodium citrate (pH 5), 22 % saturated ammonium sulfate and 12 % (w/v) PEG 4000. Cryo-protection of the crystals was achieved by 20 % glycerol, and the crystals were flash-frozen in liquid nitrogen. A complete data set out to 1.76 Å resolution was collected at a wavelength of 0.97 Å at synchrotron beamline X25 at National Synchrotron Light Source, NY (Supplementary Table 4). I-centered orthorhombic $I2_12_12_1$ crystals (a=69.7 Å, b=93.3 Å, c=128.3 Å) contained one D/L subcomplex per asymmetric unit. An excellent electron density map at 1.76 Å resolution was obtained by a combination of molecular replacement by using the Rpb3/Rpb11 structure[8] as a search model and density modification33. The initial model was refined by simulated annealing, energy minimization and individual B-factor refinement by CNS34. The presence of a 3F-4S cluster in the D subunit was confirmed by an Fe anomalous Fourier map obtained from a data collected at wavelength of 1.2818 Å (Fig. 1c and Supplementary Figs. 9c and 9d, Supplementary Table 4).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank L. Berman and A. Héroux at the National Synchrotron Light Source, D. Lessner and H. Yennawar at The Pennsylvania State University, and D. Bushnell and R. Kornberg at Stanford University for help. We thank E. P. Geiduschek, J. G. Ferry, S. A. Darst, F. Asturias, V. Lamour and R. Yajima for critiques of the manuscript. This work was supported by The Pew Scholars Program in the Biomedical Sciences and supported in part by National Institutes of Health.

References

1. Bell SD, Jackson SP. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol 1998;6:222–8. [PubMed: 9675798]
2. Geiduschek EP, Ouhammouch M. Archaeal transcription and its regulators. Mol Microbiol 2005;56:1397–407. [PubMed: 15916593]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

3. Todone F, Brick P, Werner F, Weinzierl RO, Onesti S. Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. Mol Cell 2001;8:1137–43. [PubMed: 11741548]
4. Yee A, et al. Solution structure of the RNA polymerase subunit RPB5 from Methanobacterium thermoautotrophicum. Proc Natl Acad Sci U S A 2000;97:6311–5. [PubMed: 10841538]
5. Ebright RH. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J Mol Biol 2000;304:687–98. [PubMed: 11124018]
6. Zhang G, et al. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell 1999;98:811–24. [PubMed: 10499798]
7. Murakami KS, Masuda S, Darst SA. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. Science 2002;296:1280–4. [PubMed: 12016306]
8. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 2001;292:1863–76. [PubMed: 11313498]
9. Bushnell DA, Westover KD, Davis RE, Kornberg RD. Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science 2004;303:983–8. [PubMed: 14963322]
10. Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell 2006;127:941–54. [PubMed: 17129781]
11. Bushnell DA, Kornberg RD. Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. Proc Natl Acad Sci U S A 2003;100:6969–73. [PubMed: 12746498]
12. Armache KJ, Kettenberger H, Cramer P. Architecture of initiation-competent 12-subunit RNA polymerase II. Proc Natl Acad Sci U S A 2003;100:6964–8. [PubMed: 12746495]
13. Armache KJ, Mitterweger S, Meinhart A, Cramer P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J Biol Chem 2005;280:7131–4. [PubMed: 15591044]
14. Westover KD, Bushnell DA, Kornberg RD. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell 2004;119:481–9. [PubMed: 15537538]
15. Goede B, Naji S, von Kampen O, Ilg K, Thomm M. Protein-protein interactions in the archaeal transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors. J Biol Chem 2006;281:30581–92. [PubMed: 16885163]
16. Fernandez-Tornero C, et al. Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. Mol Cell 2007;25:813–23. [PubMed: 17386259]
17. Ouhammouch M, Werner F, Weinzierl RO, Geiduschek EP. A fully recombinant system for activator-dependent archaeal transcription. J Biol Chem 2004;279:51719–21. [PubMed: 15485836]
18. Naji S, Grunberg S, Thomm M. The RPB7 orthologue E' is required for transcriptional activity of a reconstituted archaeal core enzyme at low temperatures and stimulates open complex formation. J Biol Chem 2007;282:11047–57. [PubMed: 17311916]
19. Werner F, Weinzierl RO. Direct modulation of RNA polymerase core functions by basal transcription factors. Mol Cell Biol 2005;25:8344–55. [PubMed: 16135821]
20. Chlenov M, et al. Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta' subunit. J Mol Biol 2005;353:138–54. [PubMed: 16154587]
21. Ishihama, A.; Taketo, M.; Saitoh, T.; Fukuda, R. RNA polymerase. Losick, R.; Chamberlin, M., editors. Cold Spring Harbor Laboratory; 1976. p. 485-502.
22. Zillig, W.; Palm, P.; Heil, A. RNA polymerase. Losick, R.; Chamberlin, M., editors. Cold Spring Harbor Laboratory; 1976. p. 101-125.
23. Rodriguez-Monge L, Ouzounis CA, Kyrpides NC. A ferredoxin-like domain in RNA polymerase 30/40-kDa subunits. Trends Biochem Sci 1998;23:169–70. [PubMed: 9612079]
24. Thomas MC, Chiang CM. The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol 2006;41:105–78. [PubMed: 16858867]
25. Zalenskaya K, et al. Recombinant RNA polymerase: inducible overexpression, purification and assembly of Escherichia coli rpo gene products. Gene 1990;89:7–12. [PubMed: 2197183]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

26. Murakami K, et al. Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors. Proc Natl Acad Sci U S A 1997;94:11274–8. [PubMed: 9326599]
27. Kapanidis AN, et al. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science 2006;314:1144–7. [PubMed: 17110578]
28. Vassylyev DG, et al. Structural basis for substrate loading in bacterial RNA polymerase. Nature 2007;448:163–8. [PubMed: 17581591]
29. Werner F, Weinzierl RO. A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription. Mol Cell 2002;10:635–46. [PubMed: 12408830]
30. Campbell EA, et al. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell 2001;104:901–12. [PubMed: 11290327]
31. Otwinowski, Z.; Minor, W. Macromolecular Crystallography Part A. Charles, W.; Carter, J., editors. Academic Press; 1997. p. 307-326.
32. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr 2005;61:458–64. [PubMed: 15805601]
33. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr 2000;56(Pt 8):965–72. [PubMed: 10944333]
34. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54(Pt 5):905–21. [PubMed: 9757107]
35. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr D Biol Crystallogr 2001;57:122–33. [PubMed: 11134934]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

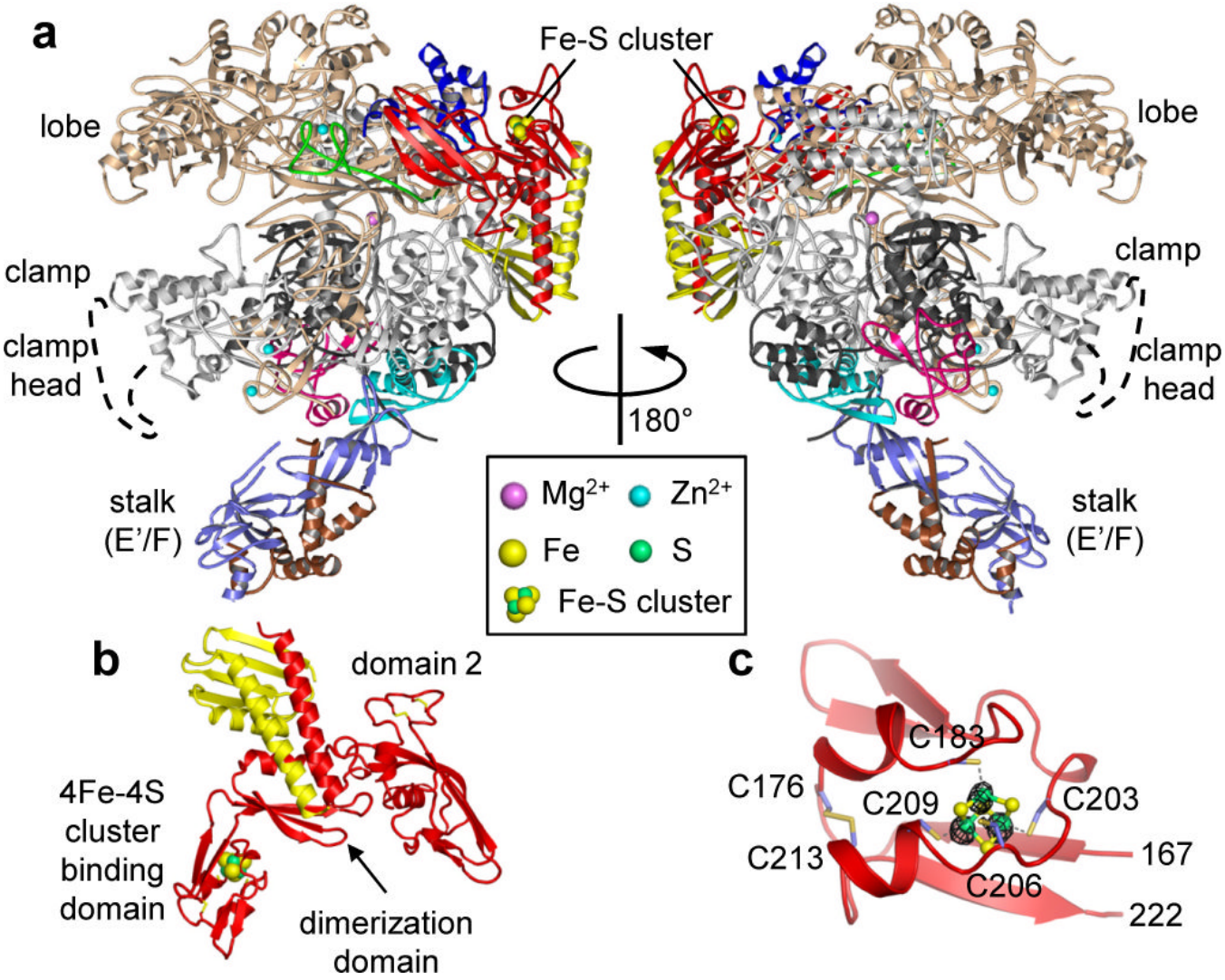

Figure 1.

Three-dimensional structure of the archaeal RNAP. a, 3.4 Å resolution crystal structure of the *S. solfataricus* RNAP. Each subunit is denoted by a unique color (see surface representations in Fig. 2 for color-code and subunit-subunit interaction). The disordered clamp head domain is indicated as a dotted line. b, 1.76 Å resolution crystal structure of the D/L subcomplex (red, D subunit; yellow, L subunit). Domain organization is shown. c, Close-up view of the 4Fe-4S cluster binding domain (residues 167–222 of the D subunit). Electron density calculated using the Fe anomalous signal is shown in black mesh (sigma cutoff =5).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

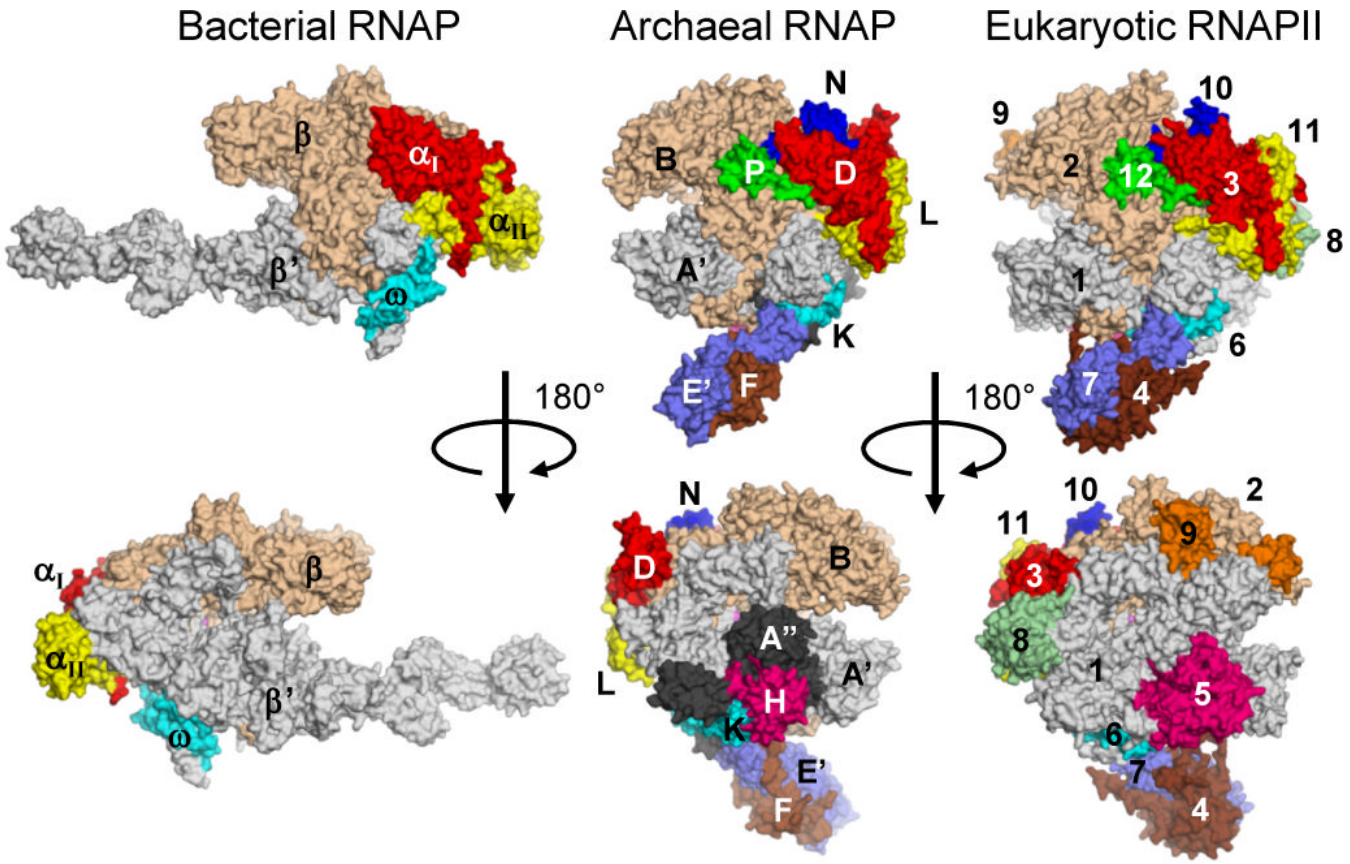

Figure 2.

Cellular RNAP structures from three domains of life. Surface representation of multi-subunit cellular RNAP structures from *Bacteria* (left, *Thermus aquaticus* core enzyme30), *Archaea* (center, *S. solfataricus*) and *Eukarya* (right, *S. cerevisiae* RNAPII13). Each subunit is denoted by a unique color and labeled. Orthologous subunits are depicted by the same color.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

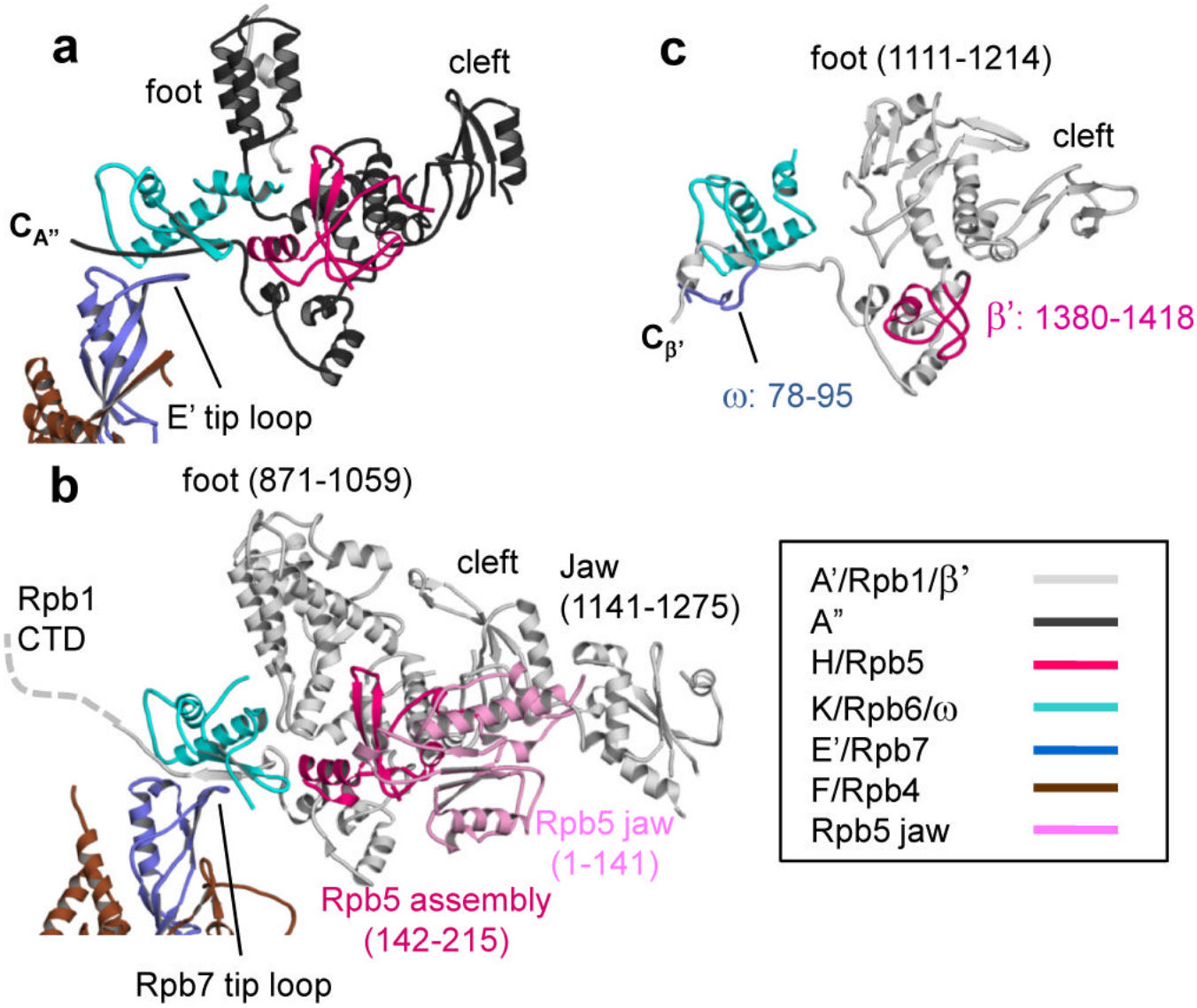

Figure 3.

Structures around the foot domains from three domains of life. a, Ribbon model of the archaeal RNAP (color coding of each subunits are indicated). Various domains and motifs are labeled. The junction between A' and A" is positioned at foot domain. Subunits H and K associate around the base of the foot. b, Ribbon model of the eukaryotic RNAPII[13]. The same four α-helix architecture, which is found in the archaeal RNAP, is conserved in the center of the RNAPII foot domain. Rpb5 and Rpb6 of also associate around the foot domain. c, Ribbon model of the bacterial RNAP30. The bacterial foot domain has a completely different architecture. The right side of the bacterial foot domain associates with the bacterial specific insertion of the β' subunit. In addition, the ω tail wraps around the C-terminus of the β' subunit.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

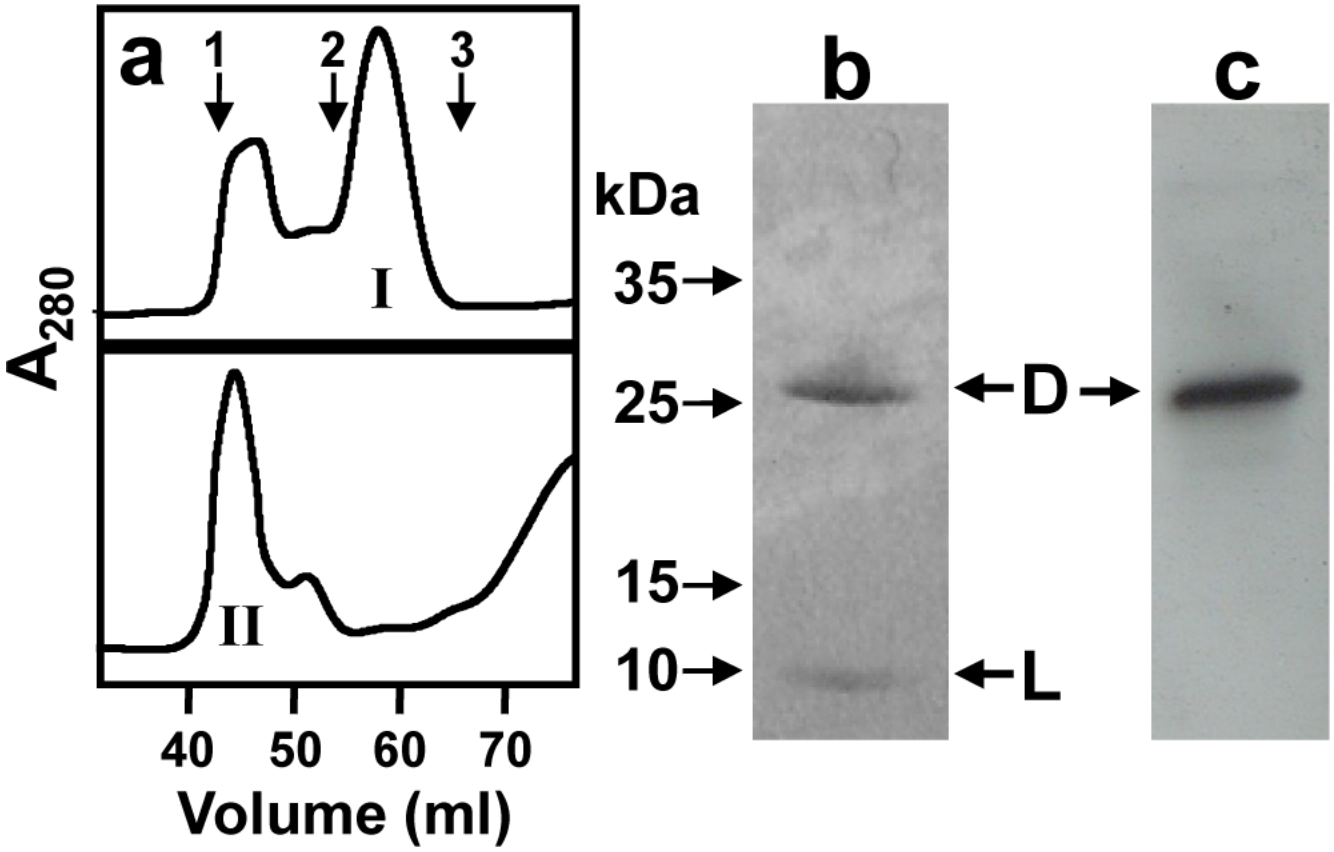

Figure 4.

The Fe-S cluster may play a structural role in D subunit folding. a, Wild-type or variant D subunits were overexpressed in *E. coli* with the L subunit. Partially purified proteins were analyzed by gel-filtration column chromatography (top: wild-type, bottom: variant). Elution of marker proteins are also shown (1: γ-globulin/158 kDa; 2: ovalbumin/44 kDa; 3: myoglobin/17 kDa). The wild-type D and L subunits formed an heterodimer, and it was eluted ~62 ml (a, top panel: peak I, mean MW: ~30 kDa, MW of the subcomplex is 40.6 kDa), which was confirmed by Coomassie stained SDS-PAGE (b). The D subunit variant was eluted ~43 ml (a, bottom panel: peak II, mean MW: ~150 kDa) without L subunit that was verified by Western blot analysis using antibody against *S. solfataricus* subcomplex D/L (c).