Structure, Volume *21*

Supplemental Information

Structure of the *Mtb* CarD/RNAP β-Lobes Complex Reveals the Molecular Basis of Interaction and Presents a Distinct DNA-Binding Domain for *Mtb* CarD

Gulcin Gulten and James C. Sacchettini

Inventory of Supplemental Information

Supplemental Data

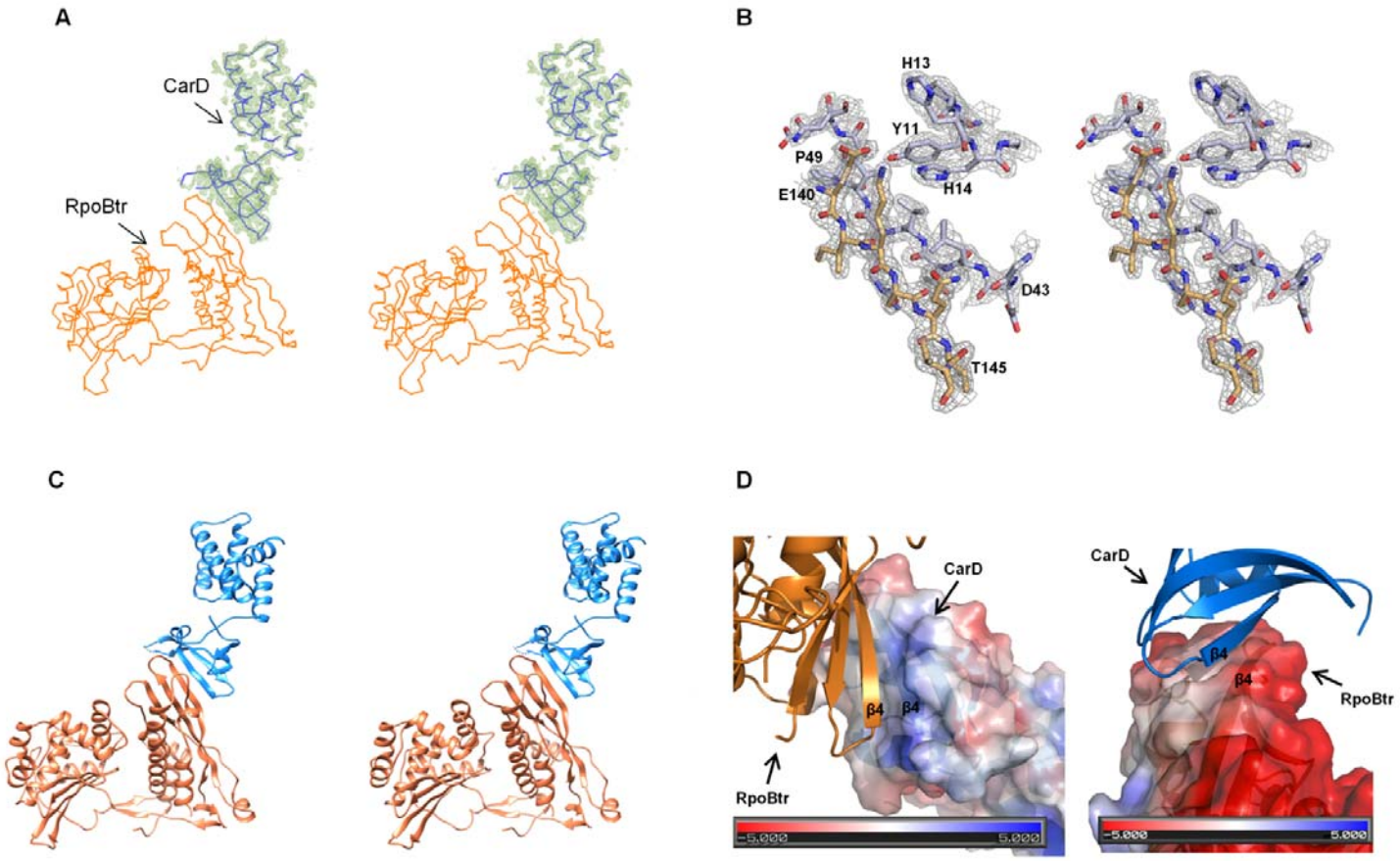

Figure S2, related to Figure 1. **Stereo images of the CarD/RNAP complex and electrostatic potential surface representation of the CarD/RNAP interface.** (A) The |Fo|-|Fc| density (shown at 1.8 sigma in green) for CarD at the protein-protein interface was unambiguous after locating RpoBtr in the complex structure by MR. C_α backbone of the final model of the complex is included for clarity. (B) The 2|Fo|-|Fc| map at 2.0 sigma of the CarD/RNAP interface after refinement. RpoBtr residues are shown in orange; CarD residues are shown in gray sticks. Images were prepared by Pymol. (C) Wall-eyed stereo image of the CarD/RNAP complex in cartoon representation. (D) CarD displays a positively charged surface (left), and RNAP has a negatively charged surface (right) at the interface. Electrostatic potential molecular surface calculations were done and images were prepared with PyMol using APBS as the

macromolecular electrostatics calculation program. In all panels, RpoBtr is shown in orange, CarD is shown in blue ribbons.

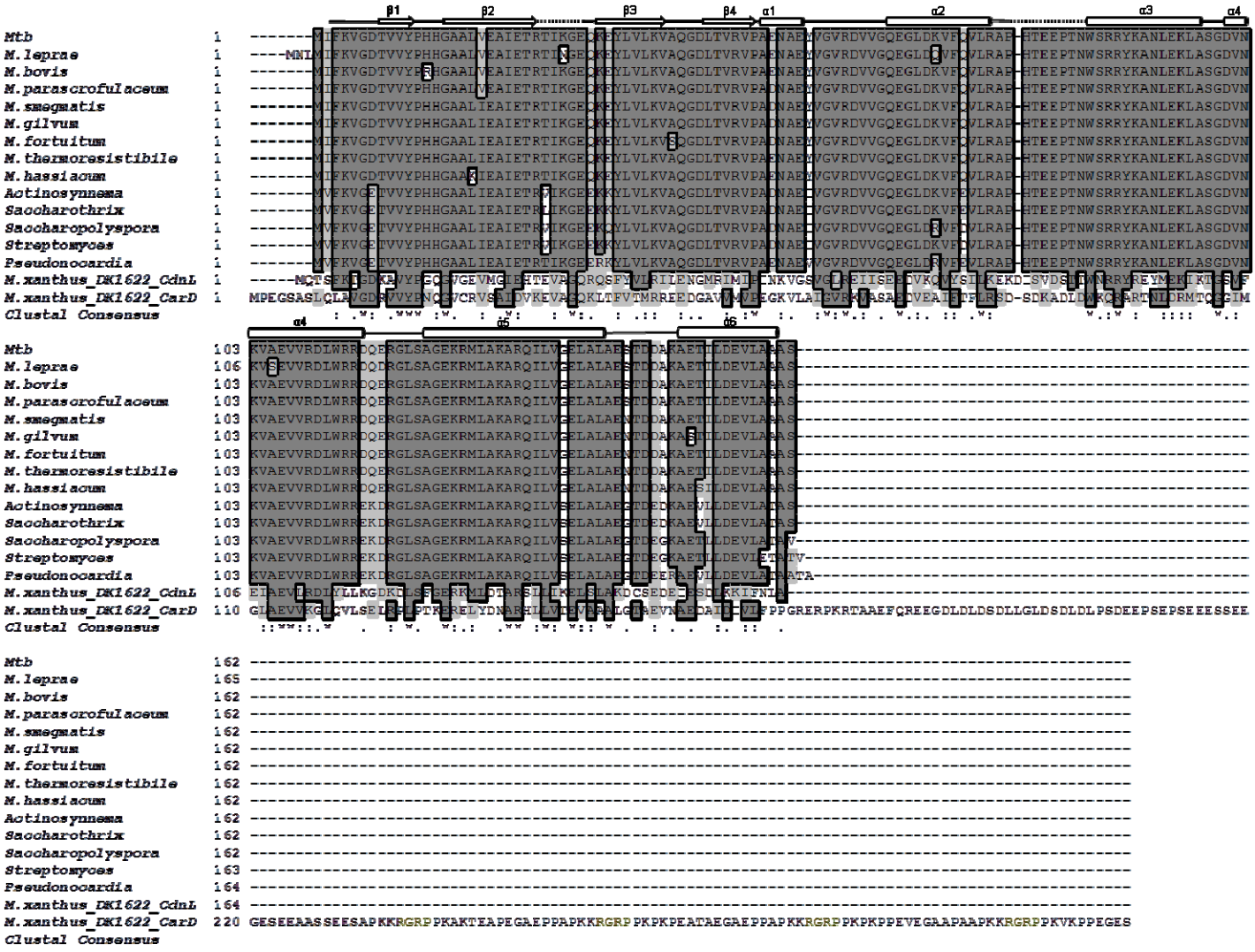

Figure S1, related to Figure 2. Protein sequence alignment of *Mtb* CarD. *Mtb* CarD sequence is aligned with the closest homologs (determined by the BLAST search) and *M. xanthus* CarD and CdnL proteins. Identical residues are colored dark gray and boxed, similar residues are shaded with light gray. *M. xanthus* CarD has additional 135 residues at the C-terminus compared to *Mtb* CarD. The AT-hook DNA binding motif (four 'RGRP' motifs) of *M. xanthus* CarD are shaded in yellow boxes. The secondary structure of *Mtb* CarD is shown above its sequence. The missing residues in the final model are indicted by dashed lines. Alignment is generated by ClustalW.

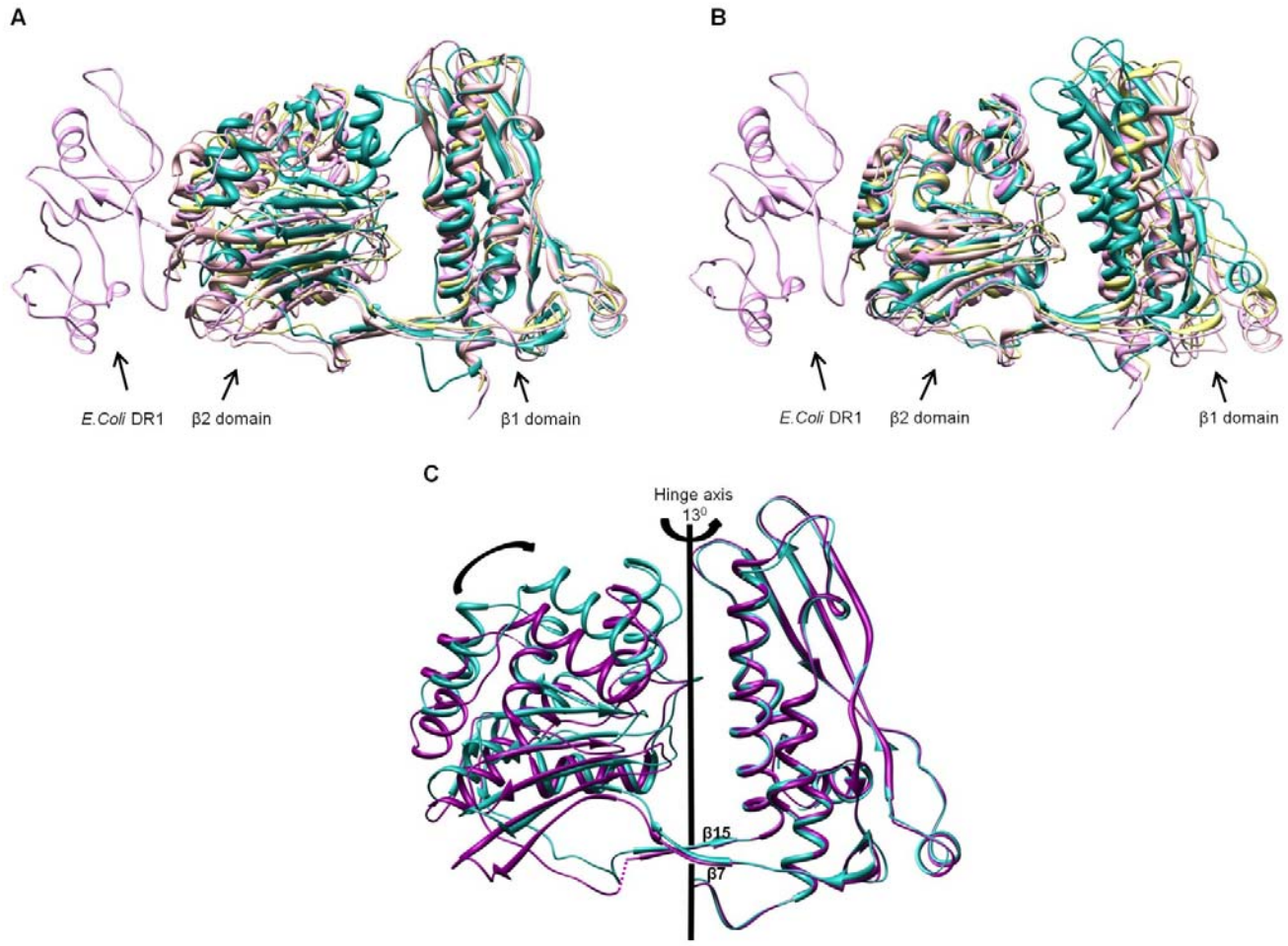

Figure S3, related to Figure 3. Comparison of the uncomplexed *Mtb* β1 and β2 domain structure (green) with the *Tth* (dark pink, PDB ID: 2O5I), *Taq* (yellow, PDB ID: 1I6V) and *E. coli* (light pink, PDB ID: 4IGC) RNAP β1 and β2 domain structures. Rest of the *Tth*, *Taq* and *E. coli* RNAP structures is omitted for clarity. Superposition of the β1 (A) and β2 (B) domains results ~10 Å RMSD over the C_α atoms of the other domain, which indicates that the relative conformation adopted by the two domains in RpoBtr is different than other bacterial core or holo RNAP structures. (C) Superposition of the uncomplexed RpoBtr β1 domains of the two molecules in the ASU (RpoBtr_A (magenta) and RpoBtr_B (green)) gives an RMSD of 5.2 Å (over 191 atom pairs) on the C_αs of the β2 domains. The conformational difference observed in the β1-β2 domain-domain orientation can be explained by a 13º rotation around the hinge axis centered on the two stranded anti-parallel β-sheet connecting the two domains.

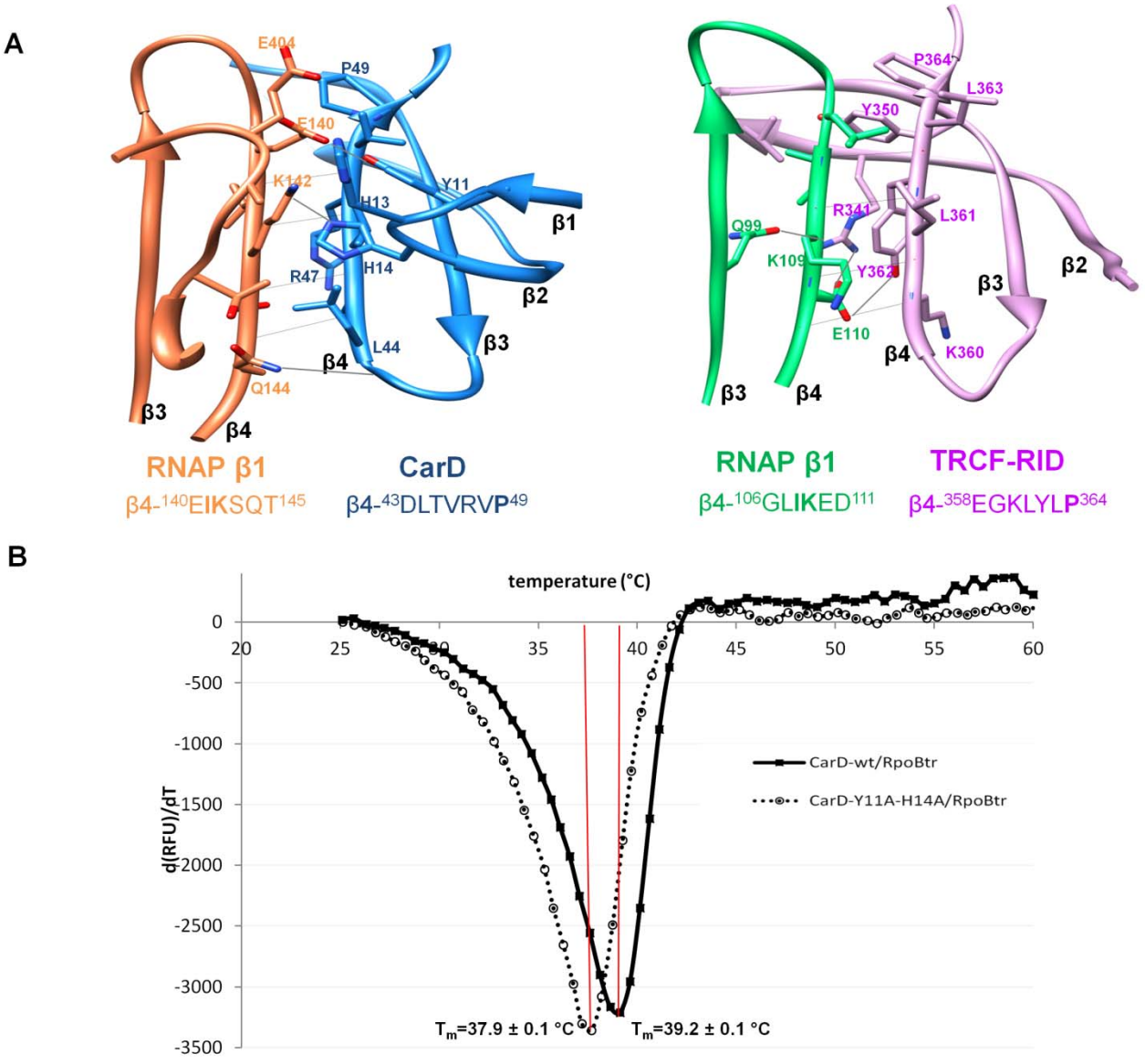

Figure S4, related to Table 2. Comparison of CarD/RNAP interaction with TRCF-RID/RNAP interaction and the DSF experiments on CarD/RpoBtr and CarD-Y11A-H14A/RpoBtr complexes. (A) The CarD/RNAP interface (this study, left panel) and the TRCF-RID/RNAP interface (PDB ID: 3MLQ, right panel) are displayed. The sequence of the β4-strands is given below each chain with conserved residues in bold. Important residues for intermolecular interaction are labeled, and H-bonds are shown with gray lines. (B) Differential scanning fluorimetry (DSF) experiments on CarD/RpoBtr (filled squares, solid

line) and CarD-Y11A-H14A/RpoBtr (open circles, dashed line) complexes. Equal amount of each complex (1 μM) in 200 mM Tris pH 7.5, 100 mM NaCl buffer was mixed with 5X Sypro orange dye. The temperature of the samples was changed from 25 to 95 °C at a heating rate of 0.5 °C/min, and the fluorescence was monitored. The first derivative of fluorescence d(RFU)/dT is plotted against temperature and the melting point (Tm, indicated with red lines) was calculated as the lowest point of the first derivative plot. Experiments were repeated five times and standard errors are calculated.

Table S1, related to Experimental Procedures. Primers that are used in this work.

Construct	Forward 5'→ 3'	Reverse 5'→ 3'
CarD-notag	GGGAAT TCCATATGATTT TCAAGGTCGGAGACACCGTTGTC	GATCCCAAGCTTTCAAGACGCGGCGGCTAAAAC CTCGTCAAG
CarD-Nterm-1-53 ($CarD_{1-53}$)	GGGAATTCCATATGATTTTCAAGGTCGGAGACACCGTTGTC	CCCAAGCTTTTAGGCGTTTTCAGCGGGAACTCG TA
CarD-Nterm-1-74 ($CarD_{1-74}$)	GGGAATTCCATATGATTTTCAAGGTCGGAGACACCGTTGTC	CCCAAGCTTTTACAACACCTGGAAAACCTTGTC CAGG
CarD-Cterm-61-162 ($CarD_{61-162}$)	AGAGAAGCATATGATGGTCGTCGGGCAGGAAGGCCTGG	GATCCCAAGCTTTCAAGACGCGGCGGCTAAAAC CTCGTCAAG
CarD-Cterm-83-162 ($CarD_{83-162}$)	AGAGAAGCATATGATGACGAACTGGTCACGTCGTTACAAGGC	GATCCCAAGCTTTCAAGACGCGGCGGCTAAAAC CTCGTCAAG
CarD_R87A-R88A-K90A	ACCGAGGAGCCGACGAACTGGTCAGCTGCTTACGCGGCGAACCTCGAG	CTCGAGGTTCGCCGCGTAAGCAGCTGACCAGTT CGTCGGCTCCTCGGT
CarD_R114A-R118A	GCGATTTGTGGCGTGCCGACCAGGAGGCTGGCTTGTC	GACAAGCCAGCCTCCTGGTCGGCACGCCACAA ATCGC
CarD_K125A-R126A-K130A	CTTGTCGGCCGGTGAGGCGGCCATGCTGGCCGCGGCCCGGCAGATT	GGCCGCGGCCAGCATGGCCGCCTCACCGGCCGA CAAG
CarD_Y11A-H14A	CGGAGACACCGTTGTCGCTCCACACGCCGGTGCTGCGTTAGTC	GACTAACGCAGCACCGGCGTGTGGAGCGACAA CGGTGTCTCCG
RpoBtr-NHis	AGAGAAGCATATGATGCTCCTTGACGTCCAGACCGATTCG	CGCGGATCCTTACATGAATTGGCTCAGCTGGCT
16S-upstream	AGACTGGCAGGGTCGCCCCGAAGCGG	CCGCCAGCGTTCGTCCTGAGCCAGGATC

Supplemental Experimental Procedures

Expression and purification

For recombinant expression of uncomplexed RpoBtr, native and mutant CarD proteins, the corresponding plasmids were transformed into *E. coli* BL21(DE3) cells. The cells were grown at 37 ºC in the presence of antibiotic to an OD_{600} of 0.6, and protein expression was induced with 1 mM IPTG at 18 ºC overnight. The cells were harvested, resuspended in buffer A (25 mM Tris pH 7.5, 500 mM NaCl, 5 mM imidazole, 2 mM β-ME, 2 mM $MgCl_2$, 40 μg/ml DNAse, and protease inhibitor cocktail) and lysed by French press. After centrifuging the lysate, the supernatant was loaded onto a His Trap FF column pre-equilibrated with buffer A. After washing the column with at least 5 column volumes of buffer A, column bound proteins were eluted with a gradient of buffer B (15-100%) (25 mM Tris pH 7.5, 500 mM NaCl, 500 mM imidazole, 2 mM β-ME). Fractions containing RpoBtr-NHis or CarD were pooled, dialyzed against buffer C (25 mM Tris 7.5, 100 mM NaCl, 2 mM DTT), concentrated, and loaded onto a Superdex 200 gel filtration column pre-equilibrated with buffer C. Final pure protein was concentrated to 15 mg/ml and stored at -80 ºC.

Selenomethionine (Se-Met) incorporated RpoBtr-NHis was expressed by using a methionine auxotrophic *E. coli* strain B834(DE3) in minimal media supplemented with Se-Met (Doublie, 1997). Purification of the Se-Met labeled protein was performed the same way as the native protein, except that the DTT concentration in buffer C was increased to 5 mM.

For co-expression of *Mtb* RNAP β1-β2 domains and *Mtb* CarD, the plasmids RpoBtr-NHis and CarD-notag were cotransformed into *E. coli* Rosetta2(DE3)pLysS cells. The cells were grown at 37 ºC in the presence of kanamycin and carbenecillin to an OD_{600} of 0.6 and were induced with 0.75 mM IPTG at 18 ºC overnight. The cells were harvested by centrifugation at

3000 x g and resuspended in buffer D (25 mM Tris pH 7.5, 200 mM NaCl, 5% (v/v) glycerol, 5 mM imidazole, 2 mM β-ME, 2 mM $MgCl_2$, 40 μg/ml DNAse, and protease inhibitor cocktail), lysed by French press, centrifuged at 16 K rpm for 45 min, and the supernatant was loaded onto a His Trap FF column pre-equilibrated with buffer D. The column was washed with at least 5 column volumes of buffer D and a stepwise gradient (5-25%) of buffer E (25 mM Tris pH 7.5, 200 mM NaCl, 300 mM imidazole, 2 mM β-ME). Column bound proteins were eluted with a gradient of buffer E (25-100%). Fractions containing the RpoBtr-NHis:CarD complex were pooled, concentrated, and loaded onto a Superdex 200 gel filtration column pre-equilibrated with buffer C (25 mM Tris 7.5, 100 mM NaCl, 2 mM DTT). The RpoBtr-NHis:CarD complex eluted as a single peak from the size exclusion column. Co-elution of RpoBtr and CarD from the IMAC and size-exclusion columns was verified by SDS-PAGE. Protein concentrations were estimated using the extinction coefficients of 17000 $M^{-1}cm^{-1}$ for CarD, 32490 $M^{-1}cm^{-1}$ for RpoBtr, and 49000 $M^{-1}cm^{-1}$ for RpoBtr/CarD complex at 280 nm using NanoDrop. Finally, the purified complex was concentrated to 10 mg/ml and stored at -80 ºC for further use.

Data collection and structure determination

Molecular replacement (MR) trials to determine the structure of *Mtb* RpoBtr using homologous *E. coli*, *Tth*, and *Taq* RNAP structures did not yield a solution. The structure of the *Mtb* RNAP β-subunit β1-β2 domains was solved by single-wavelength anomalous diffraction (SAD) using RpoBtr-NHis-SeMet crystals. Crystals belonged to the $P2_12_12_1$ space group and diffracted to 2.9 Å. Data was collected at the selenium peak wavelength (0.979 Å), processed with Denzo/Scalepack (Minor et al., 2006) and Phenix AutoSol, and Autobuilt modules were used to determine the initial phases and obtain the first atomic model (Adams et al., 2010). Subsequently, resolution of the RpoBtr structure was improved to 2.5 Å by diffraction data

obtained from native (non-SeMet) RpoBtr-NHis crystals, which also belonged to the $P2_12_12_1$ space group with two molecules in the ASU. The structure of native RpoBtr-NHis was obtained by MR using the RpoBtr-NHis-SeMet protein structure as the search model. The structure comprised residues 47-426 of the β-subunit and included six additional residues at the N-terminus between the His-tag and the first methionine residue. Refinement was performed with Phenix Refine. NCS restraints were applied to the atomic positions and individual B-factors subsequent to simulated annealing and rigid-body refinements. Iterative manual model building was done with COOT (Emsley et al., 2010). Water molecules were added to $|F_o\text{-}F_c|$ peaks >3σ and the model was further refined using Phenix Refine. The final model had R_{work} and R_{free} values of 0.21 and 0.26, respectively.

The RpoBtr-NHis:CarD complex crystals belonged to the $C222_1$ space group and the diffraction data to 2.1 Å resolution was processed with Denzo/Scalepack. The structure was solved by MR using the RpoBtr β1 and β2 domains as two different search ensembles (Phaser, CCP4) (McCoy et al., 2007; Winn et al., 2011). After locating one copy of RpoBtr in the ASU, the additional $|F_o\text{-}F_c|$ density for the CarD N-terminal domain β-strands and the C-terminal domain helices were visible with enough clarity to allow us to build CarD manually. After a few cycles of rigid-body and simulated-annealing refinements, the $|F_o\text{-}F_c|$ density for the side chains gradually improved. The model was further refined with Phenix Refine and water molecules were added to $|F_o\text{-}F_c|$ peaks >3σ. In the final stages of refinement, TLS refinement (four TLS groups: β1 domain, β2 domain, CarD N-terminal domain, and CarD C-terminal domain) was included. Iterative cycles of model building and refinement resulted in a final model which included one RpoBtr:CarD complex in the ASU with an R_{work}= 0.20 and an R_{free}= 0.23. Data collection and processing statistics are provided in Table 1.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Doublie, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods in enzymology *276*, 523-530.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta crystallographica Section D, Biological crystallography *66*, 486-501.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. Journal of applied crystallography *40*, 658-674.

Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. Acta crystallographica Section D, Biological crystallography *62*, 859-866.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.