Supplementary Materials for

Structural Basis of Non-Canonical Transcriptional Regulation by the σ^A-Bound Iron-Sulfur Protein WhiB1 in *M. tuberculosis*

Tao Wan[1+], Shanren Li[1+], Daisy Guiza Beltran[1], Andrew Schacht[1], Lu Zhang[1], Donald F. Becker[1,2], LiMei Zhang[1,2,3*]

[1]Department of Biochemistry, [2]Redox Biology Center and [3]Nebraska Center for Integrated Biomolecular Communication, University of Nebraska-Lincoln, N138 Beadle Center, 1901 Vine Street, Lincoln NE 68588, USA

[+]These authors contributed equally.

[*]To whom correspondence should be addressed: lzhang30@unl.edu

Supplementary Tables

Supplementary Table 1. Bacterial strains and plasmids used in this study.

Label	Description	Reference
Strains		
E.coli		
XL1-Blue	Host strain for routine cloning work	Stratagene
BL21-Gold (DE3)	Host strain for recombinant protein expression	Agilent Technologies
***M. smegmatis* (*Msm*)**		
WT	*Mycobacterium smegmatis* MC^2 155 (ATCC 700084), unmodified wildtype strain	ATCC
WhiB1-W3A	An unmarked *whiB1* mutant in *M. smegmatis* expressing WhiB1-W3A	This study
WhiB1-W49A	An unmarked *whiB1* mutant in *M. smegmatis* expressing WhiB1-W49A	This study
Plasmids		
Plasmids for expression and purification of protein from E.coli		
pCDF-1b-6HisMtbσ^A	The gene encoding σ^A was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the KpnI-XhoI site of pCDF-1b for expressing His_6-σ^A; $Spec^R$	This study
pCDF-1b-6HisMtb$\sigma^A{}_{C170}$	The C-terminal domain of $\sigma^A{}_{C170}$ (last 359-528 aa) gene was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted at the KpnI-XhoI site of pCDF-1b for expressing His_6-$\sigma^A{}_{C170}$; $Spec^R$	This study
pET28b-6HisMtb$\sigma^A{}_{C112}$	The gene encoding the C-terminal domain of σ^A containing the last 112 residues (417-528 aa) was amplified from pCDF-1b-6HisMtb$\sigma^A{}_{C170}$ and inserted into the NcoI-XhoI site of pET28b for expressing His_6-$\sigma^A{}_{C112}$. It was co-expressed with tagless WhiB1 for crystallization; Kan^R	This study
pCDF-1b-6HisMtb$\sigma^A{}_{C170}$-mutant (K513A, R515H, H516A, H516F, P517A, S518A or S520A)	A modification of pCDF-1b-6HisMtb$\sigma^A{}_{C170}$ by site-directed mutagenesis for expressing His_6-$\sigma^A{}_{C170}$ with a single mutation of either K513A, R515H, H516A, H516F, P517A, S518A or S520A; $Spec^R$	This study
pET19b-10HisMtbWhiB1	The full-length *whiB1* gene was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the NdeI-XhoI site of pET19b for expressing His_{10} tagged-WhiB1; Amp^R	This study

pET19b-10HisMtbWhiB1-mutants (W3A, F17A, F18A, W49A or W60A)	A modification of pET19b-10HisMtbWhiB1 by site-directed mutagenesis for expressing His_{10}-WhiB1 carrying a single mutation of either W3A, F17A, F18A, W49A or W60A; Amp^R	This study
pET28b-MtbWhiB1	The full-length *whiB1* gene was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the NcoI-XhoI site of pET28b for expressing tagless WhiB1; Kan^R	This study
pET28b-MtbWhiB1-mutant (W3A, E15A, F17A, F18A, W49A or W60A)	A modification of pET28b-MtbWhiB1 by site-directed mutagenesis for expressing tagless WhiB1 carrying a single mutation of either W3A, E15A, F17A, F18A, W49A or W60A; Kan^R	This study
pET21b-MtbWhiB1-76	The truncated *whiB1* gene (1-76 aa) was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the NdeI-XhoI site of pET21b(+) for expressing truncated WhiB1 without a tag. It was co-expressed with His_6-σ^A_{C112} for crystallization; Amp^R	This study
pETDuet-6His-SUMO-MtbWhiB1	The full length *whiB1* gene was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the NdeI/XhoI site of a modified pETDuet-SUMO for expressing His_6-SUMO-WhiB1 that was used for generating tagless WhiB1 in the pull-down assay; Amp^R	This study
pMP55	A gift from Dr. Robert Landick's group for expressing *M. tuberculosis* RNA polymerase core enzyme in *E. coli,* with a S450Y mutation in the β subunit and is resistant to rifampicin; Kan^R	(1)
pET19b-10His-3FLAG-MtbWhiB1	The full-length *whiB1* gene was amplified from *M. tuberculosis* H37Rv genomic DNA and inserted into the NdeI/XhoI site of pET19b with a 3FLAG tag fused in the N-terminus for expressing His_{10}-3FLAG-WhiB1 used for testing complex formation between WhiB1 and σ^A in the RNAP holoenzyme; Amp^R	This study
Plasmids for generating DNA probe in electrophoretic mobility shift assays (EMSA)		
pUC18-PwhiB1	A 297-bp DNA fragment containing the *whiB1* promoter (genomic region 3595416 – 3595712 bp) was amplified from *M .tuberculosis* H37Rv genomic DNA and cloned into the EcoRI/HindIII site of pUC18 used as a template for EMSA probe preparation; Amp^R	This study
Plasmids for generating mutations in M. smegmatis		
pJV53-GFP	A gift from Dr. Yicheng Sun's group for expressing recombination proteins gp60 and gp61 of mycobacteriophage Che9c and a GFP reporter. This plasmid was used for knockout deletion and mutation of *whiB1* in *M. smegmatis*; Kan^R	(2)

pKW08-Lx-Int	An integration plasmid for expressing desired genes under the inducible TetRO promoter in mycobacteria, containing *LuxAB* genes, L5 integrase gene and phage attachment *attP* sequences; Hyg^R	(3)
pKW08+3FLAGwhiB1LR	A 2121-bp DNA fragment of *whiB1* fused with 3FLAG in the N-terminus and its flanking sequence (genomic region 1999505 – 2001556 bp) was amplified from *M. smegmatis* $MC^{\wedge 2}$ 155 genomic DNA and cloned into the XbaI-HindIII site of pKW08-Lx-Int for expressing complementary *whiB1* under its native promoter; Hyg^R	This study
pUC-Zeo	A plasmid from Dr. Yicheng Sun's group contains a *dif*-flanked zeocin resistance cassette. It was used for constructing DNA fragments for gene deletion or site-directed mutation in *M. smegmatis*; Amp^R, Zeo^R	(2)
pUC-Zeo+MsmwhiB1LR	A 844-bp (genomic region 2000395 – 2001238 bp) and a 1135-bp (genomic region 1999011 – 2000145 bp) DNA fragments upstream and downstream of the *whiB1* gene, respectively, were amplified from *M. smegmatis* $MC^{\wedge 2}$ 155 genomic DNA and cloned into the left arm (KpnI-BglII site) and right arm (SpeI-HindIII site) flanking the zeocin resistance cassette in pUC-Zeo for the conditional *whiB1* deletion; Amp^R, Zeo^R	This study
pUC-Hyg	A gift plasmid from Dr. Yicheng Sun's group contains a *dif*-flanked hygromycin resistance cassette. It was used for constructing DNA fragment for deletion or site-directed mutation of *whiB1* in *M. smegmatis*; Amp^R, Hyg^R	(2)
pUC-Hyg+MsmwhiB1SDM	A 612-bp DNA fragment (genomic region 2000066 – 2000677 bp) including the *whiB1* gene and its upstream DNA and a 1135-bp DNA fragment (genomic region 1999212 – 2000065 bp) constraining the downstream DNA of the *whiB1* gene, respectively, were amplified from *M. smegmatis* $MC^{\wedge 2}$ 155 genomic DNA and cloned into the left arm (KpnI-BglII site) and right arm (NcoI-HindIII site) flanking the hygromycin resistance cassette in pUC-Hyg for the mutagenesis study of *whiB1* in *M. smegmatis*; Amp^R, Hyg^R	This study
pUC-Hyg+MsmwhiB1-(W3A,W49A)	A modification of pUC-Hyg+MsmwhiB1SDM by site-directed mutagenesis encoding a WhiB1-W3A and a WhiB1-W49A mutant, respectively, for site directed mutation of *whiB1* in *M. smegmatis*; Amp^R, Hyg^R	This study

Supplementary Table 2. Hydrophobic interactions between WhiB1 and σ^{A}_{CTD} at the [4Fe-4S] cluster binding pocket. A distance cutoff of 4.0 Å was used for this calculation.

Residues in σ^{A}_{CTD}	**Residues in WhiB1**	**Distance (Å)[a]**
His516	Trp60	3.43±0.03
His516	Phe18	3.85±0.05
His516	Phe17	3.51±0.11
Pro517	Ala7	3.85±0.01
Pro517	Trp60	3.48±0.07
Pro517	Trp3	3.79±0.12

[a]The reported distances are the average and the range of the shortest distances of the specified residues between the two copies of the WhiB1:σ^{A}_{CTD} complex molecules in an asymmetric unit.

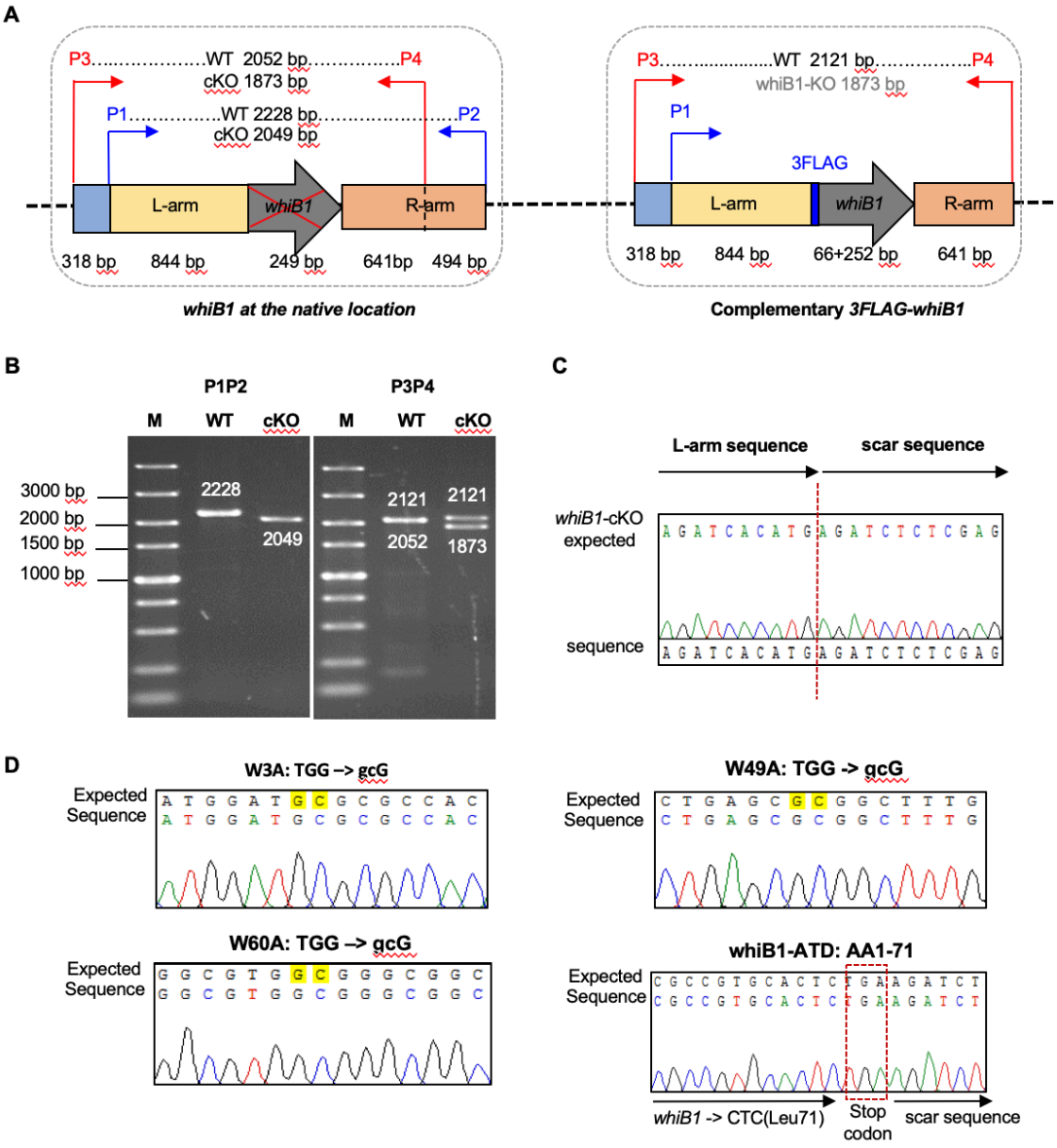
Supplementary Table 3. Polar interactions at the interface between WhiB1 and σ^{A}_{CTD}. A distance cutoff of 4.0 Å was used for this calculation.

Residues in σ^{A}_{CTD}	**Residues in WhiB1**	**Distance (Å)[a]**
At the protein surface		
Ser512/OG	Gln55/OE1	2.97±0.17
Lys513/NZ	Glu15/OE2	2.94±0.23
Arg515/O	Arg4/NH1	2.78±0.19
Ser518/OG	Cys9/O	2.79±0.05
Ser520/O	Arg10/NH2	2.66[b]
Ser520/OG	Arg4/NH1	2.61[b]
Pro517/N	Arg4/NH1	3.66±0.08
At the [4Fe-4S] cluster binding pocket		
His516/NE2	Val59/O	2.81±0.08
His516/NE2	SF4 101/S4	3.46±0.03

[a]The reported inter-atomic distances are the average and the range between the two copies of the WhiB1:σ^{A}_{CTD} complex molecules in an asymmetric unit, unless otherwise specified.

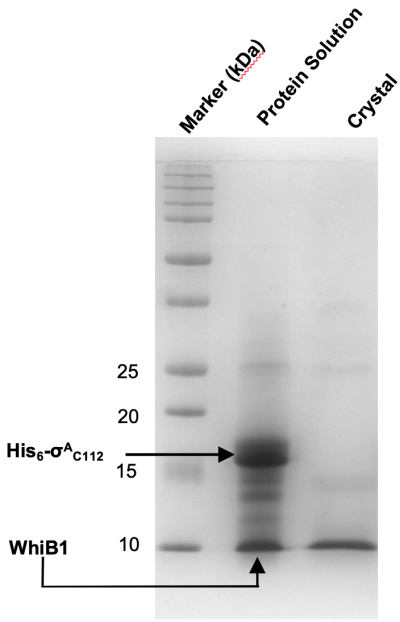
[b]The interactions are only observed in one copy of the WhiB1:σ^{A}_{CTD} complex within the distance cutoff of 4.0 Å.

Supplementary Figures

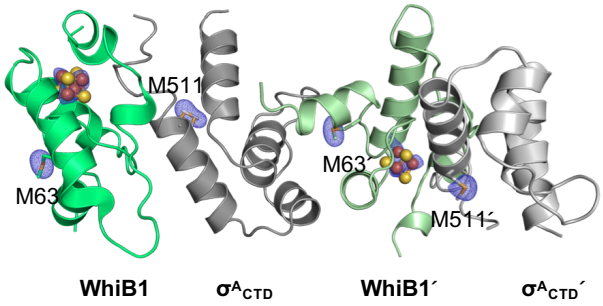


Supplementary Figure 1. Construction of the conditional *whiB1* knockout (cKO) mutant and the *whiB1* mutants by site-directed mutagenesis in *M. smegmatis*. (**A**) Schematic diagram of generating a conditional *whiB1* knockout mutant in *M. smegmatis*. The native and complementary *whiB1* genes and the corresponding upstream and downstream sequences are indicated. The 255-bp *whiB1* gene at the native location in the genome (indicated by a red cross) was removed by homologous recombination as described in the Methods. (**B**) Confirmation of the *whiB1* deletion at the native location by PCR using two pairs of primers. The first pair of the primers P1/P2 (CATCGAGGGCAAGCATCTGA/CCTATCACCGCATGGGACTG) were used to amplify the DNA sequence upstream and downstream of *whiB1* at the native location only, because the primer P2 matches the DNA sequence downstream of the native *whilB1* but out of the range of the complementary *whiB1*. The second pair of the primers P3/P4

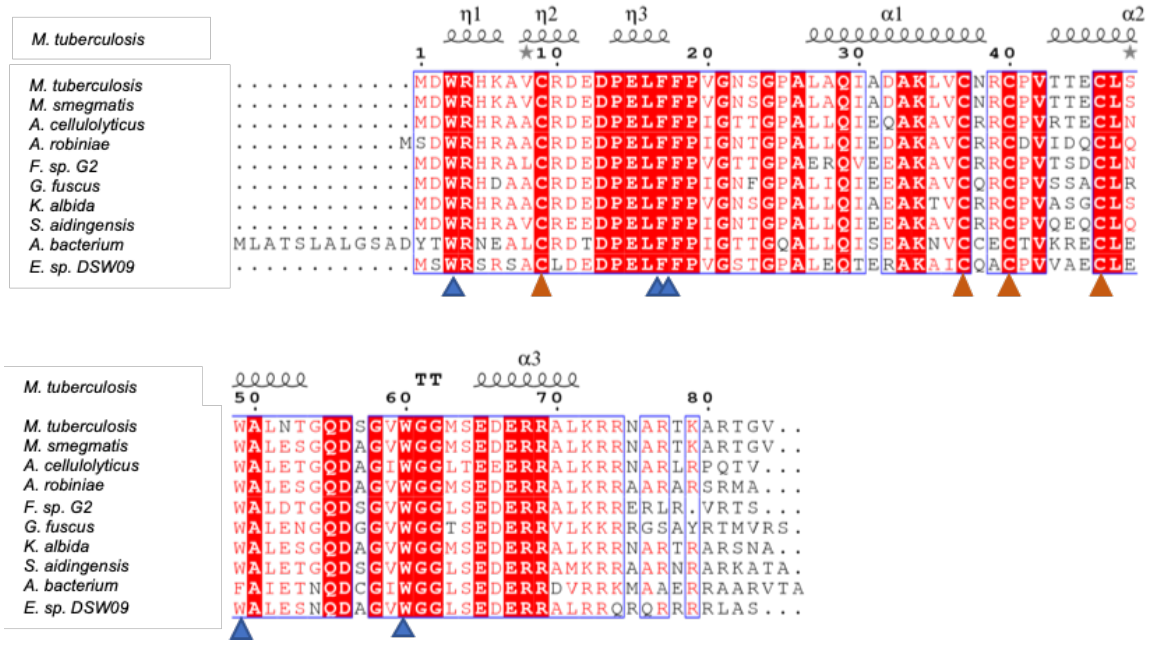
(ATACTGTTTAAACCTCTAGAGGCGCGACCTGCTGGCACA/
CGGTTGCAACGCATAAGCTTCACGACGCGTTGTCGATGTC) were used to amplifying the DNA sequences that are upstream and downstream of both *whiB1* at the native location and the complementary *whiB1*. The pair of the primers used and the expected size of the PCR products are indicated. (**C**) Confirmation of the *whiB1* gene deletion at the native location by sequencing of the PCR product (2049 bp) shown in Panel **B**. (**D**) Confirmation of the site-directed mutation and the truncated mutation in the *whiB1* gene by sequencing of the PCR products. The expected mutant sequence are shown on the top of each panel, followed by the results from DNA sequencing. The scar sequence is resulted from the excision of the *dif-zeo-dif* cassette.



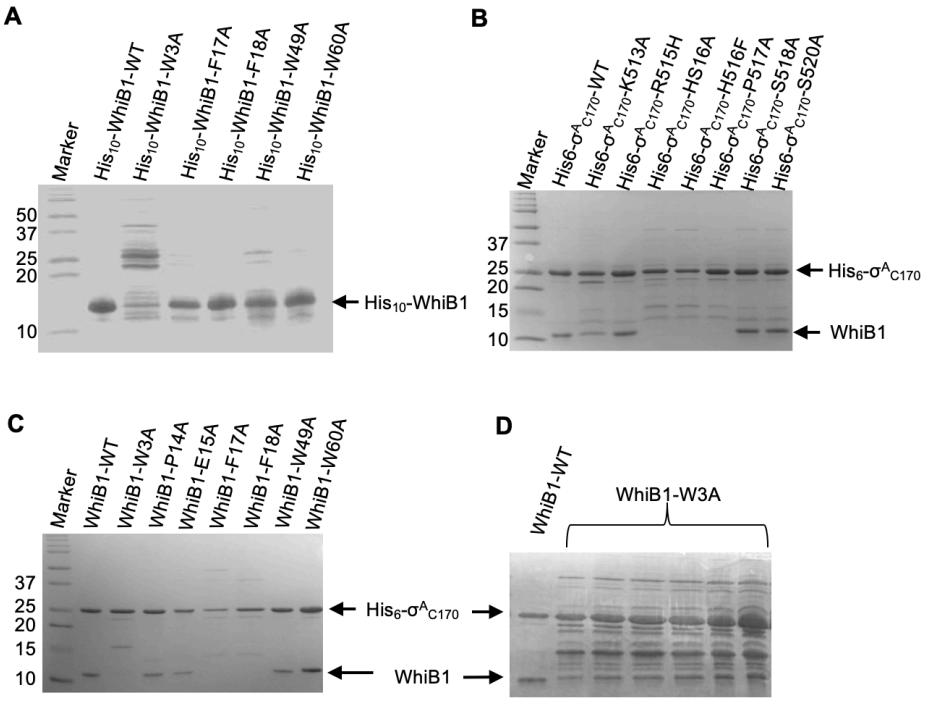
Supplementary Figure 2. SDS-PAGE analysis of the WhiB1:σ^{A}_{CTD} complex in solution and in the crystal. Lane 1, protein molecular weight marker; lane 2, solution sample of the purified WhiB1:σ^{A}_{C112} complex used for protein crystallization; lane 3, a single crystal of WhiB1:σ^{A}_{CTD} washed with the precipitation solution and dissolved in the same buffer as in the WhiB1:σ^{A}_{C112} protein solution (50 mM Tris-HCl, pH8.0, 100 mM NaCl, 1mM DTT). For the dissolved WhiB1:σ^{A}_{C112} complex crystal sample, partially degraded σ^{A}_{C112} showed an apparent size similar to WhiB1 on the SDS-PAGE gel.



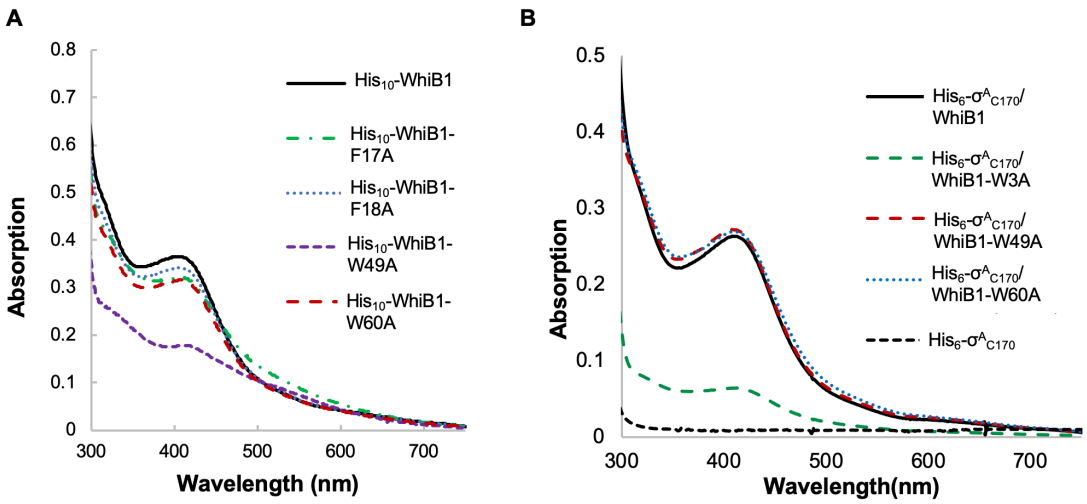
Supplementary Figure 3. Anomalous difference electron density map of the WhiB1:σ^{A}_{CTD} complexes in the asymmetric unit. The single-wavelength anomalous diffraction data were collected from a SeMet-substituted WhiB1:σ^{A}_{CTD} crystal at Se K-edge peak (12,679 eV). The anomalous difference electron density map is highlighted in light blue and contoured at 4.0 σ. The two WhiB1:σ^{A}_{CTD} complex molecules are shown in cartoon representation, with all four SeMet residues in the complexes highlighted in sticks and the [4Fe-4S] clusters in ball and sticks. Atoms are colored as follows: carbon in either green or pale green for WhiB1 and in grey for σ^{A}_{CTD}, S in yellow, Se in light orange and Fe in orange.



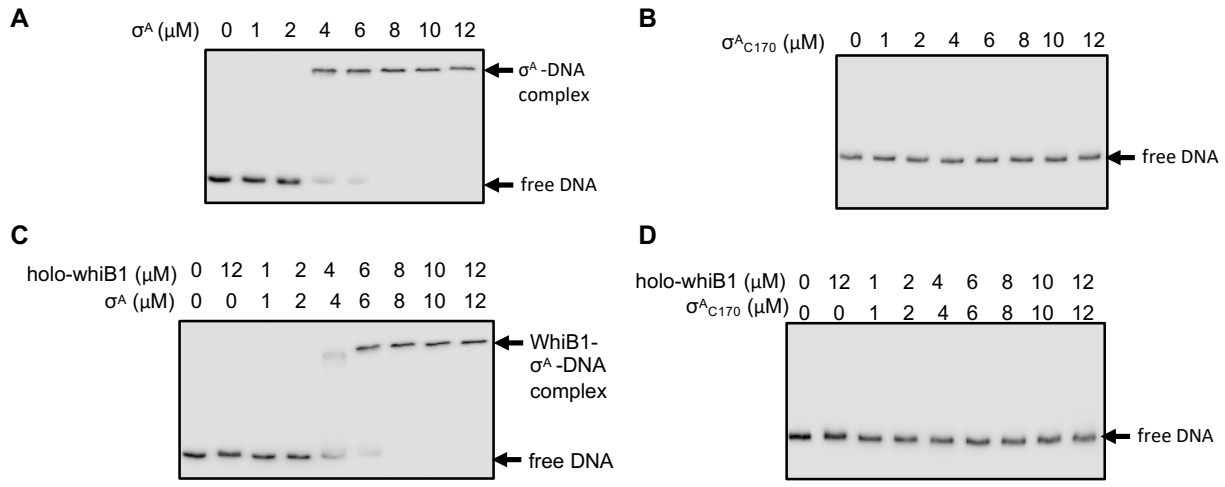
Supplementary Figure 4. Multiple sequence alignment of the WhiB1 homologues from representative bacteria in the phylum Actinobacteria. The first eight sequences are from bacterial species in different orders of the class Actinobacteria, which include *Mycobacterium tuberculosis* (*M. tuberculosis*) and *Mycobacterium smegmatis (M. smegmatis)* from the order Actinomycetale; *Acidother mus cellulolyticus* (*A. cellulolyticus*) from the order Acidothermales; *Actinospica robiniae* (*A. robiniae*) from the order Catenulisporales; *Frankia sp. G2* (*F. sp. G2*) from the order Frankiales; *Glycomyces fuscus* (*G. fuscus*) from the order Glycomycetales; *Kutzneria albida* (*K. albida*) from the order Pseudonocardiales; *Streptomyces aidingensis* (*S. aidingensis*) from the order Streptomycetales. The last two sequences are from two different classes of the phylum of Actinobacteria: *Acidimicrobiia bacterium BACL6 MAG-120924-bin43* (*A. bacterium*) from the class *Acidimicrobiia*; and *Euzebya sp. DSW09* (*E. sp. DSW09*) from the class *Nitriliruptoria*. The invariant residues in the aligned sequences are highlighted with red background, while other highly conserved residues are in red fonts. The five aromatic residues of interest (W3, F17, F18, W49 and W60 in *M. tuberculosis* WhiB1) are highlighted by blue triangles, and the orange triangles indicate the four invariant cysteines (C9, C37, C40 and C46 in *M. tuberculosis* WhiB1) coordinated to [4Fe-4S] cluster.



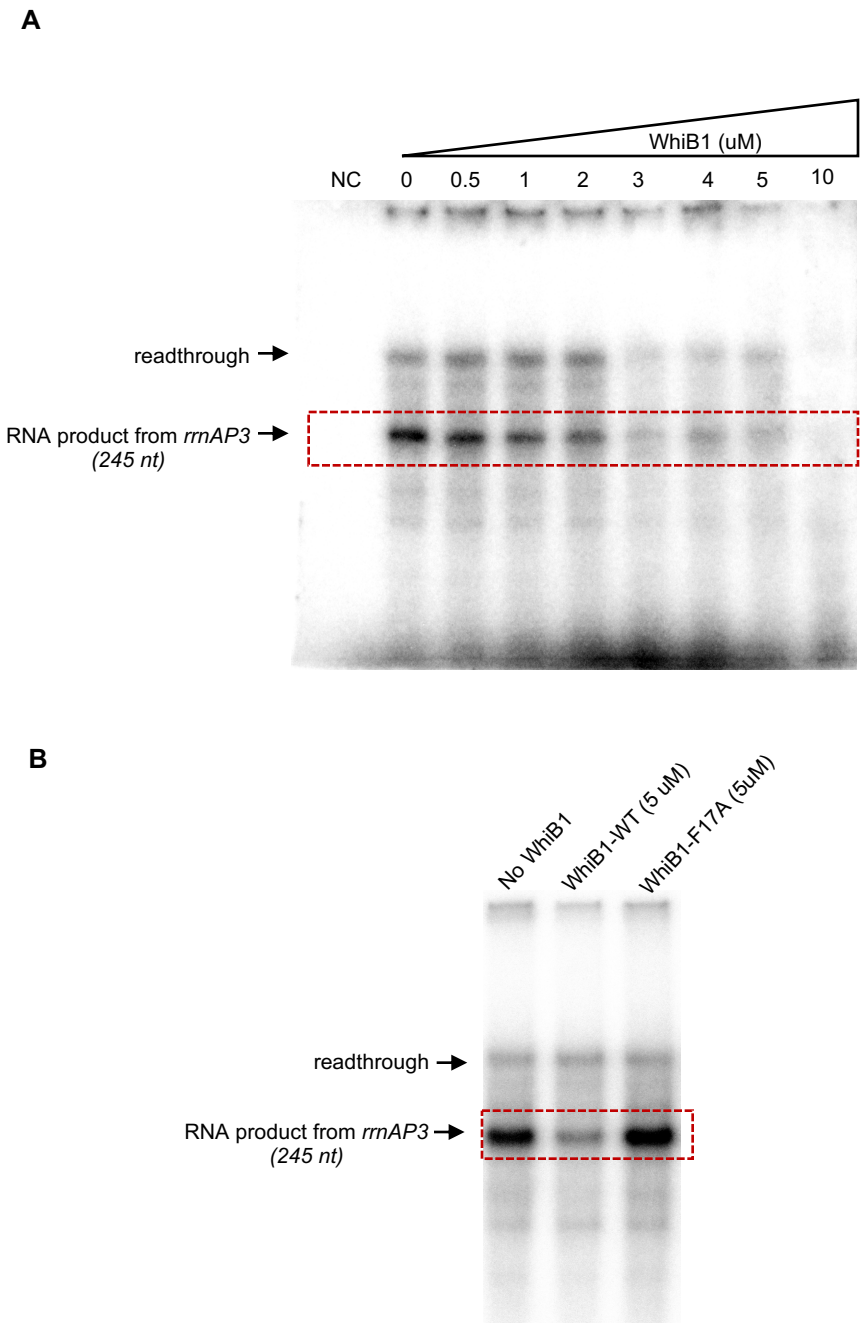
Supplementary Figure 5. SDS-PAGE analysis of the WhiB1 proteins expressed and purified alone and in complex with the His_6-σ^A_{170}. (**A**) SDS-PAGE analysis of the protein samples from the overexpression and affinity purification of His_{10}-WhiB1 (either wildtype or mutant as indicated) alone. A W3A mutation in WhiB1 (His_{10}-WhiB1-W3A) disrupted the stability of the [4Fe-4S] cluster and thus the protein stability, resulting in the low level of His_{10}-WhiB1-W3A in the whole cell supernatant and the higher level of contaminants from *E. coli* native proteins during the Ni^{2+} affinity purification as described in the MATERIALS AND METHODS. (**B**) SDS-PAGE analysis of the protein samples from co-expression and affinity purification of the His_6-σ^A_{C170} (either wildtype or mutant as indicated) with the wildtype tagless WhiB1. (**C** and **D**) SDS-PAGE analysis of the protein samples from co-expression and affinity purification of the wildtype His_6-σ^A_{C170} with the tagless WhiB1 (either wildtype or mutant as indicated). Increasing amounts of the protein sample from the co-expression and affinity purification of His_6-σ^A_{C170} with WhiB1-W3A were used in Panel **D** to show the weak band of WhiB1-W3A. The band with an apparent molecular weight of ~15 kDa shown in Panels **C** and **D** resulted from partial degradation of His_6-σ^A_{C170}, which is more significant with reduced complex formation by the WhiB1 mutants. The peptide identities of the partial degraded His_6-σ^A_{C170} and WhiB1-W3A were confirmed by mass spectrometry. Ni^{2+} Sepharose High Performance resin (GE Healthcare Life Sciences) was used for affinity purification.



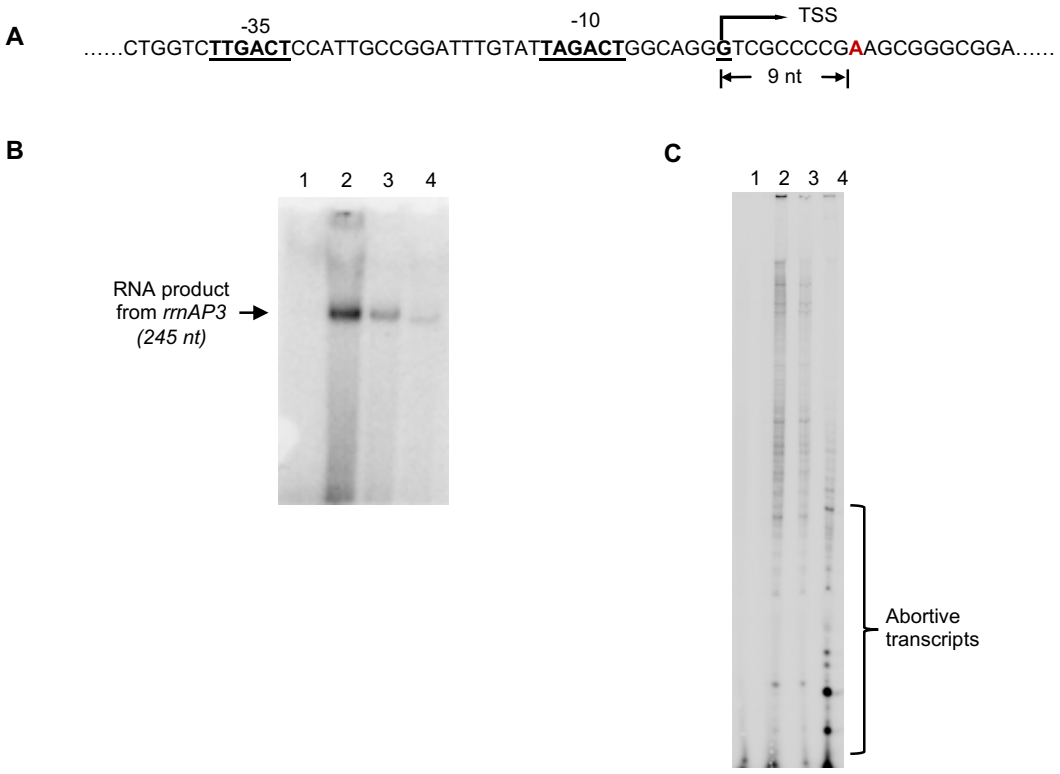
Supplementary Figure 6. UV-Visible spectra of the WhiB1 proteins alone and in complex with His_6-σ^A_{C170}. The intensity of the absorption peak around 410 nm is indicative of the occupancy of the [4Fe-4S] cluster in WhiB1. The absorption spectra were normalized at 280 nm based on the protein concentrations estimated by the Pierce Bradford Assay Kit. (**A**) His_{10}-WhiB1 wildtype and mutant proteins expressed and purified from *E. coli*. By comparison to the wildtype protein, the His_{10}-WhiB1-W49A mutants showed distinct features in the visible region and lower peak intensity around 410 nm of the absorption spectrum, indicating a lower occupancy of the [4Fe-4S] cluster. In contrast, the F17A, F18A and W60A mutations in WhiB1 showed no significant effect on the absorption spectra. (**B**) The tagless WhiB1 wildtype and mutant proteins co-expressed and purified with His_6-σ^A_{C170} from *E. coli*. The visible region of the absorption spectra for all the WhiB1 mutant proteins in complex with His_6-σ^A_{C170} are comparable to that for the WhiB1 wildtype protein alone, except for WhiB1-W3A, whose peak intensity around 410 nm is decreased to ~1/3 of the wildtype protein complex. The UV-Visible spectrum of His_6-σ^A_{C170} is shown as a reference.



Supplementary Figure 7. Effect of holo-WhiB1 on σ^A binding to the *whiB1* promoter DNA. The effect of holo-WhiB1 on σ^A binding to the *whiB1* promoter (P_{whiB1}) was determined by electrophoretic mobility shift assays (EMSAs). For all reactions, 0.2 nM biotin-labeled DNA template of P_{whiB1} was incubated with the proteins at the indicated concentrations in a 20-μl reaction buffer (100 mM NaCl, 40 mM Tris pH 8.0, 10 mM $MgCl_2$, 1 mM EDTA, 0.25 mg/ml BSA, 1 mM DTT). 50 ng/μl poly(dI-dC) was used in all reactions as a non-specific competitor. For reactions containing holo-WhiB1, σ^A (or $\sigma^A{}_{C170}$) was preincubated with an equimolar concentration of holo-WhiB1 under anaerobic conditions for 30 min to allow complex formation before adding P_{whiB1}. The reaction mixtures were separated on 6% polyacrylamide gels and the labeled DNA probes were detected using LightShift Chemiluminescent EMSA Kit (Thermo Fisher Scientific). (**A** and **C**) The gel images from the EMSAs of σ^A binding to P_{whiB1} without and with holo-WhiB1, respectively. The σ^A- P_{whiB1} complex was detected at the same concentration of σ^A with or without addition of holo-WhiB1, indicating no strong effect of holo-WhiB1 on σ^A binding to P_{whiB1}. (**B** and **D**) The gel images from the EMSAs of $\sigma^A{}_{C170}$ binding to P_{whiB1} without and with WhiB1, respectively. No protein-DNA complex formation was detected during the experimental conditions, as expected for a monomeric $\sigma^A{}_{C170}$ with a single DNA binding motif.



Supplementary Figure 8. WhiB1 represses transcription from the *rrnAP3* promoter *in vitro* by interaction with σ^A in the RNA polymerase holoenzyme. (**A**) and (**B**) are full gel images from which the cropped images (highlighted in the red frames) are shown in Fig. 5D.



Supplementary Figure 9. Effect of WhiB1 on the promoter escape during the *in vitro* transcription from *rrnAP3*. (A) The sequence of the *rrnAP3* promoter. The -10 and -35 consensus elements are underlined, and the transcription start site (TSS) is marked by an arrow. The first adenosine (A) expected in the 10th nucleotide position of the transcript product is highlighted in red. (**B**) and (**C**) are the gel images of the *in vitro* transcription products analyzed on a 8M-6% and a 8M-24% polyacrylamide gel, respectively. The *in vitro* transcription from the *rrnAP3* promoter was set up similarly as described in Fig. 5D. Lane 1, the reaction with the RNAP core enzyme alone as the negative control; Lane 2, the reaction with the RNAP holoenzyme (i.e., including σ^A) and without WhiB1; Lane 3, the reaction with the RNAP holoenzyme and 10 μM WhiB1; Lane 4, the forced abortive transcription from the *rrnAP3* promoter with the RNAP holoenzyme alone as in Lane 2 but without ATP.

Supplementary References

1. Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R. A., Landick, R., Darst, S. A., and Campbell, E. A. (2018) Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. *Elife* **7**
2. Mao, X. J., Yan, M. Y., Zhu, H., Guo, X. P., and Sun, Y. C. (2016) Efficient and simple generation of multiple unmarked gene deletions in *Mycobacterium smegmatis*. *Sci. Rep.* **6**, 22922
3. Williams, K. J., Joyce, G., and Robertson, B. D. (2010) Improved mycobacterial tetracycline inducible vectors. *Plasmid* **64**, 69-73