σ_{NCR} facilitates DNA opening at the transcription start site: Supplemental Information

Cryo-EM structure of *Escherichia coli* σ^{70} RNAP and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation

Anoop Narayanan[1,2,*], Frank S. Vago[2,*], Kungpeng Li[2,*], M. Zuhaib Qayyum[1], Dinesh Yernool[2], Wen Jiang[2,3] and Katsuhiko S. Murakami[1,3]

From the [1]Department of Biochemistry and Molecular Biology, The Center for RNA Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA, the [2]Department of Biological Sciences, Markey Center for Structural Biology, Purdue University, West Lafayette, IN 47906, USA
*contributed equally to this work

Running title: σ_{NCR} facilitates DNA opening at the transcription start site

[3]To whom correspondence should be addressed: Katsuhiko Murakami: Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802; kum14@psu.edu; Tel. (814) 865-2758 or Wen Jiang: [2]Department of Biological Sciences, Purdue University, West Lafayette, IN 47906; jiang12@purdue.edu; Tel. (765) 496-8436.

Supporting Information:

Materials included: Figures S1 – S5, Movie S1

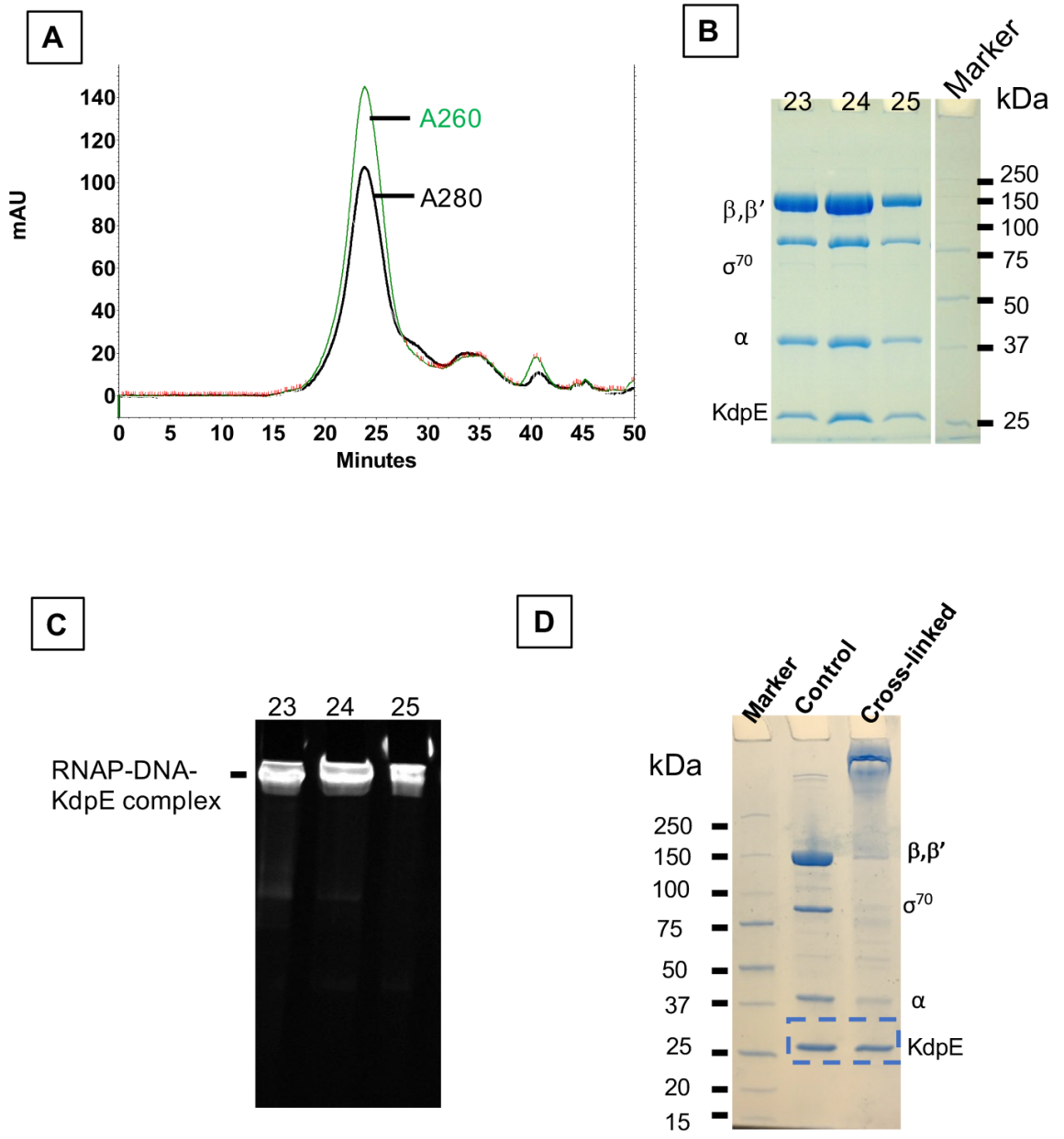

Figure S1. Preparation of the RNAP-pFABC-1-KdpE-E216A ternary complex for cryo-EM. ***A***, Size exclusion chromatography (SEC). Absorption profiles at wavelengths 260 nm and 280 nm of the assembled ternary complex separated on a Superose6 10/300 GL column (GE Healthcare). ***B***, SDS-PAGE analysis of the SEC peak fractions (23, 24 and 25) showing co-elution of RNAP subunits and KdpE protein. ***C***, Analysis of the SEC fractions (23, 24 and 25) by EMSA. The bands represent a stable ternary complex. ***D***, SDS-PAGE of purified ternary complex before (lane 1) and after cross-linking with Glutaraldehyde (lane 2). In lane 2, the top band represents cross-linked complex of RNAP subunits. Band corresponding to free KdpE protein is indicated by the rectangle with dashed border.

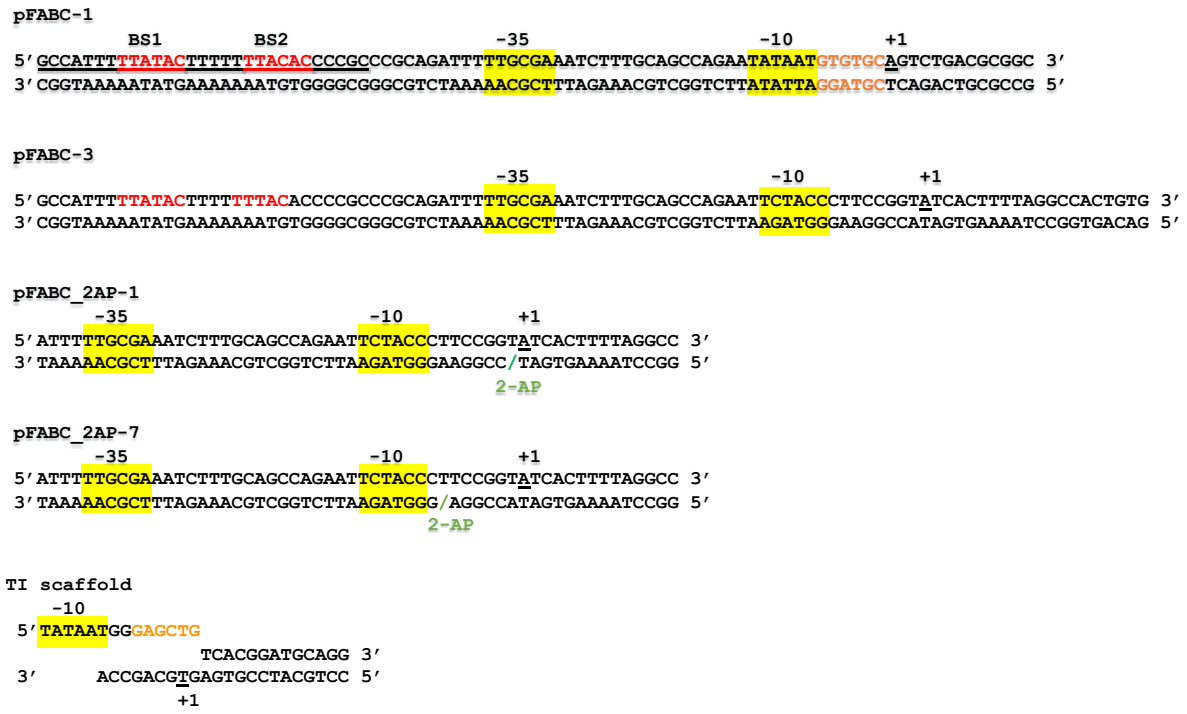

Figure S2. DNA Sequences of promoter regions used in this study.

pFABC-1: Sequence of the promoter DNA used to assemble ternary complex. Promoter regions corresponding to tandem KdpE binding sites (BS1, BS2, red letters), -35 and -10 elements (yellow boxes), non-complemental DNA sequence to form artificial bubble (orange letters), and transcription start site (+1) are indicated. Underlined sequence at the upstream DNA indicates the 29 bases missing in the cryo-EM map of RPo.

pFABC-3: Sequence of the promoter DNA used for transcription assay and EMSA (**Figs. 3A and C**). KdpE binding sites (BS1, BS2, red letters), -35 and -10 elements (yellow boxes), and transcription start site (+1) are indicated. This DNA has no artificial transcription bubble.

pFABC_2AP-1 and pFABC_2AP-7: Sequences of promoter DNA used for 2-amino purine fluorescence assay (**Fig. 3D**). -35 and -10 elements (yellow boxes) and the position of 2- amino purine substitution (green slash) are shown. These DNA have no artificial transcription bubble.

TI scaffold: Sequence of the promoter DNA with pre-melted bubble used for transcription assay (**Fig. 3F**). -10 element (yellow box) and non-complemental DNA sequence to form artificial transcription bubble (orange letters) are indicated.

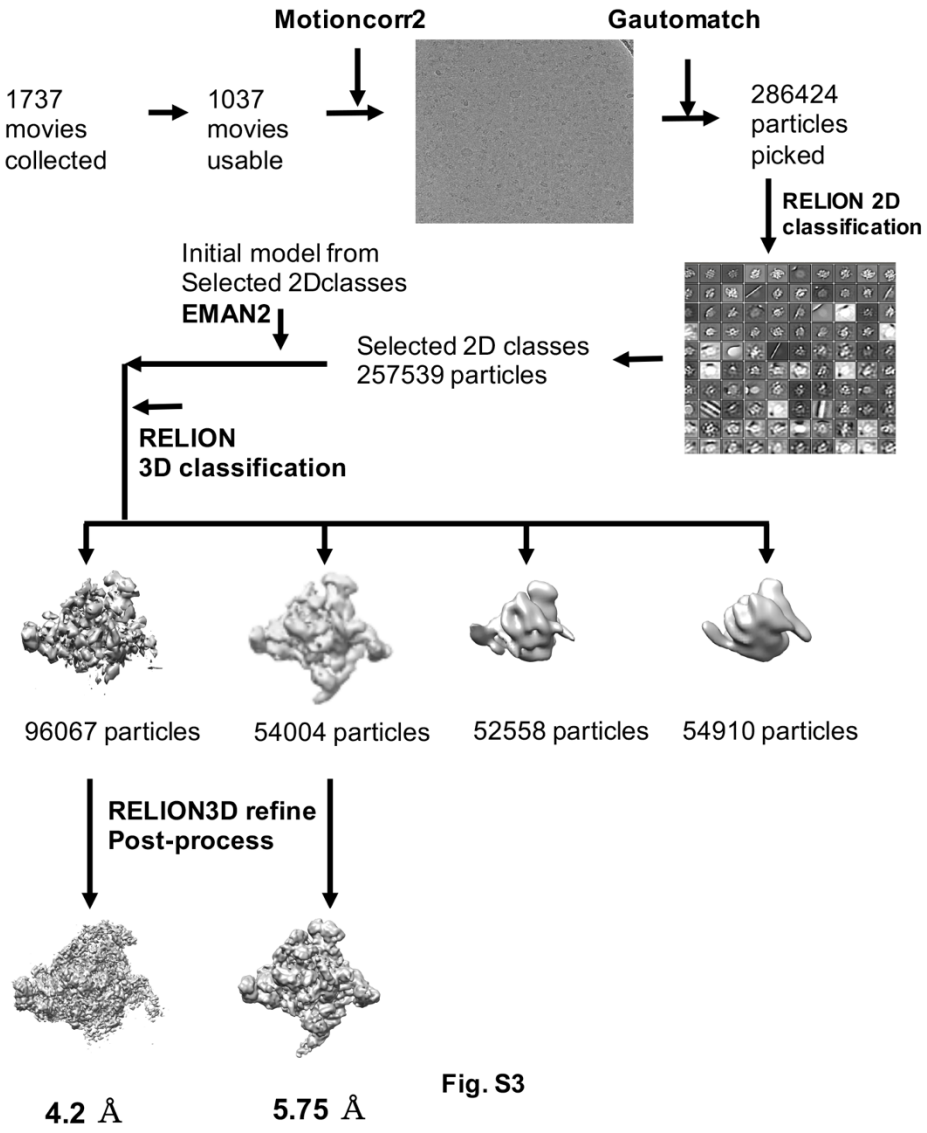

Figure S3. Cryo-EM data processing flow chart. From 1037 selected movies, 286424 particles were selected and subjected to 2D classification using RELION 2 (beta). After removing the bad classes, 257539 particles were used for subsequent 3D classification. The best two 3D classes were refined to obtain cryo-EM maps at 4.2 Å (96067 particles) and 5.7 Å (54004 particles) as indicated. Methodology is described in materials and methods section.

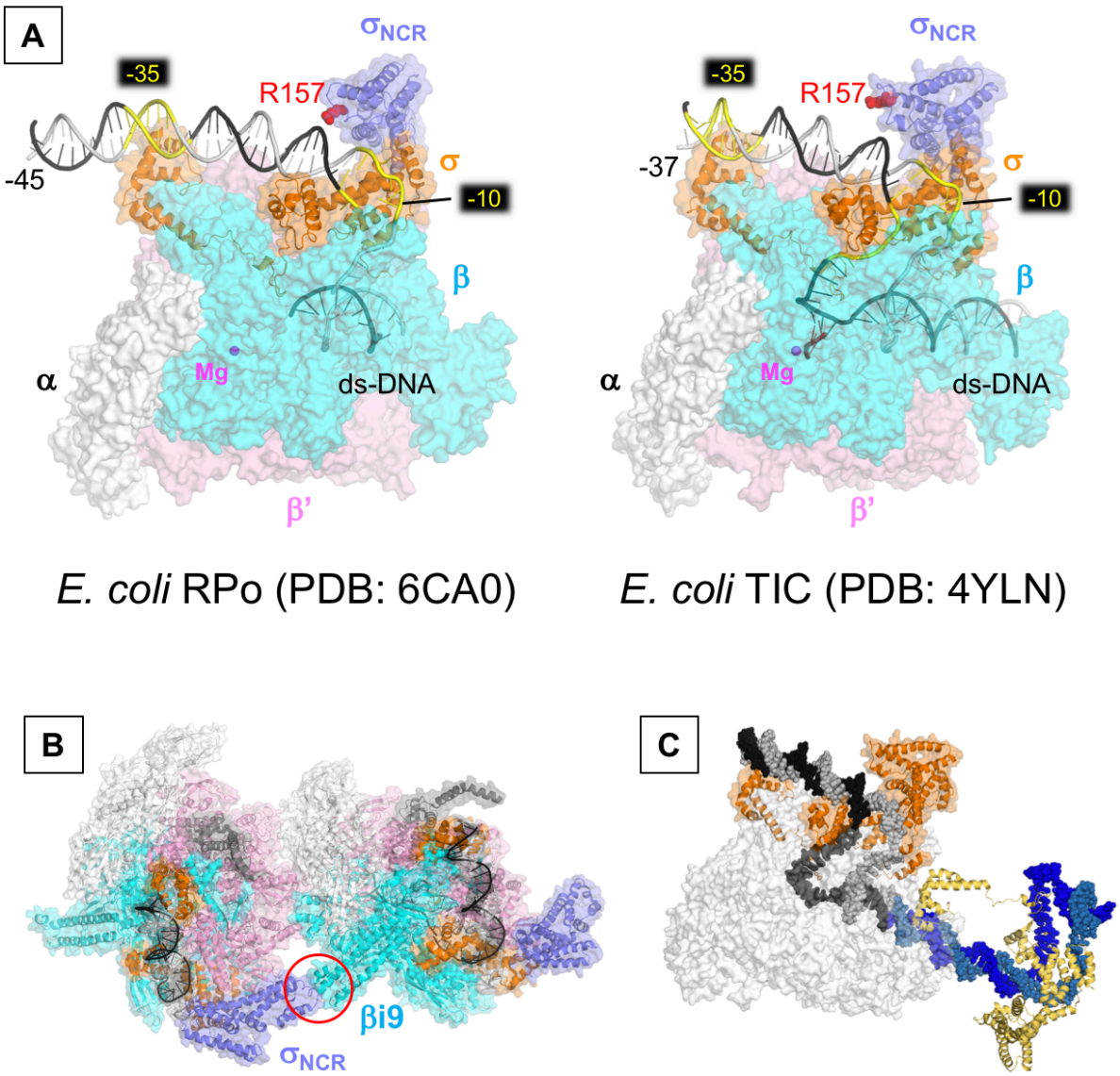

Figure S4. Structure comparison of the cryo-EM structure of RPo and the X-ray crystal structure of TIC (PDB: 4YLN). ***A***, Molecular surface representations of the cryo-EM structure of RPo (left) and the X-ray crystal structure of TIC (right) shown with ribbon representations of their respective promoter DNA and σ^{70} domain. RNAP subunits and DNA strands are colored as; α -white, β -cyan, β'-magenta, σ^{70} -orange, σ_{NCR} –purple, t-DNA- dark gray, nt-DNA-light grey, -35 and -10 elements (yellow). R157 residue of σ_{NCR} is shown as a sphere model and colored red. Active site Mg^{+} ion is indicated as a sphere. ***B***, One of the crystal packings of the TIC crystal is shown. Contact between the σ_{NCR} and the βi9 of the symmetry related molecule is highlighted by a circle. ***C***, Head-to-tail binding of downstream DNA with upstream DNA of symmetry related molecule in the crystal structure of TIC. One molecule of RNAP is shown as transparent molecular surface model with its σ^{70} as ribbon (orange) and t-DNA (dark gray) and nt-DNA (light gray) shown as sphere models. Ribbon model of σ^{70} (yellow), and sphere models of t-DNA (dark blue) and nt-DNA (light blue) of the symmetry related molecule are shown.

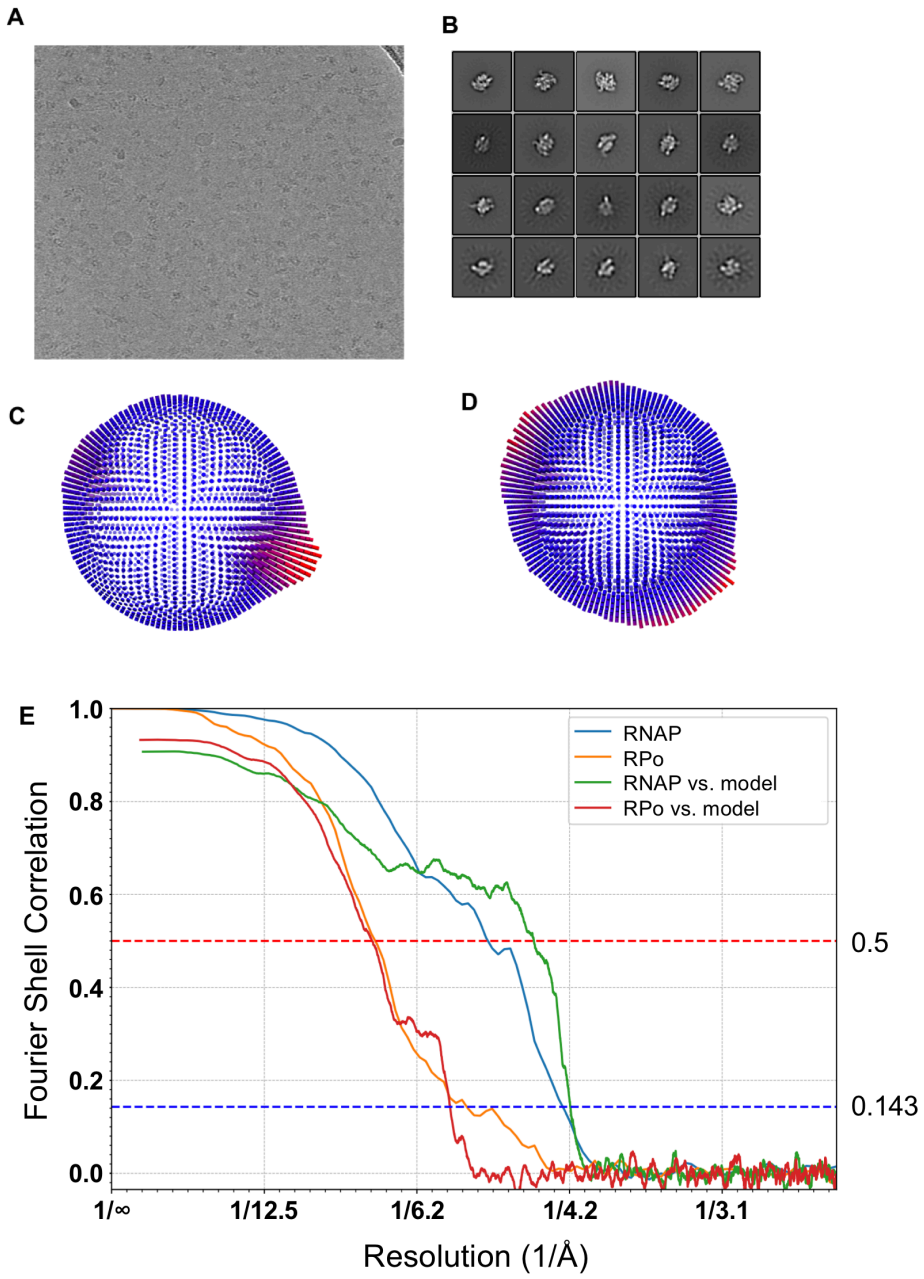

Figure S5. Cryo-EM of RNAP and RPo. ***A***, Representative motion corrected raw micrograph of frozen particles. ***B***, representative 2-D classes with particles in different views. ***C***, Euler angular distribution of the particles for 3D class representing the apo-form RNAP holoenzyme. ***D***, Euler angular distribution of the particles for 3D class representing the RPo. ***E***, Fourier Shell Correlation curves with dotted line representing 0.143 FSC cutoff (for half-maps, blue and orange lines) and 0.5 FSC cutoff (for the final model vs map, green and red lines).

Movie S1. Conformational change of the β'i6 domain (related to Fig. 2D).
This movie shows the conformational change of the β'i6 domain during the stable RPo formation, starting from the cryo-EM structure of RPo (this study) having the open-conformation of the β'i6 domain to the X-ray structure of TIC having the closed-conformation of the β'i6 domain. The movement of the β'i6 domain toward the βlobe/i4 domain results in closure of the gap between the β'jaw and βlobe/i4 domains.