Aromatic Patch in WhiB-Like Transcription Factors Facilitates Primary Sigma Factor Interaction in *Mycobacterium tuberculosis*

Daisy Guiza Beltran[1,#], Tao Wan[1,9,#], Javier Seravalli[1,2], Magdaléna Horová[1], Camden Jones[1], Shanren Li[1,10], Pengchong Ma[3], Donald F Becker[1,2], Jeffery P. Mower[4], Qiuming Yao[3], Yu Pan[5], Hongfeng Yu[3], Adrie J. C. Steyn[6,7], LiMei Zhang[1,2,8*]

[1]Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, USA

[2]Redox Biology Center, University of Nebraska-Lincoln, Lincoln, NE, USA

[3]School of Computing, University of Nebraska-Lincoln, Lincoln, NE, USA

[4]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA

[5]Department of Biological System Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA

[6]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA

[7]Africa Health Research Institute, University of KwaZulu-Natal, Durban, South Africa

[8]Nebraska Center for Integrated Biomolecular Communication, University of Nebraska-Lincoln, Lincoln, NE, USA

[9]Present address: Biortus Biosciences Co., Ltd, Jiangyin, Jiangsu, 214437, China

[10]Present address: College of Life Sciences, Fujian Normal University, Fuzhou, Fujian, 350117, China

[#]These authors contribute equally

[*]Correspondence: lzhang30@unl.edu

Keywords: WhiB-Like (Wbl) family, aromatic patch, primary sigma factor A, iron-sulfur cluster, X-ray crystallography

SIGNIFICANCE

This work addresses a critical knowledge gap in understanding how WhiB-like (Wbl) family proteins regulate gene expression and confer virulence and antibiotic resistance in *Mycobacterium tuberculosis*. WhiB-like proteins are widely distributed in Actinobacteria, the phylum that includes several deadly human pathogens and important antibiotic producers. By identifying the previously unrecognized structural motif that is highly conserved in the Wbl family and mediates Wbl interaction with the core transcriptional machinery, this work advances our understanding of the evolution, function and underlying mechanism of transcriptional regulation by this unique family of transcription factors in Actinobacteria and associated phages. These insights will be instrumental in developing new strategies to combat tuberculosis and other diseases caused by related pathogens.

ABSTRACT

WhiB-like (Wbl) family proteins are a group of small iron-sulfur ([4Fe-4S]) cluster-containing transcription factors and are uniquely found in Actinobacteria and actinobacteriophages. They play pivotal roles in diverse biological processes, including cell development, redox stress response and antibiotic resistance. However, the lack of a canonical DNA binding motif in most Wbl proteins has long obscured the mechanistic understanding of how Wbl family proteins regulate gene expression. Previous studies have shown several members of Wbl family proteins from *Mycobacterium tuberculosis* (*Mtb*) and *Streptomyces* interact with the conserved region 4 of the primary sigma factor ($\sigma^A{}_4$) in RNA polymerase holoenzyme for transcriptional regulation. The molecular interface between Wbls and $\sigma^A{}_4$ is unexpectedly dominated by hydrophobic interactions featuring a patch of conserved aromatic residues within the iron-sulfur cluster binding pocket. In this study, we present structural, molecular and biochemical evidence demonstrating that all *Mtb* Wbls bind to the same site on σ^A using a patch of conserved aromatic residues (hereby referred to as the aromatic patch) within the iron-sulfur cluster binding pocket. Moreover, our bioinformatic analyses and structural modeling of representative Wbl family proteins reveal that this aromatic patch is a conserved structural motif across the Wbl family. Our findings provide compelling evidence that Wbl proteins regulate gene expression via a shared interaction interface with $\sigma^A{}_4$ and uncover a complex evolutionary relationship of Wbls between actinobacteria and their associated phages.

INTRODUCTION

WhiB-like (Wbl) proteins are a unique family of iron-sulfur ([4Fe-4S]) cluster-containing transcription factors. The *wbl* genes were discovered in *Streptomyces coelicolor* (*Sco*) in 1972 and named after the white colonies of the *wbl* mutants due to the disruption of pigment formation (1). Subsequent genetic and bioinformatic analyses find that the Wbl proteins are broadly shared across the phylum Actinobacteria, including the species of medical and industrial significance on human diseases and economics such as *Mycobacterium tuberculosis* (*Mtb*), *Sco* and *Corynebacterium glutamicum* (*Cglu*) (2,3). Wbl proteins play critical roles in diverse biological processes in Actinobacteria, including redox stress responses, cell division, and development, nutrition starvation response, antibiotic resistance and pathogenesis (3-20). Interestingly, Wbls are also commonly found in actinobacteriophages but not in any other living organisms.

Wbl family proteins are characterized by two conserved motifs, the [4Fe-4S]-cluster binding motif consist of four conserved cysteines, and the Gly-rich motif (also referred to as the β-turn) (Fig. 1A) (21). To date, over 35,000 Wbl sequences have been deposited in the Wbl family database (Pfam ID: PF02467; InterPro ID: IPR034768). Certain actinobacterial genera, such as Cyanobacteria, Mycobacteria, *Rhodococcus* and *Streptomyces*, encode multiple Wbl paralogues (2,22). In *Mtb*, for example, there are seven Wbl members (WhiB1-7) that share 24% to 50% sequence identity. Based on their sequence similarity, Wbl family proteins can be grouped into five major subfamilies, each represented by a *Mtb* Wbl (WhiB1-4, and WhiB7, respectively) and distributed across a wide range of Actinobacterial species (2,22). The rest of the Wbl proteins seem to be species-specific, such as WhiB5 and WhiB6 found only in Mycobacteria. Several subfamily-specific functional motifs have been reported, most notably the well-characterized "AT-hook" DNA binding motif in the WhiB7 subfamily (23-26). However, most Wbl proteins lack canonical DNA binding motifs, or no other Wbl family-wide motif apart from the two conserved motifs mentioned above has been reported. This knowledge gap has long posed a challenge in understanding how Wbl proteins regulate gene expression.

The discovery that WhiB3 requires interaction with the conserved C-terminal region 4 of the primary σ^{70} sigma factor (σ^{A}_{4}) in the RNA polymerase (RNAP) holoenzyme to activate transcription marks a milestone in deciphering Wbl function (5). σ^{A}_{4} is highly conserved in the bacterial σ^{70}-family primary sigma factors, responsible for recognizing the -35 promoter element and serves as an anchor for many transcription factors for gene activation (27-29). The subsequent studies show that other *Mtb* Wbls also bind to σ^{A}_{4} (13,29,30). Recent structural studies on several Wbl proteins by our group and others – including WhiB1, WhiB3 and WhiB7 from *Mtb*, as well as *Streptomyces venezuelae* (*Sve*) WhiB (a homolog of *Mtb* WhiB2) – have revealed the unprecedented molecular interface between Wbls and σ^{A}_{4} (25,26,30-33). Distinct from other previously characterized σ^{A}_{4}-dependent transcription factors, Wbl proteins bind to the same site of σ^{A}_{4} via unexpected tight hydrophobic interactions. These Wbl proteins share a similar helical architecture surrounding the [4Fe-4S] binding pocket and the molecular interaction interface with σ^{A}_{4} despite low sequence identity (Fig. 1A; SI Appendix, Fig. S2) (25,26,30-33). H516 and P517 in the C-terminus of σ^{A}_{4} insert into the [4Fe-4S] cluster binding pocket of Wbls and interact with several aromatic residues (hereby referred to as the aromatic patch), including one residue at the beginning of the helix *h1* (corresponding to W3 in *Mtb* WhiB1, labeled as aromatic patch 1 [AP1]), two in the end of helical *h1* (corresponding to F17 and F18 in *Mtb* WhiB1 and labeled as AP2-3), one in the β-turn connecting *h3* and *h4* (corresponding to W60 in *Mtb* WhiB1 and labeled as AP4). These aromatic residues are conserved in each Wbl subfamily and play a crucial role in σ^{A}_{4} binding (SI Appendix, Fig. S1) (25,32,33). The absence of the equivalent of AP1 results in higher oxygen sensitivity of the WhiB7:σ^{A}_{4} complex relative to that of WhiB1 and makes the remaining three residues in the aromatic patch indispensable for WhiB7 binding to σ^{A}_{4} (25).

Other *Mtb* Wbl proteins, including WhiB4 and WhiB6, have also been shown to interact with σ^{A}_{4}, while it remains inconclusive for WhiB5 (29). Due to the lack of structural and biochemical information, it is unclear precisely how these *Mtb* Wbl proteins interact with σ^{A}_{4}. In this work, we present structural, molecular and biochemical evidence to demonstrate that all *Mtb* Wbl proteins, including WhiB5, bind to the same site on σ^{A}_{4} via a conserved aromatic patch containing at least two aromatic residues corresponding to AP1-4 (W3, F17, F18 and W60 in WhiB1) and at least one corresponding to AP2-3 (F17 and F18 in WhiB1). Furthermore, our structural modeling, sequence and phylogenetic analyses of Wbl proteins from diverse Actinobacteria and actinobacteriophages indicate that this aromatic patch is highly conserved among the Wbl family

members. Together, our findings uncover a previously unrecognized structural feature underlying the molecular mechanisms of action of the Wbl family proteins and offer new insights into the evolution and function of this unique family of transcription factors.

RESULTS

Crystal structure of the WhiB6:$\sigma^A{}_4$ complex

WhiB6 is a Mycobacteria-specific Wbl protein and has been shown to regulate the expression of the Type VII secretary system-1 (ESX-1) and play an important role in the reactivation of *Mtb* and other pathogenic mycobacterial species (15,34-38). Unlike other structurally characterized Wbls, *Mtb* WhiB6 and its homologs lack the two equivalent residues in the aromatic patch, AP1 and AP3 (Fig. 1A; SI Appendix, Fig. S1). To gain atomic insights into how WhiB6 interacts with $\sigma^A{}_4$, we purified and crystalized the $\sigma^A{}_4$-bound WhiB6 to determine the crystal structure of the complex. As previously described, we used the chimera protein of $\sigma^A{}_4$ fused to the RNAP β-subunit flap tip helix (β_{tip}) via an artificial linker (hereby denoted $\sigma^A{}_4$-β_{tip}) to mimic the interaction between σ^A and the β-subunit in the RNAP holoenzyme (39-41). The WhiB6:$\sigma^A{}_4$-β_{tip} complex was purified from *E. coli* by co-expressing a tagless WhiB6 and a 6xHis-tagged $\sigma^A{}_4$-β_{tip} and was crystallized under anaerobic conditions. The crystal structure of WhiB6:$\sigma^A{}_4$-β_{tip} was refined at 1.8 Å (see in SI Appendix, Materials and Methods, Tables S1 and S2).

As observed in the other Wbl:$\sigma^A{}_4$ complexes (25,26,31-33), the overall architecture of WhiB6:$\sigma^A{}_4$ is dominated by alpha helices, and the molecular interface between $\sigma^A{}_4$ and WhiB6 is centered around the [4Fe-4S] cluster binding pocket (Fig. 1B). The core helical arrangements of WhiB6 surrounding the Fe-S cluster are surprisingly similar to that of other Wbl:$\sigma^A{}_4$ complexes despite the low sequence identity. A structural overlay of $\sigma^A{}_4$–bound WhiB6 with WhiB1 shows a Cα root mean square deviation ($RMSD_{C\alpha}$) of 1.07 Å for the 37 aligned residues located in helices *h1-3* of WhiB1 (Fig. 1C). In contrast, striking differences are observed in the rest of the structural arrangements that are away from the Fe-S cluster binding pocket, in particular, WhiB6 has a significantly shorter helix *h2* and loop *l1* between *h1* and *h2*, but a much longer helix *h4* compared to WhiB1.

At the molecular interface of WhiB6:$\sigma^A{}_4$-β_{tip} complex, H516 and P517 of $\sigma^A{}_4$ interact with W21 and W56 of WhiB6 (corresponding to F17 and W60 of WhiB1) within the [4Fe-4S] cluster binding pocket of the complex (Fig. 1D). Additionally, a conserved Arg, R38, in the WhiB6 subfamily caps the Fe-S cluster pocket in the complex. R38 may serve as a functional equivalent of W3 in WhiB1 based on the structural comparison of the two complexes. As expected, substitution of either H516 in $\sigma^A{}_4$ or W21 in WhiB6 by an Ala abolishes the complex formation in the pull-down assays (Fig. 1E and F), highlighting the essential role of the aromatic patch in the formation of WhiB6:$\sigma^A{}_4$. T22 of WhiB6 in loop *l1*, which corresponds to F18 of WhiB1 (Fig. 1D), is close to the molecular interface of the WhiB6:$\sigma^A{}_4$-β_{tip} complex within the [4Fe-4S] cluster binding pocket with the hydroxyl group of T22 facing away from the [4Fe-4S] cluster. Substitution of T22 with an Ala did not affect $\sigma^A{}_4$ binding in the pull-down assay, confirming a non-essential role of WhiB6 T22 for $\sigma^A{}_4$ binding and consistent with high sequence variability at this position among the WhiB6 homologs (SI Appendix, Fig. S1).

In vitro and in vivo protein-protein binding assays between WhiB5 and $\sigma^A{}_4$

Unlike WhiB6, three out of the four residues in the aromatic patch are found in the WhiB5 subfamily, despite the low sequence identity (<30%) between WhiB5 and other *Mtb* Wbl proteins

(Fig. 1A; SI Appendix, Fig. S1). It was puzzling to note that WhiB5 did not bind to σ^A like other Wbl proteins in the *in vitro* and *in vivo* binding assays by Feng *et al.* (29). However, as the authors pointed out in their study, since the *in vitro* pull-down assay was carried out under aerobic conditions, oxidation damage of the [4Fe-4S] cluster in WhiB5 may also abolish σ^A binding.

To revisit the question of whether WhiB5 binds to $\sigma^A{}_4$, we tested the interaction between WhiB5 and $\sigma^A{}_4$ using the mycobacterial two-hybrid system assay termed mycobacterial protein fragment complementation (M-PFC) assay. The M-PFC assay is based on the functional reconstitution of two murine dihydrofolate reductase (mDHFR) domains, $mDHFR_{F[1,2]}$ and $mDHFR_{F[3]}$, respectively, by fusing the two domains independently to two interacting proteins. When the proteins of interest interact, mDHFR activity is restored, conferring bacterial resistance to trimethoprim (TRIM), an antibiotic targeting bacterial DHFR (42). The M-PFC assay was also used in the previous study of WhiB5 binding to σ^A (29), in which $mDHFR_{F[1,2]}$ was fused to the N-terminus of WhiB5. It is noted that WhiB5 has an unusually short sequence in the N-terminus prior to the first Fe-S cluster-ligating cysteine residue compared to other *Mtb* Wbl proteins (Fig. 1A). Indeed, we found that an N-terminal His-tag, but not a C-terminal tag, results in significant formation of inclusion bodies when over-expressing the recombinant WhiB5 protein in *E. coli* (data not shown). To avoid possible interference of WhiB5 folding by the N-terminal $mDHFR_{F[1,2]}$ domain and to stabilize the conformation of $\sigma^A{}_4$, we constructed the two mDHFR fragments to the C-terminus of WhiB5 and $\sigma^A{}_4$-β_{tip} denoted WhiB5-$mDHFR_{F[1,2]}$ and $\sigma^A{}_4$-$mDHFR_{F[3]}$, respectively, in our M-PFC assay (see in SI Appendix, Materials and Methods). As shown in Fig. 2A, all the strains grow similarly on the 7H10 agar plates without TRIM. In the presence of TRIM, the *Mycobacterium smegmatis* (*Msm*) strain expressing WhiB5-$mDHFR_{F[1,2]}$ and $\sigma^A{}_4$-$mDHFR_{F[3]}$, but not the negative control, grows on the TRIM agar plate comparably to the positive control, confirming the interaction of WhiB5 with $\sigma^A{}_4$ in *Msm*. This observation cements the physiological relevance of WhiB5 binding to $\sigma^A{}_4$.

We then used isothermal titration calorimetry (ITC) to quantify the WhiB5 binding to $\sigma^A{}_4$ under anaerobic conditions and compare it with WhiB1 and WhiB2. The ITC analysis shows that all three Wbl proteins bind to $\sigma^A{}_4$ with low nM dissociation constants (K_d) and the binding reactions are exothermic (Fig. 2B-E), which agrees favorably with the hydrophobic interaction-driven molecular interface of the Wbl:$\sigma^A{}_4$ complex. The binding affinity of WhiB5 to $\sigma^A{}_4$ is about two- to three-fold lower than that of WhiB1 or WhiB2, consistent with the absence of a key residue in the aromatic patch of WhiB5 for $\sigma^A{}_4$ binding compared to WhiB1 and WhiB2 (Fig. 1A).

The aromatic patch is required for $\sigma^A{}_4$ binding by WhiB4 and WhiB5

As no atomic structural information is available for the interaction between $\sigma^A{}_4$ and WhiB4 or WhiB5, we employed the structural modeling approach using SWISS-MODEL and AlphaFold3 to evaluate whether these two Wbl proteins also utilize the aromatic patch to bind to the same site on $\sigma^A{}_4$ (see in SI Appendix, Materials and Methods). In both homologous models of the WhiB4:$\sigma^A{}_4$ and WhiB5:$\sigma^A{}_4$ complexes, the conserved aromatic residues in the WhiB4 and WhiB5 subfamilies corresponding to those in the aromatic patch are found at the molecular interface within the Fe-S binding pocket (SI Appendix, Figs. S1 and S4A-B).We note that the overall architecture of homologous models of the $\sigma^A{}_4$-bound WhiB4 and WhiB5 match well with the *ab initio* models generated by AlphaFold3 (SI Appendix, Fig. S4 C-D).

We validated the role of the aromatic patch residues in WhiB4 and WhiB5 for $\sigma^A{}_4$ binding by mutagenesis and *in vitro* pull-down assays. As observed for other Wbl proteins, a H516A

substitution in $\sigma^A{}_4$ abolishes $\sigma^A{}_4$ binding to either WhiB4 or WhiB5 (Fig. 3A and B), confirming that all *Mtb* Wbl proteins, including WhiB5, bind to the same site of $\sigma^A{}_4$. Mutation of either L45 or F46 by an Ala in WhiB4, corresponding to AP2 and AP3 (F17 and F18 in WhiB1), respectively, also abolishes WhiB4 binding to $\sigma^A{}_4$ (Fig. 3C). Additionally, the F46A mutation in WhiB4, but not the L45A mutation, significantly affects the [4Fe-4S] cluster stability, indicated by a decrease in the absorption at 410 nm. This observation echoes the importance of the aromatic patch residues for the [4Fe-4S] cluster stability as previously observed in the case of WhiB7 (25). Similarly, either a W12A or F13A mutation significantly affects WhiB5 binding to $\sigma^A{}_4$ in comparison to the wildtype WhiB5 (Fig. 3D). However, we noticed that residual WhiB5 mutants were still found in the pull-down sample. We then tested and confirmed that the tagless WhiB5 proteins (either WT or mutants) do not bind to the NiNTA resin in the absence of $\sigma^A{}_4$, suggesting that both WhiB5 mutants do retain low-level affinity to $\sigma^A{}_4$, either specifically or non-specifically.

The aromatic patch used by Mtb Wbls for $\sigma^A{}_4$ binding is highly conserved in the Wbl family

Since five *Mtb* Wbl proteins are well distributed in Actinobacteria and $\sigma^A{}_4$ is highly conserved, we speculate that the aromatic patch used by *Mtb* Wbls for $\sigma^A{}_4$ binding may be found in many Wbl family proteins. To evaluate how well the aromatic patch is conserved in the Wbl family, we conducted a phylogenetic analysis of 995 Wbls from the representative proteomes of Actinobacteria and actinobacteriophages with a co-membership threshold of 15% (RP15), as defined in the Pfam database (Pfam ID: PF02467) (see in Materials and Methods). These Wbl proteins, of which 55% are from Actinobacteria and 45% are from Actinobacteriophages, were classified into 29 subfamilies (Fig. 4A). As expected, five major subfamilies, each represented by a single *Mtb* Wbl (WhiB1–4 and WhiB7), account for ~80% of the total sequences and encompass diverse actinobacterial species, which is in line with the previous bioinformatic analysis (2). Among these five major subfamilies, the WhiB2 subfamily is the largest Wbl subfamily and is found most widely distributed in Actinobacteria (2). Our analysis show that the WhiB2 subfamily has a broad mix of members from Actinobacteria (33%) and actinobacteriophages (67%), whereas the other four major subfamilies are dominated by members from Actinobacteria, with relatively few from actinobacteriophages. The 24 remaining subfamilies are much smaller in size, and the majority (19 out of 24) of them are dominated by Wbls from actinobacteriophages (Fig. 4B). The two remaining *Mtb* Wbls (WhiB5 and WhiB6) are grouped together with several unique Wbls from actinobacteriophages into one subfamily (labeled as No. 29), which is otherwise closely related with the WhiB7 subfamily. We randomly selected three Wbls from actinobacteriophages and belonging to the WhiB1 subfamily and two relatively larger species-specific Wbl subfamilies and found them all bind to $\sigma^A{}_4$ in a His516-dependent manner in the pull-down assays (SI Appendix, Fig. S5), as observed for the Wbls from *Mtb* (25,32,33).

As expected, the primary sequence alignment shows that Wbls share little in common besides the two known motifs (the four [4Fe-4S] cluster-ligating Cys and the β-turn), among which include the aromatic patch found in *Mtb* Wbls for $\sigma^A{}_4$ binding (Fig. 5). Considering the low sequence identity among the Wbl members, we generated AlphaFold structural models for these Wbl members and quantified the conservation of the aromatic patch in the Wbl family by 3D structural alignment against $\sigma^A{}_4$-bound WhiB1 (PDB ID:6ONO) (see in Materials and Methods). As shown in SI Appendix, Table S3, among the 995 sequences examined, 984 out of the 995 Wbls (98.7%) possess at least two aromatic residues (either Trp, Phe, Tyr or His) corresponding to AP1-4 (W3, F17, F18 and W60 in WhiB1) and at least one corresponding to AP2-3 (F17 and F18 in WhiB1). Moreover, eight out of these 11 outliers possess an aromatic patch similar to that observed in the

Mtb Wbl proteins when considering the aromatic residues next to the aromatic patch at the relative -1 to 1 position, while the remaining three lack an aromatic patch as defined (SI Appendix, Table S3, Fig. S6). These findings indicate the newly identified aromatic patch is widely conserved in the Wbl family. During the structural comparison, we also found three Wbls with only three Cys at the expected [4Fe-4S] cluster binding pocket in the AlphaFold models. Two of these Wbl sequences (UniProt IDs: A0A1Q9SLC3 and A0A1Q9SWN8) are obsolete, while the third one (UniProt ID: A0A1D8EX64) from *Mycobacterium phage Tortellini* misses a Cys corresponding to the first Fe-S cluster-ligating Cys (corresponding to C9 in WhiB1) (SI Appendix, Fig. S6). Blast search shows that A0A1D8EX64 shares 99% sequence identity with another Wbl protein (UniProt ID: A0A2U8UHK0) from *Mycobacterium phage Xavia*, which has an additional Cys in the N-terminus corresponding to C9 of WhiB1. Therefore, it is unclear whether these Wbl outliers are due to the uncertainty of gene sequencing or structural prediction, or they represent true evolutional variants of the Wbl family.

Discussion

Wbl family proteins have attracted a great deal of attention because of their critical roles in the many biological processes in Actinobacteria that are of significant biomedical and economic implications, and their unique mode of action as the first family of monomeric Fe-S transcription factors. However, the mechanism of action of the Wbl family proteins remains enigmatic due to the lack of conserved functional motifs other than the [4Fe-4S] cluster binding motif and the β-turn. Herein, we provide structural, molecular, and biochemical evidence indicating that all *Mtb* Wbls use an aromatic patch to bind to the same site on $\sigma^A{}_4$ in the RNAP. Moreover, our bioinformatic analysis and structural modeling indicate that the aromatic patch is highly conserved among the Wbl family members, representing a previously unrecognized motif in the Wbl family. This aromatic patch uncovered in the Wbl family provides crucial insights into the mechanism of action of how Wbl proteins interact with $\sigma^A{}_4$ to regulate the genes involved in diverse biological processes.

Our study demonstrates that Wbl proteins use a conserved aromatic patch to anchor onto $\sigma^A{}_4$ with nanomolar binding affinity. This unusual hydrophobic residue-driven interaction between Wbl proteins and $\sigma^A{}_4$ is in stark contrast to the previously characterized $\sigma^{70}{}_4$-dependent transcription activators (*i.e*. the cyclic AMP receptor protein [CRP] and the fumarate and nitrate reduction regulatory protein [FNR]) that interacts with $\sigma^{70}{}_4$ via surface charge complementarity (5,13,43). Rather, it resembles that of the "phage appropriator" AsiA of the bacteriophage T4, which hijacks the host RNAP via binding to $\sigma^{70}{}_4$ and co-activates phage genes with another phage regulator MotA (44-46). The *wbl* genes are believed to have been laterally distributed in Actinobacteria (47), while their origin remains unknown. Broad analyses of the Wbl family proteins indicate that some members are widespread among Actinobacteria, whereas others are more sporadically distributed (2). The variability in the distribution of multiple Wbl paralogs suggests that horizontal gene transfer (HGT) may have played a role in their spread, and these transfer events were possibly mediated by phages, which also harbor Wbl genes (2,24,47). In fact, the Wbl family proteins were found to be among the most abundant regulator proteins encoded by actinobacteriophage genomes, and phylogenetic analysis revealed a complex evolutionary relationship of Wbl homologs between actinobacteria and their associated phages and plasmids, supporting a history of HGT (48). To further explore the evolutionary relationships among the bacterial and phage Wbl gene family members, our phylogenetic analysis included many additional phage sequences (Fig. 4B). Like previous studies, our tree shows a complex clustering

of phage and bacterial Wbl homologs, lending further support for HGT of homologs between phage and bacteria. Notably, in many of the major clades, the earliest diverging clades are often dominated by phage sequences, while the bacterial clades often occupy more derived clades. This pattern suggests that many of these homologs were transferred from viruses to bacteria, supporting the notion that the phages may have served as vectors for the HGT of Wbl proteins among Actinobacteria. This finding implies that bacterial hosts may have sequestered the *wbl* gene from actinobacteriophages and repurposed the function of the genes for their own benefits during the early evolutionary events, followed by redistribution of the *wbl* genes among the Actinobacteria via phage-mediated HGT, in line with the previous hypothesis by Weinbauer *et al.* (49). Conversely, the phylogenetic analysis also identifies some phage Wbl members nested within otherwise bacterial clades, providing clear examples of HGT from bacteria to phage and underscoring the complex interplay of HGT in spreading *wbl* genes between Actinobacteria and their related phages.

The findings from our study raise a new question about how Wbls regulate gene expression via binding to σ^A in the RNAP, while most of them lack a recognizable DNA binding motif. Recent studies on *Sve* WhiB, a homolog of *Mtb* WhiB2, offer a working model for how other Wbl proteins without a canonical DNA binding motif may regulate gene expression via binding to σ^A (14,31,50). These studies show that *Sve* WhiB co-activates the genes involved in cell division and sporulation via binding to σ^A in the RNAP and another regulator WhiA. The single-particle cryo-electron microscopy (cryoEM) structure of the WhiA/WhiB-dependent transcription initiation complex shows that *Sve* WhiB acts as a mediator that binds to WhiA on one side and σ^A on the other side (31). The WhiA and $\sigma^A{}_4$ in the RNAP, joined by WhiB2, define promoter specificity of the WhiA/WhiB2-dependent genes. *Sve* WhiB only shows non-specific DNA contacts, explaining the absence of the recognizable DNA binding motif in the WhiB2 subfamily. Further molecular and structural characterization of the Wbl-dependent transcriptional machinery is needed for an in-depth understanding of the diverse roles of Wbl proteins in stress-responsive transcriptional regulation.

Materials and Methods

Detailed descriptions of the following are provided in SI Appendix: Constructions of strains and plasmids for expression and purification of Wbl proteins and $\sigma^A{}_4$ used in this study. A complete list of oligonucleotides, strains, and plasmids used is provided in SI Appendix, Table S1; protein purification methods and pull-down assays; crystallization of the WhiB6:$\sigma^A{}_{C82}$-β_{tip} complex; X-ray crystallographic data collection, structural determination, and analysis; isothermal titration calorimetry (ITC) measurement; mycobacterial protein fragment complementation (M-PFC) assay; bioinformatic analysis of the Wbl family proteins; 3D structural modeling and alignment of the Wbl proteins data visualization.

Data Availability

Atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank (PDB) under the accession code 8D5V for the WhiB6:$\sigma^A{}_4$-β_{tip} complex, and released for public access with the validation report.

The phylogenetic trees are available at iTOL through the following links:

https://itol.embl.de/tree/1299316123022799174675188 6 (with highlights of the 29 Wbl subfamilies)

https://itol.embl.de/tree/1299316123023080174675195 3 (with highlights of the Wbl proteins from Actinobacteriophages)

Supplemental Information

Supplemental documents can be found online at the journal website.

Acknowledgments

The authors thank Dr. Mark A. Wilson at the University of Nebraska-Lincoln for the insightful discussions. We thank the staff at beamlines 12-2 of SSRL for their assistance during the X-ray diffraction data collection. The authors would like to acknowledge the following funding sources: This work was supported by grants from the National Institutes of Health (R35 GM138157) and the National Science Foundation CAREER Award (CLP 1846908 for the work related to WhiB1) to L-M.Z., and the NIH (T32 GM136593) to D.B.G through the T32 Predoctoral Training Fellowship. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and the National Science Foundation. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393).

Author contributions

L-M.Z. directed the project. D.B.G, T.W., C. J. S. L. and M.H. performed molecular biology work, expressed, purified, and characterized the proteins. T.W. carried out the crystallographic study. J.S. and D.B.G conducted the ITC assay, with assistance from D.B. Y.P., H.Y, D.B.G and L-M.Z. carried out the structural modeling and structural analysis. Q.Y. and P. M. conducted the phylogenic analysis, J.P.M. provided expertise in the phylogenetic tree interpretation. A. S. provided the plasmids used for M-PFC assays. D.G.B, T.W and L-M.Z. wrote the manuscript. All authors contributed to the manuscript preparation.

Conflict of interest

The authors declare no competing interests.

References

1. Chater, K. F. (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. *J Gen Microbiol* **72**, 9-28
2. Chandra, G., and Chater, K. F. (2014) Developmental biology of Streptomyces from the perspective of 100 actinobacterial genome sequences. *FEMS Microbiol Rev* **38**, 345-379
3. Davis, N. K., and Chater, K. F. (1992) The *Streptomyces coelicolor* whiB gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet* **232**, 351-358
4. Molle, V., Palframan, W. J., Findlay, K. C., and Buttner, M. J. (2000) WhiD and WhiB, homologous proteins required for different stages of sporulation in *Streptomyces coelicolor A3(2)*. *J Bacteriol* **182**, 1286-1295
5. Steyn, A. J., Collins, D. M., Hondalus, M. K., Jacobs, W. R., Jr., Kawakami, R. P., and Bloom, B. R. (2002) *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect

host survival but is dispensable for *in vivo* growth. *Proc Natl Acad Sci USA* **99**, 3147-3152
6. Morris, R. P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D., Ehrt, S., Liu, Y., Heifets, L., Pieters, J., Schoolnik, G., and Thompson, C. J. (2005) Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* **102**, 12200-12205
7. Geiman, D. E., Raghunand, T. R., Agarwal, N., and Bishai, W. R. (2006) Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven *Mycobacterium tuberculosis* whiB-like genes. *Antimicrob Agents Chemother* **50**, 2836-2841
8. Singh, A., Guidry, L., Narasimhulu, K. V., Mai, D., Trombley, J., Redding, K. E., Giles, G. I., Lancaster, J. R., Jr., and Steyn, A. J. (2007) *Mycobacterium tuberculosis* WhiB3 responds to O_2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proc Natl Acad Sci USA* **104**, 11562-11567
9. Smith, L. J., Stapleton, M. R., Fullstone, G. J., Crack, J. C., Thomson, A. J., Le Brun, N. E., Hunt, D. M., Harvey, E., Adinolfi, S., Buxton, R. S., and Green, J. (2010) *Mycobacterium tuberculosis* WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. *Biochem J* **432**, 417-427
10. Chawla, M., Parikh, P., Saxena, A., Munshi, M., Mehta, M., Mai, D., Srivastava, A. K., Narasimhulu, K. V., Redding, K. E., Vashi, N., Kumar, D., Steyn, A. J., and Singh, A. (2012) *Mycobacterium tuberculosis* WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. *Mol Microbiol* **85**, 1148-1165
11. Larsson, C., Luna, B., Ammerman, N. C., Maiga, M., Agarwal, N., and Bishai, W. R. (2012) Gene expression of *Mycobacterium tuberculosis* putative transcription factors *whiB1-7* in redox environments. *PLoS One* **7**, e37516
12. Casonato, S., Cervantes Sanchez, A., Haruki, H., Rengifo Gonzalez, M., Provvedi, R., Dainese, E., Jaouen, T., Gola, S., Bini, E., Vicente, M., Johnsson, K., Ghisotti, D., Palu, G., Hernandez-Pando, R., and Manganelli, R. (2012) WhiB5, a transcriptional regulator that contributes to *Mycobacterium tuberculosis* virulence and reactivation. *Infect Immun* **80**, 3132-3144
13. Burian, J., Yim, G., Hsing, M., Axerio-Cilies, P., Cherkasov, A., Spiegelman, G. B., and Thompson, C. J. (2013) The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). *Nucleic Acids Res* **41**, 10062-10076
14. Bush, M. J., Chandra, G., Bibb, M. J., Findlay, K. C., and Buttner, M. J. (2016) Genome-wide chromatin immunoprecipitation sequencing analysis shows that WhiB is a transcription factor that cocontrols its regulon with WhiA to initiate developmental cell division in *Streptomyces*. *mBio* **7**, e00523-00516
15. Chen, Z., Hu, Y., Cumming, B. M., Lu, P., Feng, L., Deng, J., Steyn, A. J., and Chen, S. (2016) Mycobacterial WhiB6 differentially regulates ESX-1 and the Dos regulon to modulate granuloma formation and virulence in Zebrafish. *Cell Rep* **16**, 2512-2524
16. Bosserman, R. E., Nguyen, T. T., Sanchez, K. G., Chirakos, A. E., Ferrell, M. J., Thompson, C. R., Champion, M. M., Abramovitch, R. B., and Champion, P. A. (2017) WhiB6 regulation of ESX-1 gene expression is controlled by a negative feedback loop in *Mycobacterium marinum*. *Proc Natl Acad Sci USA* **114**, E10772-E10781

17. Wu, J., Ru, H. W., Xiang, Z. H., Jiang, J., Wang, Y. C., Zhang, L., and Liu, J. (2017) WhiB4 Regulates the PE/PPE Gene Family and is Essential for Virulence of Mycobacterium marinum. *Sci Rep* **7**, 3007
18. Lee, D. S., Kim, P., Kim, E. S., Kim, Y., and Lee, H. S. (2018) *Corynebacterium glutamicum* WhcD interacts with WhiA to exert a regulatory effect on cell division genes. *Antonie Van Leeuwenhoek* **111**, 641-648
19. Chawla, M., Mishra, S., Anand, K., Parikh, P., Mehta, M., Vij, M., Verma, T., Singh, P., Jakkala, K., Verma, H. N., AjitKumar, P., Ganguli, M., Narain Seshasayee, A. S., and Singh, A. (2018) Redox-dependent condensation of the mycobacterial nucleoid by WhiB4. *Redox Biol* **19**, 116-133
20. Lin, C., Tang, Y., Wang, Y., Zhang, J., Li, Y., Xu, S., Xia, B., Zhai, Q., Li, Y., Zhang, L., and Liu, J. (2022) WhiB4 Is Required for the Reactivation of Persistent Infection of *Mycobacterium marinum* in Zebrafish. *Microbiol Spectr* **10**, e0044321
21. Guiza Beltran, D., Wan, T., and Zhang, L. (2024) WhiB-like proteins: Diversity of structure, function and mechanism. *Biochim Biophys Acta Mol Cell Res* **1871**, 119787
22. Bush, M. J. (2018) The actinobacterial WhiB-like (Wbl) family of transcription factors. *Mol Microbiol* **110**, 663-676
23. Burian, J., Ramon-Garcia, S., Sweet, G., Gomez-Velasco, A., Av-Gay, Y., and Thompson, C. J. (2012) The mycobacterial transcriptional regulator *whiB7* gene links redox homeostasis and intrinsic antibiotic resistance. *J Biol Chem* **287**, 299-310
24. Saini, V., Farhana, A., and Steyn, A. J. (2012) *Mycobacterium tuberculosis* WhiB3: a novel iron-sulfur cluster protein that regulates redox homeostasis and virulence. *Antioxid Redox Signal* **16**, 687-697
25. Wan, T., Horova, M., Beltran, D. G., Li, S., Wong, H. X., and Zhang, L. M. (2021) Structural insights into the functional divergence of WhiB-like proteins in *Mycobacterium tuberculosis*. *Mol Cell* **81**, 2887–2900
26. Lilic, M., Darst, S. A., and Campbell, E. A. (2021) Structural basis of transcriptional activation by the *Mycobacterium tuberculosis* intrinsic antibiotic-resistance transcription factor WhiB7. *Mol Cell* **81**, 2875-2886 e2875
27. Gomez, M., Doukhan, L., Nair, G., and Smith, I. (1998) sigA is an essential gene in Mycobacterium smegmatis. *Molecular Microbiology* **29**, 617-628
28. Lee, D. J., Minchin, S. D., and Busby, S. J. (2012) Activating transcription in bacteria. *Annu Rev Microbiol* **66**, 125-152
29. Feng, L., Chen, Z., Wang, Z., Hu, Y., and Chen, S. (2016) Genome-wide characterization of monomeric transcriptional regulators in *Mycobacterium tuberculosis*. *Microbiology* **162**, 889-897
30. Kudhair, B. K., Hounslow, A. M., Rolfe, M. D., Crack, J. C., Hunt, D. M., Buxton, R. S., Smith, L. J., Le Brun, N. E., Williamson, M. P., and Green, J. (2017) Structure of a Wbl protein and implications for NO sensing by *M. tuberculosis*. *Nat Commun* **8**, 2280
31. Lilic, M., Holmes, N. A., Bush, M. J., Marti, A. K., Widdick, D. A., Findlay, K. C., Choi, Y. J., Froom, R., Koh, S., Buttner, M. J., and Campbell, E. A. (2023) Structural basis of dual activation of cell division by the actinobacterial transcription factors WhiA and WhiB. *Proc Natl Acad Sci USA* **120**, e2220785120
32. Wan, T., Horova, M., Khetrapal, V., Li, S., Jones, C., Schacht, A., Sun, X., and Zhang, L. (2023) Structural basis of DNA binding by the WhiB-like transcription factor WhiB3 in *Mycobacterium tuberculosis*. *The Journal of biological chemistry* **299**, 104777

33. Wan, T., Li, S., Beltran, D. G., Schacht, A., Zhang, L., Becker, D. F., and Zhang, L. (2020) Structural basis of non-canonical transcriptional regulation by the σ^A-bound iron-sulfur protein WhiB1 in *M. tuberculosis*. *Nucleic Acids Res* **48**, 501-516
34. Abdallah, A. M., Weerdenburg, E. M., Guan, Q., Ummels, R., Borggreve, S., Adroub, S. A., Malas, T. B., Naeem, R., Zhang, H., Otto, T. D., Bitter, W., and Pain, A. (2019) Integrated transcriptomic and proteomic analysis of pathogenic mycobacteria and their esx-1 mutants reveal secretion-dependent regulation of ESX-1 substrates and WhiB6 as a transcriptional regulator. *PLoS One* **14**, e0211003
35. Bosserman, R. E., Nguyen, T. T., Sanchez, K. G., Chirakos, A. E., Ferrell, M. J., Thompson, C. R., Champion, M. M., Abramovitch, R. B., and Champion, P. A. (2017) WhiB6 regulation of ESX-1 gene expression is controlled by a negative feedback loop in Mycobacterium marinum. *Proc Natl Acad Sci U S A* **114**, E10772-E10781
36. Solans, L., Aguilo, N., Samper, S., Pawlik, A., Frigui, W., Martin, C., Brosch, R., and Gonzalo-Asensio, J. (2014) A specific polymorphism in Mycobacterium tuberculosis H37Rv causes differential ESAT-6 expression and identifies WhiB6 as a novel ESX-1 component. *Infect Immun* **82**, 3446-3456
37. Sala, C., Odermatt, N. T., Soler-Arnedo, P., Gulen, M. F., von Schultz, S., Benjak, A., and Cole, S. T. (2018) EspL is essential for virulence and stabilizes EspE, EspF and EspH levels in Mycobacterium tuberculosis. *PLoS Pathog* **14**, e1007491
38. Vijayaraj, M., Abhinand, P. A., and Ragunath, P. K. (2019) Virtual screening of a MDR-TB WhiB6 target identified by gene expression profiling. *Bioinformation* **15**, 557-567
39. Wan, T., Horova, M., Beltran, D. G., Li, S., Wong, H. X., and Zhang, L. M. (2021) Structural insights into the functional divergence of WhiB-like proteins in Mycobacterium tuberculosis. *Mol Cell* **81**, 2887-2900 e2885
40. Blanco, A. G., Canals, A., Bernues, J., Sola, M., and Coll, M. (2011) The structure of a transcription activation subcomplex reveals how σ^{70} is recruited to PhoB promoters. *EMBO J* **30**, 3776-3785
41. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* **9**, 527-539
42. Singh, A., Mai, D., Kumar, A., and Steyn, A. J. (2006) Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. *Proc Natl Acad Sci USA* **103**, 11346-11351
43. Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J Mol Biol* **284**, 1353-1365
44. Lambert, L. J., Wei, Y., Schirf, V., Demeler, B., and Werner, M. H. (2004) T4 AsiA blocks DNA recognition by remodeling σ^{70} region 4. *EMBO J* **23**, 2952-2962
45. Gregory, B. D., Nickels, B. E., Garrity, S. J., Severinova, E., Minakhin, L., Urbauer, R. J. B., Urbauer, J. L., Heyduk, T., Severinov, K., and Hochschild, A. (2004) A regulator that inhibits transcription by targeting an intersubunit interaction of the RNA polymerase holoenzyme. *Proc Natl Acad Sci USA* **101**, 4554-4559
46. Shi, J., Wen, A., Zhao, M., You, L., Zhang, Y., and Feng, Y. (2019) Structural basis of sigma appropriation. *Nucleic Acids Res* **47**, 9423-9432
47. Chater, K. F., and Chandra, G. (2006) The evolution of development in Streptomyces analysed by genome comparisons. *FEMS Microbiol Rev* **30**, 651-672

48. Sharma, V., Hardy, A., Luthe, T., and Frunzke, J. (2021) Phylogenetic Distribution of WhiB- and Lsr2-Type Regulators in Actinobacteriophage Genomes. *Microbiol Spectr* **9**, e0072721
49. Weinbauer, M. G., and Rassoulzadegan, F. (2004) Are viruses driving microbial diversification and diversity? *Environ Microbiol* **6**, 1-11
50. Bush, M. J., Bibb, M. J., Chandra, G., Findlay, K. C., and Buttner, M. J. (2013) Genes Required for Aerial Growth, Cell Division, and Chromosome Segregation Are Targets of WhiA before Sporulation in *Streptomyces venezuelae*. *mBio* **4**

FIGURE LEGENDS

Figure 1. The aromatic patch facilitates *Mtb* WhiB6 binding to $\sigma^A{}_4$. (A) Sequence alignment of the seven *Mtb* Wbl proteins. The two previously identified motifs in the Wbl family, including the four conserved cysteines (the black triangle) in the [4Fe-4S] cluster binding motif and the Gly-rich β-turn, are highlighted. The assignment of the secondary structures (helices *h1-h4*, loop *l1-l2*, β-turn) is based on the $\sigma^A{}_4$-bound WhiB1 (PDB ID:6ONO). The residues in the aromatic patch in *Mtb* Wbls corresponding to W3, F17, F18, and W60 in WhiB1 for binding to $\sigma^A{}_4$ are highlighted in red fonts within the black frame. **(**B) Cartoon representations of the WhiB6:$\sigma^A{}_4$-β_{tip} complex. (C-D) Structural overlay of $\sigma^A{}_4$-bound WhiB6 (pale green) with WhiB1 (pink, PDB ID:6ONO) and a close-up view of the Wbl:$\sigma^A{}_4$ interface around the Fe-S cluster binding pocket. $\sigma^A{}_4$ and the [4Fe-4S] cluster of WhiB1 are not shown for clarity. In all structures, the helices *h1-4* and β-turn are indicated. (E and F) UV-visible spectra and SDS-PAGE analyses of the pull-down samples of WhiB6 and $\sigma^A{}_4$-β_{tip} (either wildtype [WT] or mutant as indicated). The color of the UV-visible spectra matches that of the sample labels. The intensity of the absorption peak around 410 nm, highlighted with dashed lines, is indicative of the occupancy of the [4Fe-4S] cluster in the purified samples. The absorption spectra were normalized at 280 nm. The uncropped SDS PAGE images are shown in Fig. S3.

Figure 2. *Mtb* WhiB5 binds to $\sigma^A{}_4$ with high affinity. (A) Detection of the interactions between *Mtb* WhiB5 and $\sigma^A{}_4$ in *Mycobacterium smegmatis* (*Msm*) by Mycobacterial Protein Fragment Complementation (M-PFC) assay. The *Msm* strain, labeled as WhiB5:$\sigma^A{}_4$-β_{tip}, expressing *Mtb* WhiB5 fused to the N-terminus of the murine DHFR fragment ($mDHFR_{F[1,2]}$) and $\sigma^A{}_{C82}$-β_{tip} fused to the other half of murine DHFR ($mDHFR_{F[3]}$) was used the test the interaction between the two proteins on the 7H10 broth agar plate in the absence or presence of 5 μg/ml trimethoprim (TRIM) (see in Materials and Methods). Cell growth in the presence of TRIM is indicative of the protein-protein interaction. Positive control, the *Msm* strain expressing the *Saccharomyces cerevisiae* leucine-zipper sequence homodimerization domain GCN4 fused to the N terminus of $mDHFR_{F[1,2]}$ and $mDHFR_{F[3]}$, respectively; Negative control, the *Msm* strain expressing the two fragments $mDHFR_{F[1,2]}$ and $mDHFR_{F[3]}$ separately. (B-D) Representative results for Isothermal titration calorimetry (ITC) assay of the interactions between $\sigma^A{}_4$ and *Mtb* Wbls as indicated. The calorimetric titration is shown in the top panel, with the integrated injection heat derived from the titration and the best-fit curve (dash lines) of a simple 1:1 interaction model shown in the bottom

panel. (E) Summary of the binding parameters from the best fit, including the association constant (K_a), dissociation constant (K_d), changes in enthalpy (ΔH) and entropy (ΔS) upon binding, and the goodness of the fit (Chi^2/DoF).

Figure 3. Role of the aromatic patch in *Mtb* WhiB4 and WhiB5 for $\sigma^A{}_4$ binding. (A-D) UV-visible spectra and SDS-PAGE analysis of the samples from co-expression and purification to test the $\sigma^A{}_4$ binding by *Mtb* WhiB4 and WhiB5 (either wildtype or mutant as indicated). For Panels A, B, and D, His_6-$\sigma^A{}_{C170}$ (either wildtype pr mutant as indicated) was used as the bait for the pull-down assays with the tagless *Mtb* WhiB4 (wildtype) and WhiB5 (either wildtype or mutant as indicated), respectively. For Panel C, WhiB4-His_6 (either wildtype or mutant as indicated) was used as the bait to test the interaction with Strep-$\sigma^A{}_4\beta_{tip}$. The absorption spectra were normalized at 280 nm. The colors of the UV-Vis spectra match that of the sample labels. The uncropped SDS PAGE images are shown in Fig. S3.

Figure 4. Phylogenetic analysis of the Wbl proteins from the representative genomes of Actinobacteria and actinobacteriophages. The unrooted phylogenetic tree of the 995 Wbl family proteins from the representative genomes of Actinobacteria and actinobacteriophages using the Pfam PF02467 (RP15), with highlights of the 29 Wbl subfamilies (A) and the Wbls from actinobacteriophages in cyan (B). The five major Wbl subfamilies are labeled with the representative Wbls in *Mtb* (WhiB1-4 and WhiB7, respectively). The remaining smaller subfamilies are color-coded and assigned sequential numbers in the tree.

Figure 5. Conservation of residues across the Wbl proteins from the representative genomes of Actinobacteria and actinobacteriophages. The sequence logo of the 995 Wbl family proteins from the representative genomes of Actinobacteria and actinobacteriophages (Pfam PF02467, RP15) was generated by WebLogo.

Figures

Figure 1

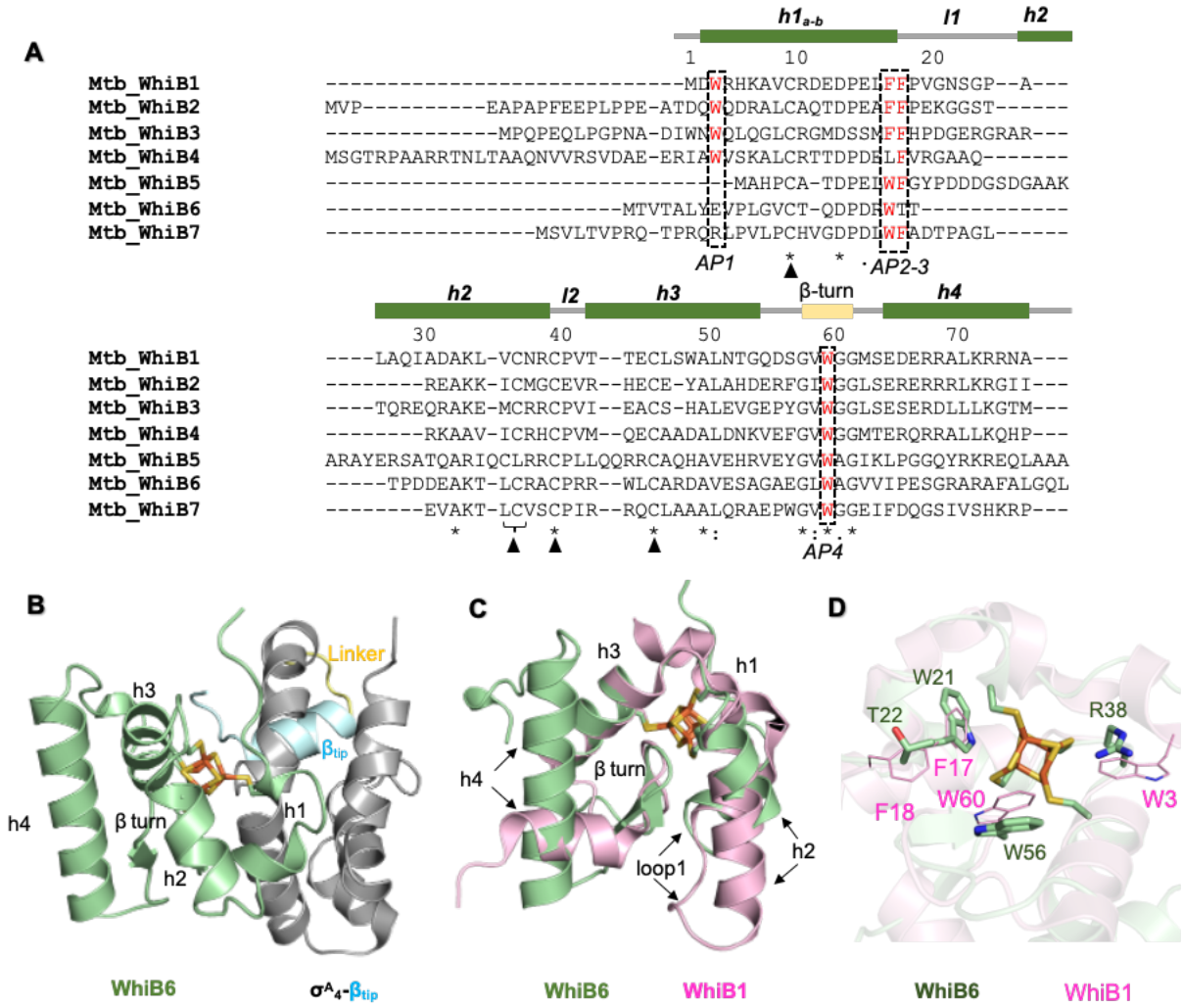
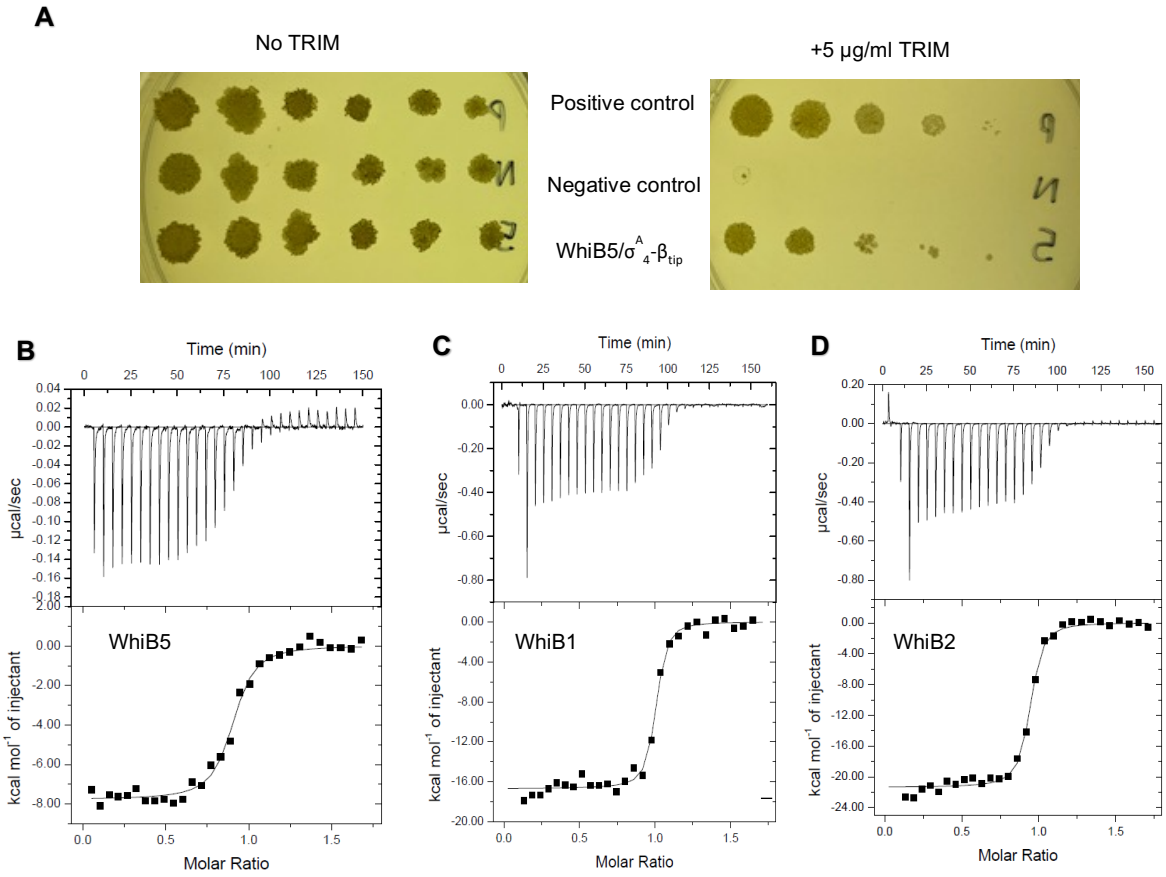


Figure 2



E

Sample	Ka (x10^{7} M^{-1})	K_d (nM)	ΔH (kcal/mol)	−TΔS (kcal/mol)	n	Chi²/DoF
WhiB1	8.9 ± 2.3	12.0± 3.0	-16.68 ± 0.18	7.18	0.981 ± 0.005	4.11E5
WhiB2	6.1 ± 1.1	16.9± 3.0	-21.36 ± 0.19	13.39	0.919 ± 0.004	3.804E5
WhiB5	2.1 ± 0.4	50.3± 10.0	-7.76 ± 0.10	2.77	0.884 ± 0.007	1.008E5

Figure 3

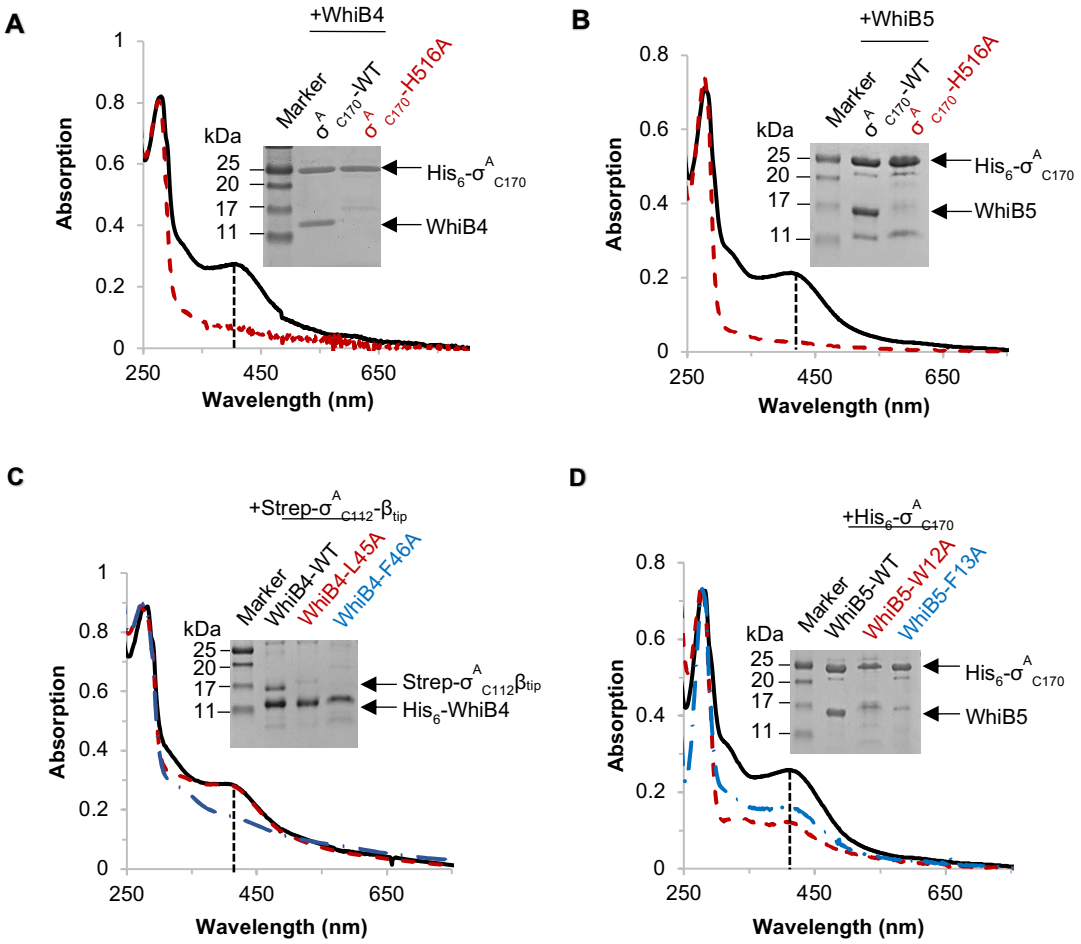
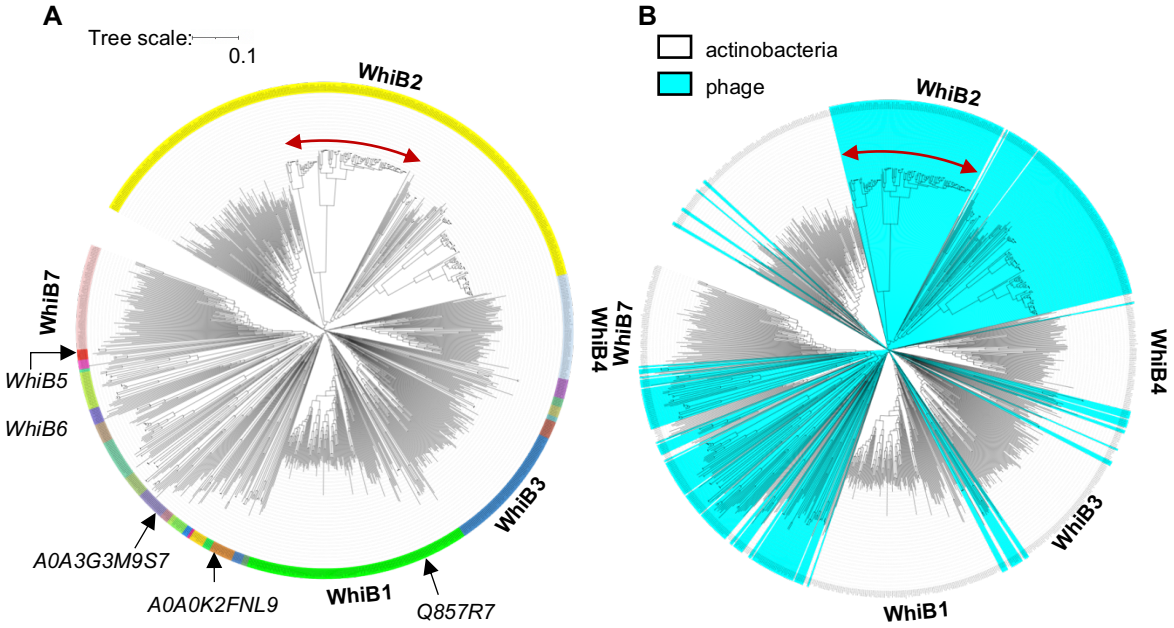

A
+WhiB4
Absorption
Wavelength (nm)
kDa
Marker
σA C170-WT
σA C170-H516A
His6-σA C170
WhiB4
B
+WhiB5
Absorption
Wavelength (nm)
kDa
Marker
σA C170-WT
σA C170-H516A
His6-σA C170
WhiB5
C
+Strep-σA C112-βtip
Absorption
Wavelength (nm)
kDa
Marker
WhiB4-WT
WhiB4-L45A
WhiB4-F46A
Strep-σA C112βtip
His6-WhiB4
D
+His6-σA C170
Absorption
Wavelength (nm)
kDa
Marker
WhiB5-WT
WhiB5-W12A
WhiB5-F13A
His6-σA C170
WhiB5

Figure 4


A
Tree scale:
0.1
WhiB2
WhiB7
WhiB5
WhiB6
WhiB3
A0A3G3M9S7
A0A0K2FNL9
WhiB1
Q857R7
B
actinobacteria
phage
WhiB2
WhiB4
WhiB7
WhiB4
WhiB3
WhiB1

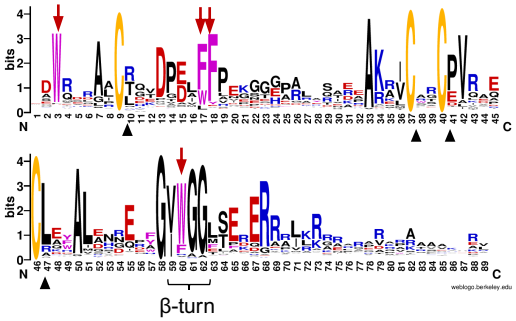



Figure 5