Figures and figure supplements
Structure of a bacterial RNA polymerase holoenzyme open promoter complex

Brian Bae, et al.
Figure 1. Structure of RPo. (A) Oligonucleotides used for RPo crystallization. The numbers above denote the DNA position with respect to the transcription start site (+1). The DNA sequence is derived from the full con promoter (Gaal et al., 2001). The $-35$ and $-10$ (Pribnow box) elements are shaded yellow, the extended $-10$ (Keilty and Rosenberg, 1987) and discriminator (Feklistov et al., 2006; Haugen et al., 2006) elements purple. The nt-strand DNA (top strand) is colored dark grey; t-strand DNA (bottom strand), light grey; RNA transcript, red. (B) Overall structure of RPo. The nucleic acids are shown as CPK spheres and color-coded as above. The Taq $\Delta 1.1\sigma A$ is shown as a molecular surface ($\alpha I$, $\alpha II$, $\omega$, grey; $\beta$, light cyan; $\beta'$, light pink; $\Delta 1.1\sigma A$, light orange), transparent to reveal the RNAP active site $\text{Mg}^{2+}$ (yellow sphere) and the nucleic acids held inside the RNAP active site channel. (C) Electron density and model for RPo nucleic acids. Blue mesh, $2F_o - F_c$ maps for nucleic acids (contoured at 0.7$\sigma$).

DOI: 10.7554/eLife.08504.003
Figure 1—figure supplement 1. (Left) Synthetic oligonucleotides used for us-fork (−12 bp) crystallization. The numbers above the sequence denote the DNA position with respect to the transcription start site (+1). The DNA sequence is derived from the full con promoter (Gaal et al., 2001). The −35 and −10 (Pribnow box) elements are shaded yellow, the extended −10 (Keilty and Rosenberg, 1987) and discriminator (Feklistov et al., 2006; Haugen et al., 2006) elements purple. The nt-strand DNA (top strand) is colored dark grey; the t-strand DNA (bottom strand), light grey. (Right) Electron density and model for nucleic acids. Blue mesh, $2F_o - F_c$ maps for nucleic acids (contoured at 0.7σ). DOI: 10.7554/eLife.08504.004
Figure 1—figure supplement 2. Data and model quality for us-fork (−12 bp) and RPo complexes. Plots relating data quality with model quality at 4.0 Å-resolution (using the Pearson correlation coefficient (CC) analysis described by Karplus and Diederichs (2012). CC1/2 (red squares) was determined from the unmerged diffraction data randomly divided in half. Since CC1/2 underestimates the information content of the data (since it’s calculated by dividing the dataset in half), CC* was calculated from an analytical relation to estimate the information content of the full data (Karplus and Diederichs, 2012). CC* provides a statistic that assesses data quality as well and also allows direct comparison of crystallographic model quality and data quality on the same scale through CC_work and CC_free, the standard and cross-validated correlations of the experimental intensities with the intensities calculated from the refined model. A CC_work/CC_free smaller than CC* indicates that the model does not account for all of the signal in the data, meaning it is not overfit. Plotted also are the standard <I>/σ_I for the diffraction data, as well as the R_work/R_free for the refined models. (Left) Data for Taq Δ1.1σ_A/us-fork (−12 bp) at 4.0 Å-resolution. (Right) Data for Taq Δ1.1σ_A RPo (with 4-nt RNA primer) at 4.0 Å-resolution. DOI: 10.7554/eLife.08504.005
Figure 1—figure supplement 3  Sequence alignment of regions 2–4 of selected bacterial RNAP primary (Group I) σ subunits. Sequence alignment of regions 2–4 (Lonetto et al., 1992) of bacterial RNAP primary (Group I) σ subunits. The sequences shown were selected from diverse phyla/groups taken from a much larger alignment of 1002 sequences. The σA sequences shown are from the following organisms chosen to represent the preceding phylum/group: Deinococcus-Thermus, Thermus aquaticus; γ-Proteobacteria, Escherichia coli; Actinobacteria, Mycobacterium tuberculosis; Acidobacteria-Candidatus Solibacter usitatus; α-Proteobacteria, Rickettsia bellii; Aquificae, Desulfurobacterium thermolithotrophum; Chlamydae, Chlamydae trachomatis; Cyanobacteria, Mastigocoleus testarum; δ-Proteobacteria; Desulfobulbus propionicus; Firmicutes, Bacillus cereus; Spirochaetes, Treponema pallidum; Thermodesulfobacteria, Thermodesulfatator atlanticus. The sequences are shaded according to conservation within the sub-alignment; red shading indicates 100% identity, blue shading indicates >50% identity. The histogram at the top represents the sequence conservation within the entire 1002 sequence alignment (red bar, 100% identity; orange, 83–99%; green, 68–83%; blue, 50–67%).

DOI: 10.7554/eLife.08504.006
Figure 2. Protein interactions with duplex DNA upstream of the transcription bubble and recognition of the extended −10 element. (A) (Left) Overall view of RPo structure (similar to Figure 1B). The boxed area is magnified on the right. (Right) Magnified view showing protein interactions (β' and σA) with duplex DNA from −18 to −12. Proteins are shown as backbone worms (β', light pink; σA, light orange) with interacting side chains shown in stick format (β', pink; σA, orange). Likely polar interactions are denoted with grey dashed lines. (B) Same as (A) (right) but rotated 180˚ about the x-axis. (C) Similar view as (A) (right). Superimposed is the simulated annealing omit map (grey mesh, 2Fo – Fc, contoured at 1σ), calculated from a model where the following protein segments were removed (β' 33–36; σA 259–292) and shown only within 2 Å of omitted atoms. (D) Similar view as (B). Superimposed is the simulated annealing omit map (grey mesh, 2Fo – Fc, contoured at 1σ), calculated from a model where the following protein segments were removed (β' 33–36; σA 259–292) and shown only within 2 Å of omitted atoms. DOI: 10.7554/eLife.08504.010
Figure 3. Structures maintaining the upstream ds/ss junction of the transcription bubble and directing the t-strand DNA to the RNAP active site. (A) (Left) Overall view of RPo structure (similar to Figure 1B). The boxed area is magnified on the right. (Right) Magnified view showing the upstream ds/ss junction of the transcription bubble in RPo (the RNAP β subunit, which obscures the view, has been removed). RNAP is shown as a molecular surface, except side chains of key σA residues (R217, R220, W256, R288, R291) are shown (orange). The orthogonal directions of the ss nt- and t-strand DNA following the upstream ds/ss junction are denoted by black arrows. The dashed, curved line denotes the potential path of the t-strand −11 base from its position in the duplex DNA (base-paired to A−11(nt)) to its position in the structure. (B) Structure of Taq σA2 bound to the ss, nt-strand −10 element.
(PDB ID 3UGO) (Feklistov and Darst, 2011) showing the disposition of the universally conserved σ^70 W-dyad (Tag σ^70 W256/W257). Shown is the ss DNA from −14 to −7 (−10 element colored yellow), the σ^70 α helix (light orange) and the W-dyad (orange side chains with transparent CPK atoms). W256 makes an edge-on interaction with the face of W257, as observed in all other σ^70/σ^A structures in many different contexts (Malhotra et al., 1996; Campbell et al., 2002; Vassylyev et al., 2002; Murakami et al., 2002a, 2002b; Feklistov and Darst, 2011; Zhang et al., 2012). (C) Disposition of the W-dyad in RPo (containing upstream ds/ss junction, shown schematically above). Only the nt-strand DNA from −14 to −7, the σ^70 α helix, and the W-dyad are shown (as in B). (D) Same view as (C). Superimposed is the simulated annealing omit map (grey mesh, 2F_o − F_c, contoured at 1σ), calculated from a model where the following segments of σ^70 were completely removed (216–221, 255–258, and 287–292) and shown only within 2 Å of omitted atoms. (E) Similar view as (A) (right). Superimposed is the simulated annealing omit map (grey mesh, 2F_o − F_c, contoured at 1σ), calculated from a model where the following segments of σ^70 were removed (216–221, 255–258, and 287–292) and shown only within 2 Å of omitted atoms. Clear Fourier density for σ^A Y217 and R288 is shown.

DOI: 10.7554/eLife.08504.011

Figure 3—figure supplement 1. Stereo view of RPo model and resulting electron density map (grey mesh, 2F_o − F_c, contoured at 0.7σ). The view is similar to Figure 3A.

DOI: 10.7554/eLife.08504.012
Figure 4. The $\sigma^*$ Y217 may stack on the $T_{-11(t)}$ base orphaned by the flipped out $A_{-11(nt)}$ base. (A) Synthetic oligonucleotides used for us-fork ($-11$ bp) crystallization. The numbers above the sequence denote the DNA position with respect to the transcription start site ($+1$). The DNA sequence is derived from the full con promoter (Gaal et al., 2001). The $-35$ and $-10$ (Pribnow box) elements are shaded yellow, the extended $-10$ element (Keilty and Rosenberg, 1987) purple. The nt-strand DNA (top strand) is colored dark grey; the t-strand DNA (bottom strand), light grey; the RNA transcript, red. (B) The $T_{-11(t)}$ base orphaned by the flipped out $A_{-11(nt)}$ stacks on $\sigma^*$ Y217 in the us-fork ($-11$ bp) structure. The 4.6 Å-resolution electron density map (contoured at $0.7\sigma$) is shown (grey mesh). Also superimposed is the simulated annealing omit map (green mesh, $F_o - F_c$ contoured at $3\sigma$), calculated from a model where $\sigma^*$ Y217 was mutated to Ala and the $T_{-11(t)}$ nucleotide was deleted.

DOI: 10.7554/eLife.08504.013
Data and model quality for us-fork (-11 bp) complex. Plots relating data quality with model quality at 4.6 Å-resolution using the Pearson correlation coefficient (CC) analysis described by Karplus and Diederichs (2012). CC1/2 (red squares) was determined from the unmerged diffraction data randomly divided in half. Since CC1/2 underestimates the information content of the data (since it’s calculated by dividing the dataset in half), CC* was calculated from an analytical relation to estimate the information content of the full data (Karplus and Diederichs, 2012). CC* provides a statistic that assesses data quality as well and also allows direct comparison of crystallographic model quality and data quality on the same scale through CC_work and CC_free, the standard and cross-validated correlations of the experimental intensities with the intensities calculated from the refined model. A CC_work/CC_free smaller than CC* indicates that the model does not account for all of the signal in the data, meaning it is not overfit. Plotted also are the standard <I>/<σ for the diffraction data, as well as the R_work/R_free for the refined models.

DOI: 10.7554/eLife.08504.014
Figure 5. Functional role of Eco σ70 W433 and Y394 in RPo formation. (A) Simplified, two-step kinetic scheme for RPo formation (Roe et al., 1984; Buc and McClure, 1985) (R, RNAP; P, promoter; RPi, intermediate complex). (B) Representative time trace of fluorescence increase (from Cy3 labelled promoter DNA) during RPo formation. The solid red line illustrates the non-linear regression fit to a single-exponential model (see 'Materials and methods'), which described >90% of the fluorescence amplitude rise. (C) The RNAP-concentration dependence of the observed rate ($k_{\text{obs}}$) of RPo formation detected by Cy3 fluorescence (Ko and Heyduk, 2014) for Eco holoenzymes with σ70 (wt) as well as σ70 carrying substitutions W433A or Y394A. Error bars denote standard errors of the mean for ≥three independent measurements. (D) Summary of effects of σ70 W433A and Y394A substitutions on thermodynamic and kinetic parameters of RPo formation. The data was normalized to the % observed with wild-type Ecoσ70. (E) Equilibrium binding of ss nt-strand oligos of λ Pr promoter –10 element detected in the fluorescent RNAP beacon assay (Feklistov and Darst, 2011; Mekler et al., 2011) to Eco holoenzymes with σ70, as well as σ70 carrying substitutions W433A or Y394A. DOI: 10.7554/eLife.08504.015
Figure 5—figure supplement 1. RPo dissociation data. (Left) Representative time trace of fluorescence decay after rapid mixing of pre-formed Eco RPo (with wild-type σ70) into 1.1 M NaCl (Gries et al., 2010). The solid line illustrates the non-linear regression fit to a single-exponential model. (Right) Representative dissociation curves for holoenzymes containing wild-type σ70 and W433A and Y394A substitutions.

DOI: 10.7554/eLife.08504.016
Figure 6. Structural role of the $\sigma_{3.2}$-loop. (Left) Overall view of RPo structure, colored as in Figure 1 except $\sigma^A$ is orange. The RNAP $\beta$ and $\beta'$ subunits are transparent to reveal the RNAP active site $\text{Mg}^{2+}$ (yellow sphere) and the nucleic acids held inside the RNAP active site channel. The ss nt-strand DNA is omitted for clarity. The boxed area is magnified on the right. (Right) Magnified view showing a cross-section of the RNAP active site channel. For clarity, the RNAP $\beta$, $\beta'$, and $\sigma_{2}^{A}$ domains are shown mostly as outlined shapes, with $\beta$ transparent. The ss t-strand DNA ($-11$ to $-4$) is directed towards the RNAP active site through a tunnel between the $\sigma_{3.2}$-loop and the $\beta$-lobe. The 4-nt RNA transcript ($-3$ to $+1$) contacts the distal tip of the $\sigma_{3.2}$-loop. Further elongation of the RNA would require displacement of the $\sigma_{3.2}$-loop.

DOI: 10.7554/eLife.08504.017