Published online 30 July 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 19* ***9941–9952***
doi:10.1093/nar/gks692

Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA

Magdalena N. Wojtas[1], Maria Mogni[2], Oscar Millet[1], Stephen D. Bell[2,*] and Nicola G. A. Abrescia[1,3,*]

[1]Structural Biology Unit, CIC bioGUNE, CIBERehd, 48160 Derio, Spain, [2]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE, UK and [3]IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain

Received May 1, 2012; Revised June 22, 2012; Accepted June 25, 2012

ABSTRACT

Multi-subunit RNA polymerases (RNAPs) in all three domains of life share a common ancestry. The composition of the archaeal RNAP (*a*RNAP) is not identical between phyla and species, with subunits Rpo8 and Rpo13 found in restricted subsets of archaea. While Rpo8 has an ortholog, Rpb8, in the nuclear eukaryal RNAPs, Rpo13 lacks clear eukaryal orthologs. Here, we report crystal structures of the DNA-bound and free form of the *a*RNAP from *Sulfolobus shibatae.* Together with biochemical and biophysical analyses, these data show that Rpo13 C-terminus binds non-specifically to double-stranded DNA. These interactions map on our RNAP–DNA binary complex on the downstream DNA at the far end of the DNA entry channel. Our findings thus support Rpo13 as a RNAP–DNA stabilization factor, a role reminiscent of eukaryotic general transcriptional factors. The data further yield insight into the mechanisms and evolution of RNAP–DNA interaction.

INTRODUCTION

The archaeal transcription apparatus is a simplified form of the eukaryal machinery (1). In addition to the conservation of RNA polymerase (RNAP) subunits, the archaeal general transcription factors are related to those of eukaryotes, with orthologs of the TATA-box binding protein (TBP), TFIIB and the alpha subunit of TFIIE present in archaea. While homologs of the XPB and XPD helicase subunits of TFIIH are found in archaeal genomes, there is no evidence for them playing roles in transcription (2). To date, no archaeal counterparts of TFIIA or TFIIF have been detected. Intriguingly, the RNAP II accessory factor TFIIF subunits (Rap74 and Rap30 in mammalian and Tfg1 and Tfg2 in yeast) are homologous to constitutively attached subunits of RNAP I (A49 and A34.5) and RNAP III (C37 and C53), suggesting that the last common ancestor of all three nuclear RNAPs possessed a TFIIF-like factor or subunit(s) (3,4). TFIIF facilitates promoter binding by RNAP II via contacts with DNA downstream of the transcription start site (TSS) (5). In contrast to the three task-specific nuclear RNAPs present in Eukarya, whose subunit composition ranges from 12 to 17, Archaea possess only one single RNAP. Nevertheless, the presence of subunits Rpo8 and Rpo13, described in the structure of the complete 13-subunit archaeal RNAP (*a*RNAP) from *Sulfolobus shibatae* (*Ssh*RNAP) (6), mainly distinguish the architecture of *a*RNAPs between different species and phyla.

Given the pivotal location proximal to the downstream-DNA binding cleft, the α-helical secondary structure and the helix-turn-helix (HTH) topology of Rpo13, we investigated the role of Rpo13 in transcription, reporting here the X-ray crystal structures of the DNA-bound *a*RNAP binary complex at 4.3 Å and of the naked *Ssh*RNAP improved at 3.2 Å resolution. Aided by structure comparisons, circular dichroism (CD), NMR spectroscopy, chromatin immuno-precipitation sequencing and electrophoretic mobility essays (EMSAs), we have built an atomic and molecular model for the *a*RNAP–DNA complex. Our model unveils a mechanism of RNAP–DNA complex stabilization via Rpo13 that might pre-date the separation of archaeal and eukaryotic RNAPs. Interestingly, CD and NMR data reveal that Rpo13 behaves as a molten-globule protein with intrinsically disordered regions (IDRs).

These results yield knowledge into archaeal transcription and Rpo13's function and provide mechanistic insights into the evolution of the RNAP–DNA stabilization mechanisms across the three domains of life.

*To whom correspondence should be addressed. Tel: +34 946 572 523; Fax: +34 946 572 502; Email: nabrescia@cicbiogune.es
Correspondence may also be addressed to Stephen D. Bell. Tel: +44 1865 275 564; Fax: +44 1865 275 515; Email: stephen.bell@path.ox.ac.uk

MATERIALS AND METHODS

Crystallization of naked *a*RNAP and *a*RNAP–DNA complex

The *a*RNAP was produced as previously described (7). Crystals were obtained using vapour diffusion technique using nanolitre sitting-drop (400 nl) set up by a Mosquito robot. Crystallization conditions were set up around the initial hit (6) by increasing the concentration of KCl up to 500 mM and including 2 mM $ZnCl_2$. Crystals were flash-frozen in liquid nitrogen using the crystallization solution with 25% glycerol as a cryo-protectant. To obtain crystals of the RNAP in complex with DNA (Figure 1A), prior to DNA soaking the crystals were de-salted using a newly implemented protocol (M.N.W. and N.G.A.A. submitted for publication). Of the designed deoxy-oligonucleotides inspired on the T6 promoter that we screened, the successful sequence is shown in Figure 1B and incorporates a mismatch region mimicking a bubble formation. Diffraction data were collected at the ESRF (*a*RNAP–DNA complex at 4.3 Å resolution) and at the SLS (apo-*a*RNAP at 3.2 Å resolution) and processed using HKL2000 (8) (Table 1).

Structure solution and refinement

First, we solved the apo-*a*RNAP structure by molecular replacement technique in Phaser (9) using the previous published *a*RNAP model (PDB ID 2WAQ). The refinement to 3.2 Å resolution was carried out using Phenix version 1.6.2 (10) with tight X-ray/stereochemistry weight (0.01), secondary structure restraints, TLS and B-group. The quality of the resulting $(2F_{obs}-F_{calc})e^{i\Psi c}$ and $(F_{obs}-F_{calc})e^{i\Psi c}$ maps at different stages allowed the manual re-building in COOT (11) of several regions including subunit Rpo13 (Supplementary Methods).

The DNA-bound polymerase structure was solved using Phaser with the refined 3.2 Å naked-RNAP structure. The two molecules in the asymmetric unit were rigid-body refined with one rigid body per subunit, each of which with a group B-factor (Wilson $B_{factor} = 105 Å^2$). One TLS was assigned to each RNAP subunit. The outcoming $(2Fo-Fc)e^{i\Psi c}$ and $(Fo-Fc)e^{i\Psi c}$ maps showed clear density within the DNA entry channel. Then, to aid the fitting of the DNA molecule, the $(2Fo-Fc)e^{i\Psi c}$ map was 2-fold averaged and solvent flattened using the General Averaging Program (Stuart, D.I. and Grimes, J.M., unpublished software available on request). The averaging procedure allowed the comparison of the initial $(2Fo-Fc)e^{i\Psi c}$ map, the averaged map $F_{ave}e^{i\Psi ave}$ and the $(2Fo-Fc)e^{i\Psi ave}$ map for the DNA region (Supplementary Figure S1). The unequivocal density corresponding to double-stranded DNA (dsDNA) was then fitted in COOT with an idealized B-form dsDNA model and then rigid-body refined. Nucleotides with no 2Fo–Fc density were removed and the remaining ones grouped in 2-bp steps and rigid-body refined. After minor adjustments, careful refinement resumed with five cycles of individual positions refinement with tight geometry restraints (0.01), 2-fold NCS and secondary structure restraints including Watson–Crick base pairing.

The final R_{work}/R_{free} (%) of the free and DNA-bound RNAPs are, respectively, 24.1/30.0 and 29.2/31.1% (Table 1). Figures were created with Pymol software (http://www.pymol.org/).

Chromatin immuno-precipitation sequencing

Chromatin immuno-precipitation (ChIP) was performed essentially as described (12). Formaldehyde was added to a final concentration of 1% to an exponentially growing culture of *Sulfolobus acidocaldarius* which was then cooled to room temperature for 20 min. The reaction was quenched by the addition of glycine to 125 mM. Cells were harvested and washed with phosphate-buffered saline, then resuspended and lysed in 4 ml TBSTT (20 mM Tris-Cl [pH 7.5], 150 mM NaCl, 0.1% Tween 20, 0.1% Triton X-100). The extract was sonicated and clarified by centrifugation. Immuno-precipitations were performed by incubating 3 µl α-Rpo2 [a kind gift from S. Qureshi (13)] or α-Rpo13 antiserum with 10 µg of extract in 100 µl TBSTT with shaking at 4°C for 3 h. After addition of 25 µl of 50% slurry of protein A sepharose, incubation was continued further for an hour. Immune complexes were collected by centrifugation and washed by five consecutive 5 min incubations with 1.2 ml TBSTT with vigorous shaking at room temperature. Beads were then washed once with TBSTT containing 500 mM NaCl and once with TBSTT containing Tween 20 and Triton X-100 at 0.5%. Finally, immune complexes were disrupted by resuspending the beads in 20 mM Tris-Cl [pH 7.8], 10 mM EDTA and 0.5% SDS and heating to 65°C for 30 min. Beads were removed, and DNA was recovered by treating samples with 10 µg/ml proteinase K for 6 h at 65°C then 10 h at 37°C, followed by extraction with phenol/chloroform/isoamyl alcohol and back-extraction of the organic phase with TE buffer, then chloroform extraction. Finally, DNA was precipitated with ethanol in the presence of 20 µg glycogen then resuspended in 50 µl TE buffer. Input samples were treated as above without the addition of antiserum and beads. DNA was processed for next-generation sequencing on an Illumina HiSeq 2000 by Source Bioscience (Nottingham, UK). More than 60 000 000 reads were obtained per sample, each yielding over 5×10^9 bases, of which >85% were of Q30 or higher quality and over 70% mapped to the *S. acidocaldarius* genome. Reads were mapped to the *S. acidocaldarius* genome using SAMtools and BEDTools (http://samtools.sourceforge.net/ http://code.google.com/p/bedtools/) to generate .bed files which were then visualized and quantified using SeqMonk (http://www.bioinformatics.bbsrc.ac.uk/projects/seqmonk/) with read counts evaluated over a running window probe size of 20 nt and normalized to input DNA.

DNA–Rpo13 binding assay

EMSAs involving Rpo13 were performed with radio-labelled single-stranded 40 nt DNA (ORB1 mut2 top 5′-ATATTTACCTTAAGTTCTAACGTGGAAACAAA GGGGTTTT-3′), where the double-stranded probe was

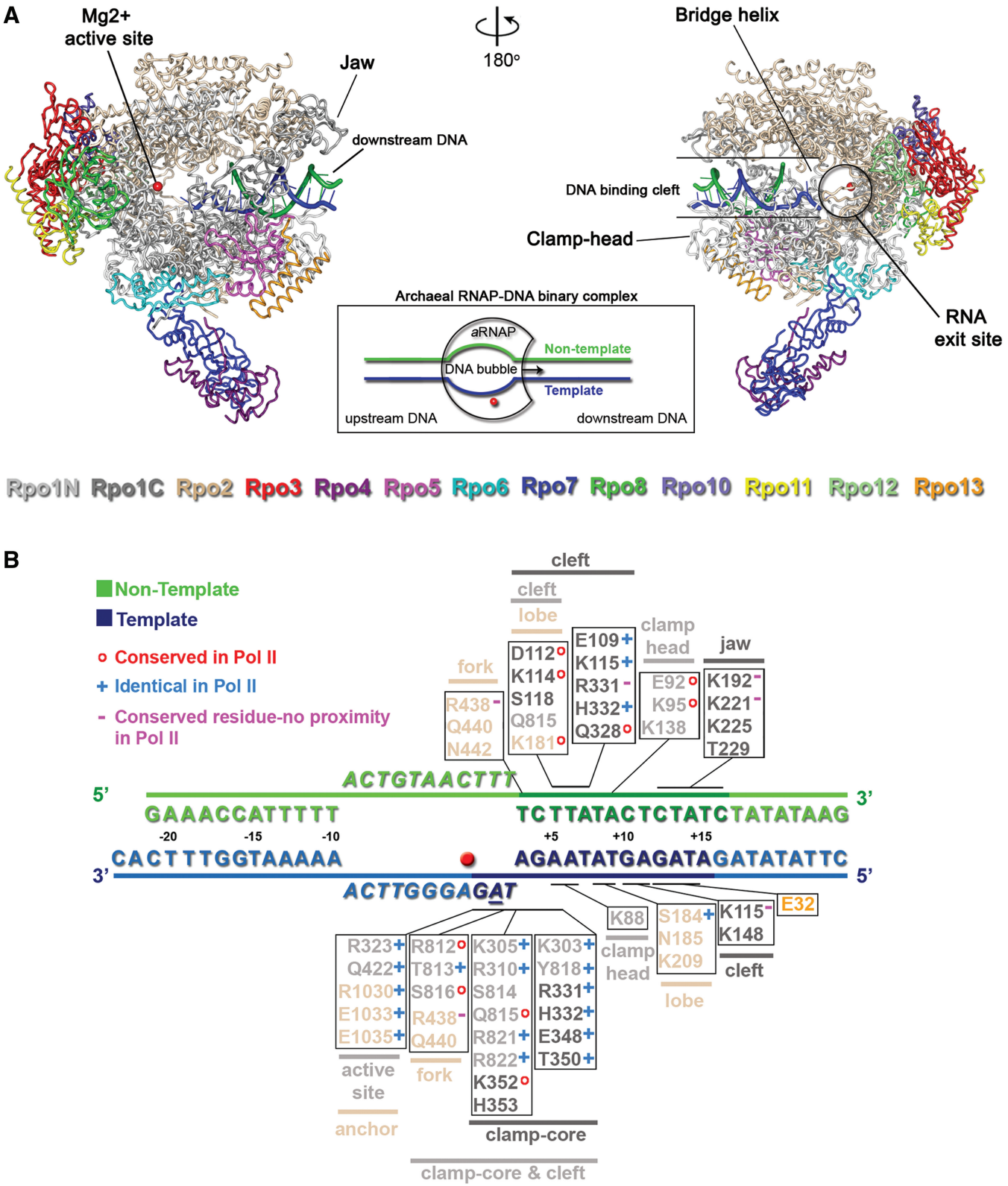

Figure 1. *a*RNAP–DNA binary complex. (**A**) DNA-bound *a*RNAP structure represented as cartoon tube with schematic in the centre; DNA and *a*RNAP colour coded as in keys (centre and below). (**B**) DNA sequence used for soaking coloured as in (A); characters with darker colours indicate the structurally ordered nucleotides; nucleotides outside the box pre-form the DNA bubble. Negative and positive numbers indicate upstream and downstream nucleotide/position with respect to the start-site *A* (due to the limited resolution this sequence assignment is not unique; see Supplementary Figure S2). Residues within 8Å distance from the DNA are indicated and colour-coded according to subunit and domain to which they belong. The red circle indicates the magnesium ion at the catalytic site.

Table 1. Data collection and refinement statistics

Data collection	Swiss light synchrotron	European synchrotron radiation facility
	apo-RNAP	RNAP + DNA
Beamline	SLS–PXIII	BM14
Wavelength (Å)	1.000	1.033
Space group	$P2_12_12$	$P2_1$
Unit cell parameters (Å, °)	$a = 195.7$, $b = 212.4$, $c = 128.8$	$a = 134.1$, $b = 199.4$, $c = 214.2$, $\beta = 103.5$
Molecule per AU	1	2
Resolution range (Å)	40.1–3.20 (3.27–3.20)	60–4.30 (4.45–4.30)
No. of images, $\Delta\Phi$(°)	720, 0.5	440, 0.5
Unique reflections	88 449	65 048
Redundancy[a]	15.2 (15.3)	4.1 (4.0)
Completeness[a] (%)	99.5 (99.1)	92.1 (77.8)
I/σ (I)[a]	13.9 (2.1)	6.0 (2.0)
R_{merge}[a] (%)	18.0 (–)	21.1 (66.7)
Refinement statistics		
Resolution range (Å)	40.1–3.20	50.4–4.32
No. of reflections (total/test)	88 376/4426	64 957/3301
R_{work}/R_{free} (%)	24.1/30.0	29.2/31.1
No. of protein atoms	27 103	54 193
Wilson/average B_{factor} (Å^2)	65.0/96.5	105.2/173.9
RMS Δ bond/angle (Å, °)	0.007/0.762	0.005/0.667
Ramachandran statistics[b]: Favoured/Allowed/ Outliers (%)	85.3/11.5/3.2	85.6/11.5/2.9

[a]Statistics for the highest shell are given in parentheses.
[b]Using Rampage software (9).

generated by annealing the above oligonucleotide to its complement (5′-AAAACCCCTTTGTTTCCACGTTAG AACTTAAGGTAAATAT-3′). Binding reactions were performed in 50 mM Tris-Cl [pH 8], 25 mM $MgCl_2$, 75 mM KCl, 10 mM DTT, 5% glycerol and 0.1 mg/ml of BSA for 15 min at 48°C (14). Protein–DNA complexes were electrophoresed on native 10% acrylamide gels in Tris-glycine buffer at 13 V cm^{-1} before drying and phosphoimagery.

RESULTS

X-ray structure of the *a*RNAP–DNA binary complex

We solved the structure of *Ssh*RNAP bound to a duplex DNA oligonucleotide with an internal pre-formed bubble to facilitate loading into the active site (Figure 1). The DNA fragment was soaked in pre-treated naked RNAP crystals (M.N.W. and N.G.A.A., submitted for publication). Of the several crystals screened, only one, in space group $P2_1$, diffracted successfully to 4.3 Å resolution (Table 1). The structure was solved by molecular replacement using the improved 3.2 Å apo-*a*RNAP structure. The two RNAP molecules composing the asymmetric unit clearly show density corresponding to bound DNA (Supplementary Figure S1). This *a*RNAP–DNA complex structure (Figure 1A) lacks general transcription factors and thus cannot strictly depict an open complex (OC) state. However, we note that the *Sulfolobus* TFB promoter has a start site that is factor independent and therefore our complex could either mimic an open complex-like state (13) or be akin to an elongation complex without an RNA primer in the transcription bubble. These complexes can also be formed by the eukaryotic RNAPs independently of RNA primers (15). Because of the relatively low resolution of the *a*RNAP–DNA structure, the DNA sequence cannot be unambiguosly assigned in the structure and in principle alternative DNA binding modes can be envisaged (See Supplementary Figure S2). However, crystal packing and DNA melting temperature considerations favour the scheme illustrated in Figure 1.

Our RNAP–DNA binary complex informs on the possible interactions that allow positioning of the nucleic acid within the DNA binding cleft, highlighting those residues that are conserved across Pol II and *a*RNAP and those that are archaeal specific (Figure 1B). Furthermore, the archaeal binary complex reveals that from the fork-loop 2 (Rpo2: amino acids 434–445), the dsDNA unwinds at position +3 leaving 10 bp (from +4 to +10) in canonical B-form (Figure 2). Fork-loop 2 and switches 1 (Rpo1C: amino acids 329–351) and 2 (Rpo1N: amino acids 301–319) are in close proximity to the template strand (Figure 2A). Densities for nucleotides in the template strand at positions +2 and +1 (TSS) are visible although weaker for the +1 base. Upstream of the TSS, the density becomes increasingly weaker, indicating that the nucleic acid strand is not further ordered (Figure 2). Similarly, there is no evidence of ordered nucleotides upstream of position +2 in the non-template strand. Downstream, at positions +14 and +15, Watson–Crick pairing is lost and bases appear to swing out and the remaining nucleotides are not resolved (Figures 1B and 2).

The overall geometry of the translocation coordinator bridge helix (Rpo1N: amino acids 791–827) is bent, as in the eukaryal counterpart OC, (Figure 2B) with a maximum bending angle of ~24.6° in correspondence of S814 (~22.4° in correspondence of K830 in Pol II OC) which is the closest residue to the start site phosphate group (~5 Å Cα-P). Finally, residues from 87 to 99 within the trigger loop (Rpo1C: amino acids 71–116) remain disordered as it is also the case of the eukaryal Pol II OC (PDB ID 4A3I) and transcription elongation complex (TEC) (PDB ID 1R9T).

*a*RNAP–DNA *versus* eukaryotic RNAP–DNA complexes

While the interactions of the *Ssh*RNAP with the DNA approaching the active site are equivalent to those observed in the eukaryotic Pol II OC and TEC (15–18) (Figures 1B and 2A) four key differences distinguish the DNA-bound *a*RNAP. First, stabilization of the downstream DNA occurs at positions +13 and +12 of the non-template strand via a lysine-rich region of the jaw domain (Figures 1B, 3A, left and 3B) that is constitutively further in the cleft; ~20° rotation inwards is necessary for the eukaryotic jaw domain to lock onto the archaeal counterpart, a movement hindered by the contacts with the adjacent eukaryotic specific Rpb9 subunit (absent in Archaea) (Figure 3A, right). Second, the downstream

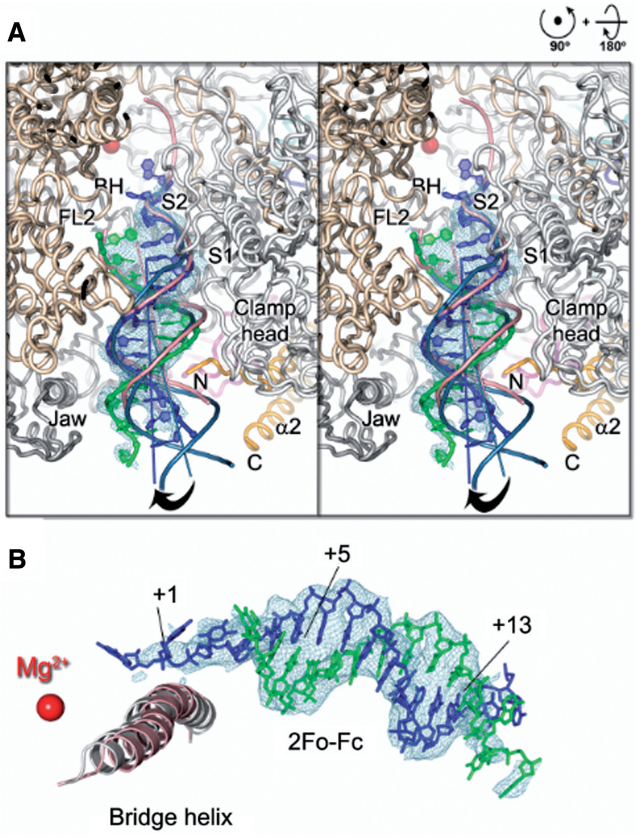

Figure 2. *a*RNAP-DNA binding cleft interactions and bridge helix. (**A**) Stereo-figure of the DNA with 2Fo–Fc map (slate-blue mesh at 1.0σ) with refined DNA phases included; *a*RNAP's subunits as cartoon tube with semi-transparent surface and DNA coloured as Figure 1; in light pink and steel blue the DNA backbone (as tube) of the eukaryotic Pol II–DNA open complex (PDB ID 4A3I) and of Pol II DNA/RNA elongation complex (PDB ID 1R9T) superimposed onto the DNA bound to our *a*RNAP [corresponding superimpositions of Rpb1 onto Rpo1, RMSD 2.8 Å, 1183 Cαs equivalences and 2.8Å, 1174 Cαs equivalences were performed using the Structural Homology Program (SHP) (33)]; dark blue, light pink, and steel blue straight lines correspond to helix axes; black curved arrow indicates the tilting required (∼17°) for superimposing the DNA of Pol II TEC (steel blue axis) onto the DNA bound to *a*RNAP (dark blue axis); a tilting of ∼11° in the same direction is required to superimpose the DNA of Pol II OC (light pink axis) onto the DNA bound to *a*RNAP (dark blue axis). Labelled some of the relevant structural domains; geometrical symbols on the right corner of panel relate the current view relative to the RNAP as viewed in Figure 1A, left. (**B**) Superimposition of bridge helices of the archaeal RNAP (white smoke) and Pol II OC (light pink) represented as cartoon with DNA bound to *a*RNAP depicted and density contoured as in (A) with numbering describing the nucleotide positions with respect to the TSS +1.

DNA path deviates, respectively, by ∼11° and ∼17° from the helix direction observed in the downstream dsDNA helix-axis in the binary Pol II–DNA OC (15) and Pol II–DNA/RNA TEC (18) (Figures 2A and 3B). In eukarya, the N-terminal domain of Rpb5 appears to act as a 'guard rail' for DNA docking, thus helping define its trajectory (Figure 3B). Notably, archaeal Rpo5 lacks this domain, its place, in *Sulfolobus*, instead being occupied by Rpo13 (Figure 3B).

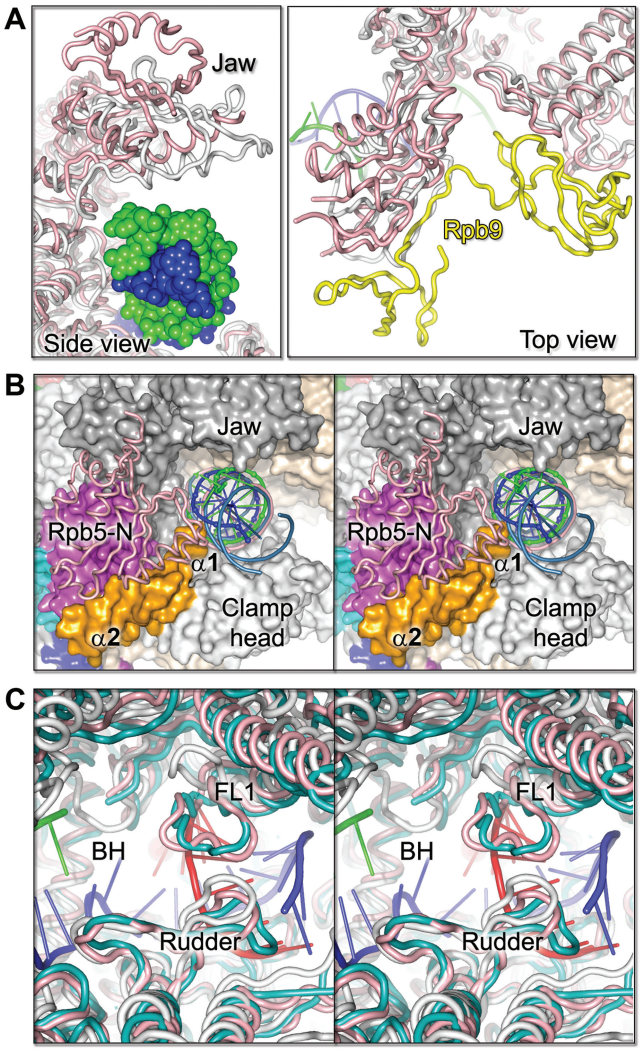

Figure 3. Comparisons of archaeal and eukaryotic structural domains downstream the DNA binding cleft and RNA exit channel. (**A**) Left, superimposition of the eukaryotic Rpb1 (light pink cartoon tube; PDB ID 4A3I) onto the archaeal Rpo1N and Rpo1C (white smoke as cartoon tube) complexed with DNA (this study) represented as spheres (non-template and template strands in green and blue, respectively) viewed *quasi* along the DNA axis; analysis of the hinge movement (∼20°) that the eukaryotic Jaw domain (amino acids 1141–1275) would need to superimpose onto the archaeal counterpart was performed with SHP (Similar results were obtained using the eukaryotic Rpb1 in PDB IDs 1Y1W and 1R9T). Right, same superimposition as in left but viewed from the top of the Jaw domain with eukaryotic Rpb9 coloured in yellow interacting with the eukaryotic Jaw domain; DNA depicted as backbone. (**B**) Stereo-view of the DNA binding cleft viewed along the archaeal DNA helix axis; *a*RNAP is represented as surface and colour coded as in Figure 1A; in cartoon tube (light pink) the corresponding location of Rpb5 N-terminal subunit in Pol II and corresponding DNA helix axes as in Figure 2A. (**C**) Stereo-view of the region corresponding to the RNA exit channel with the *a*RNAP–DNA (white smoke), Pol II open (light pink) and transcription (blue aquamarine) complexes superimposed; the shorter FL1 in the *a*RNAP–DNA is constitutively in 'up' conformation in contrast to its counterpart in the Pol II open complex (light pink). The DNA (template in blue and non-template in green) in the binding cleft corresponds to the current study whereas the RNA (red) and DNA (blue) in the exit channel corresponds to the elongation complex (PDB ID 1R9T), all represented as cartoon tube.

Third, contrary to the disorder/order transition observed in some important structural elements [e.g. Fork loops 1 (FL1) and 2 (FL2)] between eukaryotic naked Pol II (PDB ID 1WCM) and Pol II–DNA/RNA complexes upon nucleic acids binding (for comparison see PDB IDs 4A3I, 1R9T and 1Y1W), the archaeal apo- and DNA-bound RNAP forms do not show any striking order conversions or major re-adjustments upon DNA binding. All switches, 1–5, are ordered in the apo- and DNA-bound *a*RNAP structures, however, we note that switch 3, facing the RNA exit path, might readjust in presence of RNA whereas the flap remains disordered as in the case of the eukaryal apo- and nucleic acid bound Pol II. Interestingly, FL1 is also ordered in both archaeal structures (in apo-Pol II it is disordered; PDB ID 1Y1W), probably due to being shorter by ~6 residues compared with the equivalent FL1 in Pol II, but it adopts an 'up' conformation (Figure 3C) that may lock down once RNA is synthesized. As with FL1, FL2 is well-defined (Figure 2A) as is its eukaryal counterpart in the complete elongation complex (PDB ID 1Y1W). However, this same loop is not ordered in the Pol II open and transcribing complexes (PDB IDs 4A3I, 1R9T). These findings indicate that *a*RNAPs require only minimal adjustments for the transition from the DNA unbound-to-bound form.

Fourth, structural comparison of the individual subunits (and sub-domains) between the archaeal apo- and the DNA-bound form of the *a*RNAP shows no open–close movement of the clamp upon DNA binding. We point out the obvious caveat that alternative conformations of the RNAP clamp could exist in solution. Nonetheless, minor movements ($<2^\circ$ rotation) can be detected in the jaw domain (aa 147–234 in Rpo1C) and in the clamp-head (amino acids 86–208 in Rpo1N) (Figure 4A). However, DNA-binding does not induce folding of Rpo13's glutamate-rich N- and lysine-rich C-terminal ends, though deviation of the DNA helical path and additional density (0.7σ) at the N-terminus suggest interactions with the DNA (Figures 2A, 3B and Supplementary Figure S3). The improved model of Rpo13 from the naked RNAP structure (see below; Figure 4B) together with the first archaeal DNA-bound RNAP structure (Figure 1) prompted a thorough biochemical, biophysical and systemic analysis of Rpo13.

Apo-*a*RNAP at 3.2 Å: structural basis of Rpo13-RNAP specificity

Prior to obtaining the *a*RNAP–DNA crystal, we solved the structure of the apo-form of *Ssh*RNAP at 3.2 Å, improving on the previous 3.4 Å structure (6). The apparent minimal increase in resolution translates into ~29 000 additional measured reflections and into a higher signal-to-noise ratio (overall and in the highest resolution shell) and thus in higher completeness, contributing to a more stable structure refinement and a more reliable *a*RNAP model (Table 1). The most important result, a consequence of the improved quality of the electron density, has been the unequivocal register of the sequence of Rpo13 (Supplemental Methods). This has revealed the basis of its interaction with the RNAP. The core of Rpo13 forms a HTH comprising helix-1 (α_1; residues 38–56) and helix-2 (α_2; residues 61–82) with the first visible Cα N-terminal residue (F33) facing the DNA-binding cleft (Figure 4B). Density at both the N- and C-termini is relatively weak (0.6σ) and CD and NMR analysis of recombinant Rpo13 (see below) show that residues outside the core HTH motif are indeed inherently flexible. The two helices' axes are almost parallel to each other and this conformation appears to be stabilized by inter-helix contacts mainly of hydrophobic nature (I50, I59, A64, L67). α_1 docks into the RNAP complex (Figure 4B) with hydrophobic contributions from residues F58 and L44, whereas α_2 is the helix most exposed to the solvent and thus, in principle, available for interactions with additional ligands. α_2 is characterized by polar residues. Interestingly, this region shows some similarity to the charged region of Tfg1-like factors of TFIIF, Supplementary Figure S4). Residues N47 and W51 with T48 and D52 of Rpo13's α_1 confer specificity for binding to the RNAP. The former amino acids bind to the Rpo5 subunit and the latter two to the clamp-head domain of Rpo1N. Two tryptophan residues, one from Rpo13 (W51) and the other from the clamp-head (W125) reciprocally dock via hydrogen bonds onto G60 of Rpo5 and T48 of Rpo13, respectively (Figure 4B). Further hydrogen bonds between Rpo13 N47 and P58 of Rpo5 and Rpo13 D52 and R124 of the clamp-head domain strengthen the binding of Rpo13 to the crevice formed by Rpo5 and the clamp head (Figure 4B).

Moreover, the improved data allowed us to put in register segments of sequence in Rpo1 and Rpo2 subunits and to model a fully ordered clamp-core and -head domains in the Rpo1N subunit (residues 32–42 and 155–177); the latter adopting a three-antiparallel strands motif (Figure 4B). In Rpo2, residues 55–58, 180–187 and 195–199 of the protrusion domain, residues 589–603 of the external two domains and residues 1071–1091 of the hybrid-binding domain have been also re-modelled.

Rpo13 subunit: a molten-globule protein

Rpo13's sequence is characterized by a strikingly polarized distribution of negative and positive charges (14 glutamates in the first 32 residues and 11 lysines in the last 22 residues) (Figure 4B). These terminal ends are not seen in any of the *a*RNAP structures so far solved (6,19,20). Preliminary biophysical analysis on Rpo13 purified following expression in *E. coli* has shown that in solution Rpo13 forms predominantly dimers (20), a behaviour displayed by other transcription factors, including TFIIF which contains a specific dimerization domain preceding the charged region (3,5). CD spectroscopy in the far UV region on Rpo13 reveals a protein with ~35% α-helical, 48% random coil and 17% beta-sheet content (Figure 5A, left; Supplementary Methods). This secondary structure assignment of Rpo13 in solution fits the X-ray secondary structural description of Rpo13 bound to the entire enzyme where only 39 residues out of 104 are ordered. Also, controlled heat-induced unfolding experiments over the temperature range 22°C to 82°C

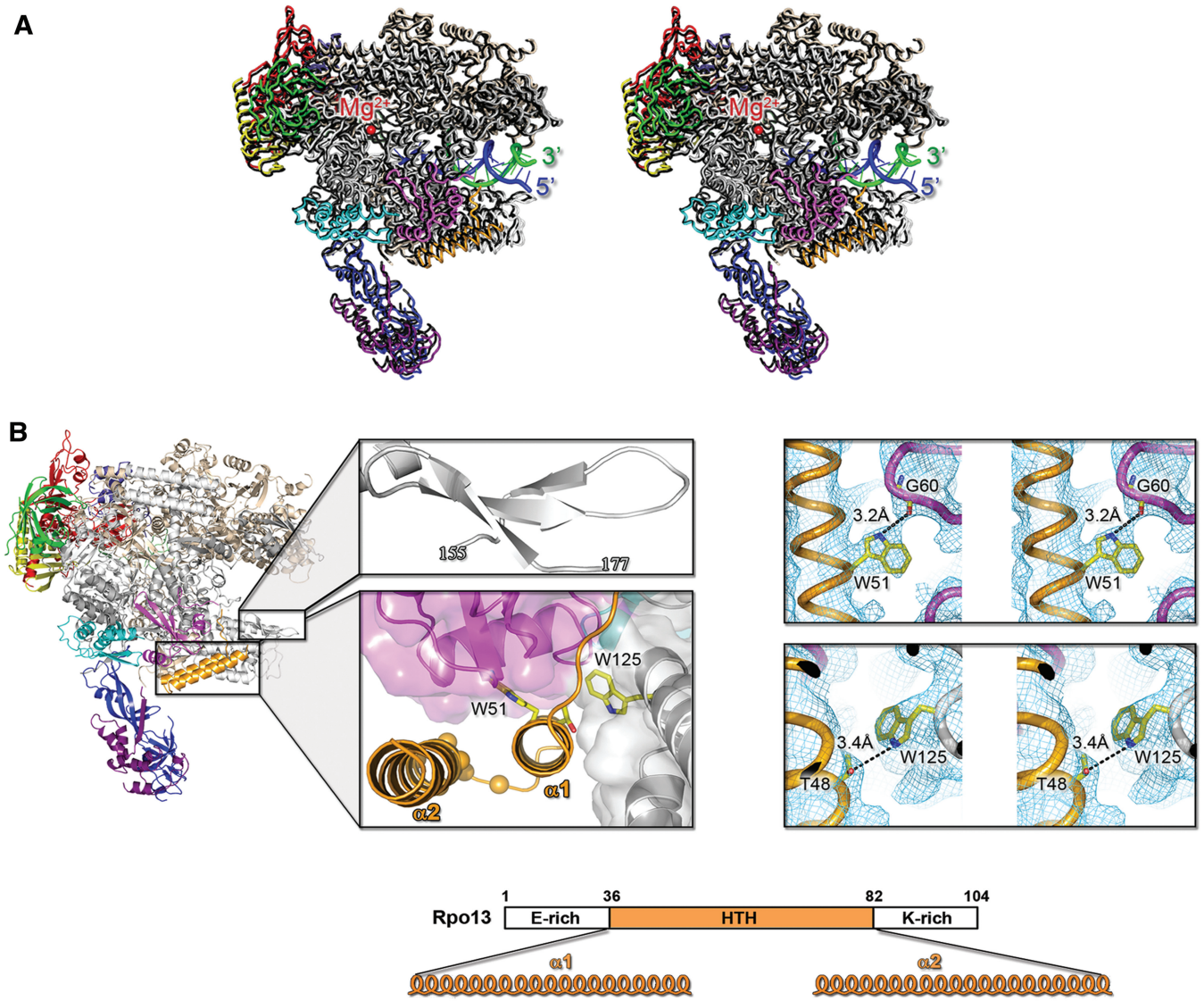

Figure 4. Apo-*a*RNAP structure at 3.2 Å and comparison with DNA-bound form. (**A**) Top, stereo-view of superimposed apo-*a*RNAP structure (black) onto DNA-bound *a*RNAP (coloured as Figure 1A); as red sphere, the magnesium ion in the active site. Superimposition of the apo-*a*RNAP onto the *a*RNAP–DNA complex was carried out with the SHP software (33) using the active-site (amino acids 319–483) as pivot. (**B**) Left, apo-*a*RNAP structure colour-coded by subunit as Figure 1A. Centre top, the modelled clamp head-domain. Centre below, the newly modelled Rpo13 fitting the crevice formed by Rpo5 and clamp-head (as cartoon and semi-transparent surface); as spheres the Cα of Rpo13's conserved patch ISVDEA (see Supplementary Figure S4); in sticks the tryptophans with binding partners. Right, stereo-view of details of the Rpo13 α_1 helix and Rpo5 (top) and clamp-head domain (below) interactions with 2Fo–Fc density (slate blue mesh at 1.0σ) providing Rpo13 specificity for the RNAP; below panels Rpo13 sequence layout.

(Figure 5A, right) show that the percentage of the α-helical content is marginally affected at 82°C (approximately 25%), perhaps not surprising, given the extreme growth conditions of *Sulfolobus* species (75–80°C) (21). On the other hand, the thermal denaturation curves obtained by monitoring the CD signal at 222 nm in both directions overlap but do not show a characteristic sigmoidal unfolding curve of a protein with a well-defined tertiary structure (Supplementary Figure S5). This observation suggests that the folding into a topological HTH motif is coupled to the association with the rest of the archaeal enzyme.

Full-length Rpo13 was also studied by NMR spectroscopy (Supplemental Methods). The ^{1}H–^{15}N–HSQC spectrum (Figure 5B) shows a poor dispersion in the proton chemical shift, indicating a loss in both secondary and tertiary structure. Interestingly, the analysis of the peaks cannot account for all residues (even considering signal overlap) and the spectrum is showing resonances for a subset of residues only. The missing peaks that include the tryptophan residue used for the specific interaction with Rpo1N, may have disappeared because of conformational exchange in the microsecond–millisecond timescale or because of segmental aggregation. This

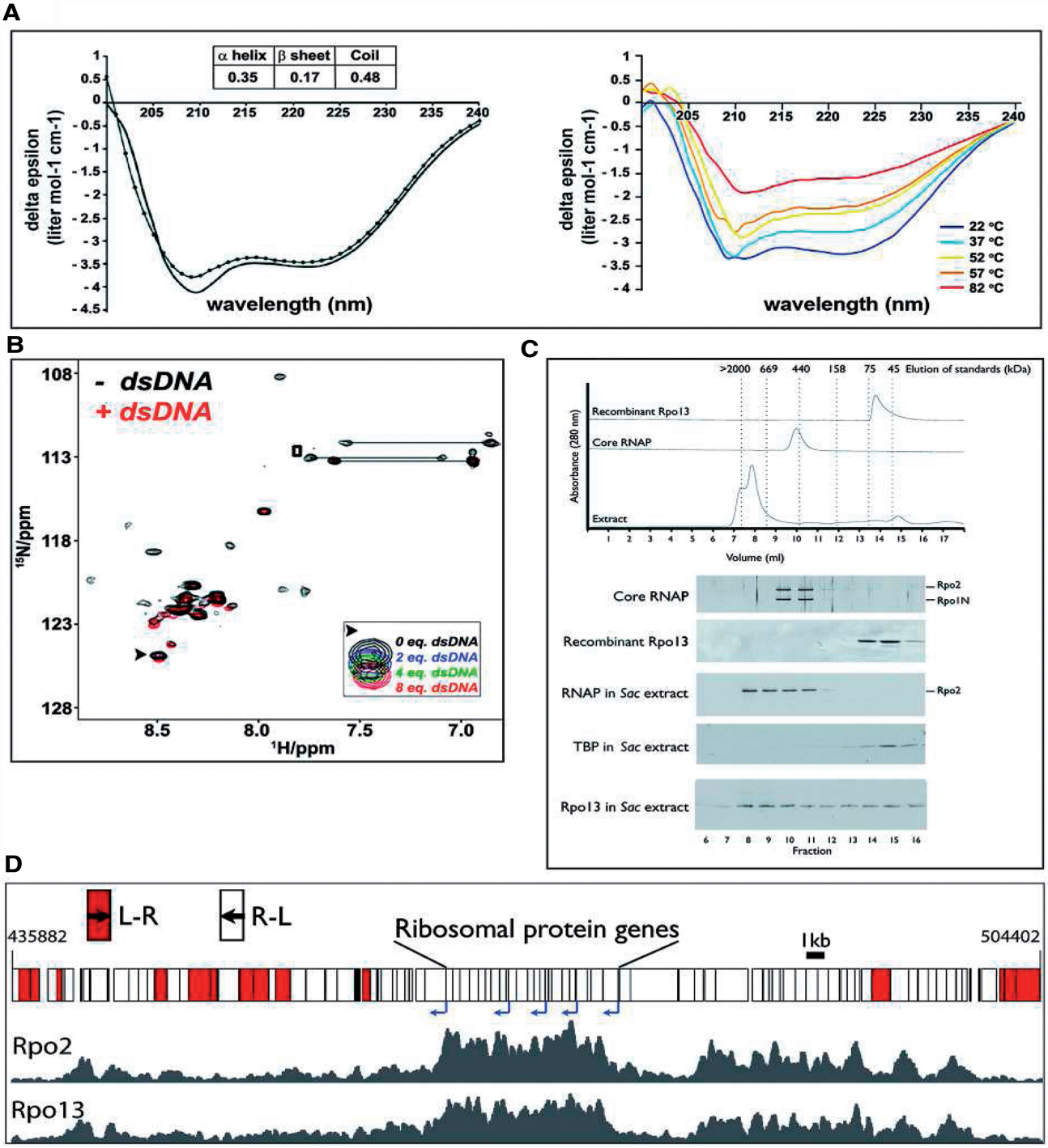

Figure 5. Characterisation of Rpo13 and its association with RNAP. (**A**) Left, CD spectrum of recombinant Rpo13 collected from 200 to 240 nm at room temperature with secondary structure assignment displayed in the inset table. The continuous line represents the experimental data whereas the dotted line is the reconstructed data by K2d software in Dichroweb (34). Right, CD spectra of Rpo13 recorded at different temperatures ranging from 22 to 82°C; measurements were taken every 5°C with 1°C /min slope. (**B**) Contour plot of 2D ^{1}H–^{15}N HSQC spectrum from based titration experiments showing peaks with chemical shift variation (filled arrow head) and chemical shift disappearance (open square) upon dsDNA binding. Both effects shall be attributed to the binding event and are probably due to the different exchange timescale introduced with the DNA binding. The solid horizontal lines connect the two peaks for the N and Q side-chain amide groups. (**C**) Upper panel, elution profiles, following absorbance at 280 nm, of recombinant *S. acidocaldarius* (*Sac*) Rpo13, purified core *Sac* RNAP and *Sac* whole cell extract over a Superdex 200 column. Lower panel, protein content of the 1 ml fractions collected from the Superdex 200 column (fractions are aligned to the chromatographic traces in the upper panel). The Rpo2 and Rpo1N components of pure RNAP were detected by silver staining of the gel following SDS PAGE, recombinant Rpo13 by Coomassie staining, proteins in the remaining panels were detected by western blotting with the indicated antisera. (**D**) ChIP-Seq data generated using antisera against Rpo2 (middle track) or Rpo13 (bottom track) are shown for coordinates 435 882–504 402 of the *S. acidocaldarius* chromosome. Open reading frames are shown as boxes in the top track; red indicates transcription left to right and white the opposite direction, a region that encodes a highly transcribed cluster of ribosomal protein genes is indicated. The arrows underneath show the positions of promoters for the polycistronic units [data from *S. solfataricus* (35), the gene order of the ribosomal protein cluster is conserved in *S. acidocaldarius*].

complex behaviour implies a molten globular structure that undergoes exchange between several conformations.

Rpo13, an *in vivo* constituent RNAP subunit that binds to dsDNA

The Rpo13–DNA interactions inferred from our *a*RNAP-DNA X-ray structure and Rpo13's molten-globule nature prompted us to address the related questions of the nature of the association of this protein with the RNAP and the role of Rpo13 in DNA binding. We therefore sought to address whether (i) Rpo13 is only found associated with soluble RNAP or if a free pool of the Rpo13 could exist in cells, (ii) Rpo13 is associated with the RNAP during *in vivo* transcription and (iii) the recombinant protein might bind DNA.

First, we prepared a cell-free extract from *S. acidocaldarius* and fractionated this over a Superdex 200 gel filtration column. The elution profiles of bulk RNAP and Rpo13 were followed by western blotting with α-Rpo2 and α-Rpo13 antisera, respectively (Figure 5C). Purified recombinant Rpo13 (see Supplementary Methods) and purified native RNAP were also passed over the column. Purified Rpo13 eluted with a retention time compatible with formation of dimers in solution, as seen previously (20). Purified RNAP eluted at ~450 kDa. RNAP in cell-free extracts, monitored by western blotting using anti-Rpo2 antisera, had a broader elution profile than the purified core enzyme, eluting as a broad peak between the void volume of the column and the same fractions that the purified enzyme occupied. We note that the samples were pre-treated with ethidium bromide to prevent interactions with DNA, suggesting that the higher molecular weight complexes may contain additional protein factors in a holo-enzyme complex. As seen previously (13), the general transcription factor, TBP, elutes as a dimer and is not detectably associated with the RNAP. Notably, while we observe co-elution of Rpo13 with Rpo2, we also detect Rpo13 over a broad range with a minor peak (lanes 14–15 in Figure 5C) corresponding to the elution of purified recombinant Rpo13, indicating that the extract has a population of free Rpo13 protein.

The co-purification and crystallization of Rpo13 with RNAP clearly indicates a stable association in solution. We therefore wished to address whether Rpo13 is also associated with RNAP during transcription elongation *in vivo*. Accordingly, we utilized the α-Rpo2 and α-Rpo13 antisera to perform ChIP-sequencing (ChIP-Seq) experiments to profile the genomic localization of these proteins in exponentially growing asynchronous *S. acidocaldarius* cell cultures. From the examples shown in Figure 5D, it can be seen that both proteins show enrichment over transcribed regions with the highest peaks corresponding to heavily transcribed ribosomal protein genes.

We also detect elevated signals of both proteins in some intragenic regions, perhaps indicative of polymerase pausing [an extensive genome-wide analysis will be published elsewhere (M.M. and S.D.B., unpublished data)]. Taken together, these data indicate that Rpo13 co-localizes with elongating RNAP.

Finally, using EMSAs (Figure 6) we could readily detect interaction between purified recombinant Rpo13 and dsDNA with an apparent K_d of ~2 μM, consistent with our NMR data (Figure 5B). We could not detect any significant single stranded DNA (ssDNA) binding activity (Figure 6A, left). We additionally performed supershift experiments to identify Rpo13 as the DNA binder (Figure 6A, right). Neither major groove nor minor groove interacting compounds (methyl green and netropsin, respectively) had a major impact on binding to dsDNA (Supplementary Figure S6).

To identify which regions within Rpo13 were responsible for DNA binding, we prepared deletion constructs of Rpo13 lacking either or both N- and C-terminal tails and tested these in EMSAs (Figure 6B and 6C). Deletion of the N-terminal tail rendered the protein prone to aggregation, but DNA binding was still detectable at low concentrations (Figure 6C, top). Importantly, however, deletion of the C-terminal tail (amino acids 85–104) had no detectable effect on protein stability or solubility but abrogated DNA binding (Figure 6C, below).

DISCUSSION

We have provided insight into archaeal transcription and Rpo13's structure and function. We have demonstrated that the binary interactions between DNA and *a*RNAP towards the active site are ancestrally equivalent to those of eukaryotic Pol II bound to DNA and that the *a*RNAP does not show major structural rearrangements or element folding upon DNA binding. However, significant differences occur at the far end of the DNA-binding channel. Our structural and functional findings reveal how the RNAP of the model archaea of the genus *Sulfolobus* achieves and enhances DNA downstream stabilization via two universal mechanisms: (i) by a hinge movement of the jaw domain constitutively locking onto the downstream DNA and (ii) by the interaction of a novel subunit (Rpo13) that is able to bind dsDNA but not ssDNA.

We note that the N-terminal domain of Rpb5 is conserved in all eukaryal nuclear RNAPs but absent from archaea. This, together with the acquisition of the Pol II-specific Rpb9 subunit, that constrains the adjacent jaw domain, indicates that eukarya have evolved their own mechanisms for binding and orienting downstream DNA in the enzyme cleft.

Rpo13's structure from the apo- and DNA-bound *a*RNAP crystals and from the isolated Rpo13 in solution studies suggest (i) that Rpo13's secondary structure purified as individual protein recapitulates the secondary structure of Rpo13 in the crystal structure of the entire enzyme and (ii) that the unstructured component of Rpo13 considered either as individual protein or as subunit of the RNAP complex is an inherent property of the protein. Thus, Rpo13 can be classified as a molten-globule containing IDRs (22). Indeed many proteins which possess disordered regions [including 49% of human transcription factors (23,24)] function by binding

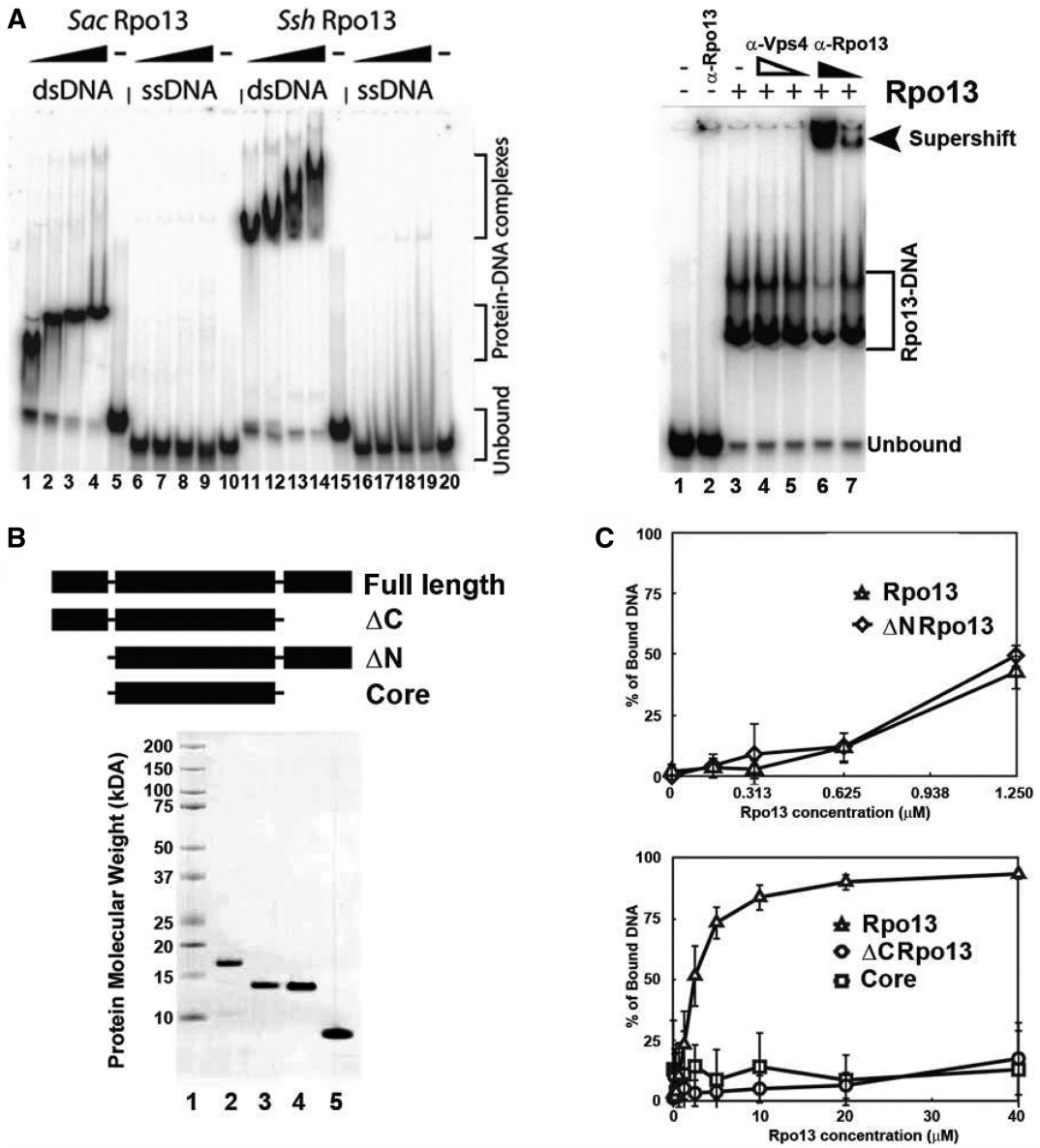

Figure 6. EMSA analysis of Rpo13 DNA binding activity. (**A**) Left, varying amounts of *S. acidocaldarius* (*Sac*) or *S. shibatae* (*Ssh*) Rpo13, purified as recombinant proteins from *E. coli*, were incubated with single-stranded 40 nt DNA or double-stranded 40 bp DNA prior to electrophoresis on a 10% polyacrylamide gel in Tris-glycine. Reactions contained 0.1 nmol (lanes 1, 6, 11 and 16), 0.2 nmol (lanes 2, 7, 12 and 17), 0.4 nmol (lanes 3, 8, 13 and 18), 0.8 nmol (lanes 4, 9, 14 and 19) or 0 nmol (lanes 5, 10, 15 and 20) of the relevant Rpo13. Right, 'supershift' analysis of Rpo13–DNA complexes. Binding reactions containing 0.0164 nmol *Sac* Rpo13 (lanes 3 and 4) were incubated with 2 and 1 μl of antisera raised against Rpo13 (lanes 6 and 7) or, as a non-specific control, the cell division protein Vps4 (lanes 4 and 5). The reaction electrophoresed in lane 2 lacked Rpo13 but contained 2 μl of α-Rpo13 antisera. (**B**) Cartoon of wild-type and deletion derivatives of Rpo13. ΔN corresponds to residues 34–104 of the full-length protein, ΔC to residues 1–80 and Core to residues 34–80. 3 μg of each of the purified proteins were resolved on a 4–12% gradient NuPAGE gel and visualized by staining with Coomassie Brilliant Blue. (**C**) Upper panel, the N-terminal region of Rpo13 contributes to protein stability but is dispensable for DNA binding. Quantitation of EMSA experiments to determine the relative DNA binding activities of full length (triangle) and ΔN (rhombus) proteins. The experiments were performed at lower protein concentrations that those in the panel below because of the propensity of the ΔN protein to form insoluble aggregates at concentrations above 1.25 μM. Right, the basic C-terminal region of Rpo13 is critical for DNA binding. Lower panel, quantitation of EMSA experiments to determine the relative DNA binding activities of full length (triangle), ΔC (circle) and core (square) proteins. All experiments were performed in triplicate and error bars indicate the standard deviation.

specifically to other proteins, DNA or RNA, and a recent computational study has also revealed the presence and the relevance of IDRs in the archaeal proteome (25). Furthermore, current views on intrinsically disordered proteins (IDP) and IDRs, which are often characterized by an enrichment of charged residues, suggest that they may evolve at accelerated rates, providing enhanced diversity upon which selective processes can act (26,27). Significantly, it has also been proposed that they can modulate affinity in DNA–protein interactions (28).

In our DNA-bound *a*RNAP structure, the N- (E-rich) and C-terminal (K-rich) regions do not become ordered

upon interaction with the DNA, thus precluding a clear snapshot of direct contacts at the C-terminal end (there are however some indications at the N-terminus; see Supplementary Figure S3). Indeed, the non-sequence specific but discriminatory binding to dsDNA displayed by Rpo13 in EMSA experiments as well as its demonstrated molten-globule character provides a clear justification for this. It is known that lysines interact non-specifically with the DNA, a clear example being the lysine-rich tails of histones, physiologically relevant in stabilizing chromatin structure but whose promiscuous binding to DNA makes them unfavourable for crystallographic snapshots (29). It is enticing to note that this basic segment of Rpo13 possesses a lysine residues that is a target for methylation in *S. solfataricus* (30). Comparisons may also be drawn to polyamine polymers, such as spermine/spermidine, which function to stabilize the double helical structure of DNA but are invisible in co-crystal structures with DNA (31,32).

Taken together, our X-ray structures of DNA-bound and naked *a*RNAP structures, Rpo13's molten-globule nature manifesting in the inherent flexibility of the N- and the C-termini, the biophysical properties of Rpo13 and the *in vivo* ChIP data, implicate Rpo13 in both transcription initiation and elongation by interactions with dsDNA at the downstream end of the RNAP's cleft. The weak DNA binding affinity ($K_d \sim 2\,\mu M$) is compatible with Rpo13 not compromising transcriptional processivity but rather modulating the balance between non-specific and specific DNA interactions.

In all eukaryotes, the Rpb5 subunit is embellished with an additional N-terminal domain when compared with the archaeal Rpo5 and, as stated above, this eukaryote-specific domain acts as a 'guard rail' for DNA entering the enzyme (see Figure 3B). However, no archaeal species from any phylum possesses this domain and we have shown that its place in the Sulfolobales, and by inference the Desulfurococcales, Acidilobales and Korarchaea, is occupied by Rpo13 (Supplementary Figures S7 and S8). We speculate that either the low sequence complexity of Rpo13 has prevented its detection in archaea beyond these species or that other archaeal species have analogous rather than orthologous factors or subunits that occupy the predicted and conserved void between Rpo5 and Rpo1N. In this scenario, Rpo13 may represent a 'fossil', supplanted by the appearence of the Rpb5 N-terminal domain in eukaryotes.

ACCESSION NUMBERS

Protein Data Bank: Atomic coordinates and structure factors have been deposited under the accession numbers 4b1o and 4b1p (aRNAP-DNA) and 4ayb (apo-aRNAP).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–9, Supplementary Methods and Supplementary References [36–39].

ACKNOWLEDGEMENTS

We thank Yakov Korkhin for useful discussions, Marina Ondiviela for biotechnological assistance, Blanca Lopez for recording the NMR spectra and Rachel Samson for advice on the ChIP Seq experiments. We acknowledge the Swiss Light Synchrotron (SLS), the European Synchrotron Radiation Facility (ESRF) and Diamond Light Source for provision of synchrotron facilities. We also thank the staff at the beamlines PXIII at SLS, BM14 at ESRF and I24 at Diamond.

FUNDING

Basque Government [PI2010-20]; the Spanish Ministerio de Ciencia y Tecnología [BFU2009-08123]; CICbioGUNE (to M.N.W. and N.G.A.A.); Wellcome Trust and Edward Penley Abraham Trust (to S.D.B.); European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n.° 226716. Funding for open access charge: The Spanish Ministerio de Ciencia y Tecnología [BFU2009-08123].

Conflict of interest statement. None declared.

REFERENCES

1. Werner,F. and Grohmann,D. (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.*, **9**, 85–98.
2. Richard,D.J., Bolderson,E., Cubeddu,L., Wadsworth,R.I., Savage,K., Sharma,G.G., Nicolette,M.L., Tsvetanov,S., McIlwraith,M.J., Pandita,R.K. *et al.* (2008) Single-stranded DNA-binding protein hSSB1 is critical for genomic stability. *Nature*, **453**, 677–681.
3. Geiger,S.R., Lorenzen,K., Schreieck,A., Hanecker,P., Kostrewa,D., Heck,A.J. and Cramer,P. (2010) RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol Cell*, **39**, 583–594.
4. Werner,M., Thuriaux,P. and Soutourina,J. (2009) Structure-function analysis of RNA polymerases I and III. *Curr. Opin. Struct. Biol.*, **19**, 740–745.
5. Chen,Z.A., Jawhari,A., Fischer,L., Buchen,C., Tahir,S., Kamenski,T., Rasmussen,M., Lariviere,L., Bukowski-Wills,J.C., Nilges,M. *et al.* (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.*, **29**, 717–726.
6. Korkhin,Y., Unligil,U.M., Littlefield,O., Nelson,P.J., Stuart,D.I., Sigler,P.B., Bell,S.D. and Abrescia,N.G. (2009) Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol.*, **7**, e102.
7. Korkhin,Y., Littlefield,O., Nelson,P.J., Bell,S.D. and Sigler,P.B. (2001) Preparation of components of archaeal transcription preinitiation complex. *Methods Enzymol.*, **334**, 227–239.
8. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
9. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
10. Adams,P.D., Grosse-Kunstleve,R.W., Hung,L.W., Ioerger,T.R., McCoy,A.J., Moriarty,N.W., Read,R.J., Sacchettini,J.C., Sauter,N.K. and Terwilliger,T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1948–1954.
11. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.

12. Bell,S.D., Cairns,S.S., Robson,R.L. and Jackson,S.P. (1999) Transcriptional regulation of an archaeal operon in vivo and in vitro. *Mol. Cell*, **4**, 971–982.
13. Qureshi,S.A., Bell,S.D. and Jackson,S.P. (1997) Factor requirements for transcription in the Archaeon Sulfolobus shibatae. *EMBO J.*, **16**, 2927–2936.
14. Hudepohl,U., Reiter,W.D. and Zillig,W. (1990) In vitro transcription of two rRNA genes of the archaebacterium Sulfolobus sp. B12 indicates a factor requirement for specific initiation. *Proc. Natl Acad. Sci. USA*, **87**, 5851–5855.
15. Cheung,A.C., Sainsbury,S. and Cramer,P. (2011) Structural basis of initial RNA polymerase II transcription. *EMBO J.*, **30**, 4755–4763.
16. Kettenberger,H., Armache,K.J. and Cramer,P. (2003) Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell*, **114**, 347–357.
17. Liu,X., Bushnell,D.A., Silva,D.A., Huang,X. and Kornberg,R.D. (2011) Initiation complex structure and promoter proofreading. *Science*, **333**, 633–637.
18. Westover,K.D., Bushnell,D.A. and Kornberg,R.D. (2004) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science*, **303**, 1014–1016.
19. Hirata,A., Klein,B.J. and Murakami,K.S. (2008) The X-ray crystal structure of RNA polymerase from Archaea. *Nature*, **451**, 851–854.
20. Wojtas,M., Peralta,B., Ondiviela,M., Mogni,M., Bell,S.D. and Abrescia,N.G. (2011) Archaeal RNA polymerase: the influence of the protruding stalk in crystal packing and preliminary biophysical analysis of the Rpo13 subunit. *Biochem. Soc. Trans.*, **39**, 25–30.
21. Brock,T.D., Brock,K.M., Belly,R.T. and Weiss,R.L. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Microbiol.*, **84**, 54–68.
22. Dyson,H.J. and Wright,P.E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, **6**, 197–208.
23. Minezaki,Y., Homma,K., Kinjo,A.R. and Nishikawa,K. (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J. Mol. Biol.*, **359**, 1137–1149.
24. Turjanski,A.G., Gutkind,J.S., Best,R.B. and Hummer,G. (2008) Binding-induced folding of a natively unstructured transcription factor. *PLoS Comput. Biol.*, **4**, e1000060.
25. Xue,B., Williams,R.W., Oldfield,C.J., Dunker,A.K. and Uversky,V.N. (2010) Archaic chaos: intrinsically disordered proteins in Archaea. *BMC Syst. Biol.*, **4(Suppl. 1)**, S1.
26. Brown,C.J., Johnson,A.K., Dunker,A.K. and Daughdrill,G.W. (2011) Evolution and disorder. *Curr. Opin. Struct. Biol.*, **21**, 441–446.
27. Schlessinger,A., Schaefer,C., Vicedo,E., Schmidberger,M., Punta,M. and Rost,B. (2011) Protein disorder–a breakthrough invention of evolution? *Curr. Opin. Struct. Biol.*, **21**, 412–418.
28. Vuzman,D. and Levy,Y. (2012) Intrinsically disordered regions as affinity tuners in protein-DNA interactions. *Mol. Biosyst.*, **8**, 47–57.
29. Ramakrishnan,V., Finch,J.T., Graziano,V., Lee,P.L. and Sweet,R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219–223.
30. Botting,C.H., Talbot,P., Paytubi,S. and White,M.F. (2010) Extensive lysine methylation in hyperthermophilic crenarchaea: potential implications for protein stability and recombinant enzymes. *Archaea*, **2010**, Article ID 106341, 6 pages, doi:10.1155/2010/106341.
31. Korolev,N., Lyubartsev,A.P., Nordenskiold,L. and Laaksonen,A. (2001) Spermine: an "invisible" component in the crystals of B-DNA. A grand canonical Monte Carlo and molecular dynamics simulation study. *J. Mol. Biol.*, **308**, 907–917.
32. Venkiteswaran,S., Thomas,T. and Thomas,T.J. (2006) Role of polyamines in regulation of sequence-specific DNA binding activity. In: Wang,J.-Y. and Casero,R.A. (eds), *Polyamine Cell Signaling, Physiology, Pharmacology and Cancer Research*. Humana Press, Totowa, NJ, pp. 91–122.
33. Stuart,D.I., Levine,M., Muirhead,H. and Stammers,D.K. (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 A. *J. Mol. Biol.*, **134**, 109–142.
34. Whitmore,L. and Wallace,B.A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*, **89**, 392–400.
35. Wurtzel,O., Sapra,R., Chen,F., Zhu,Y., Simmons,B.A. and Sorek,R. (2010) A single-base resolution map of an archaeal transcriptome. *Genome Res.*, **20**, 133–141.
36. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
37. Berrow,N.S., Alderton,D., Sainsbury,S., Nettleship,J., Assenberg,R., Rahman,N., Stuart,D.I. and Owens,R.J. (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.*, **35**, e45.
38. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
39. Gouet,P., Courcelle,E., Stuart,D.I. and Metoz,F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.