Genome-wide characterization of monomeric transcriptional regulators in *Mycobacterium tuberculosis*

Lipeng Feng, Zhenkang Chen, Zhongwei Wang, Yangbo Hu and Shiyun Chen

Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, PR China

Gene transcription catalysed by RNA polymerase is regulated by transcriptional regulators, which play central roles in the control of gene transcription in both eukaryotes and prokaryotes. In regulating gene transcription, many regulators form dimers that bind to DNA with repeated motifs. However, some regulators function as monomers, but their mechanisms of gene expression control are largely uncharacterized. Here we systematically characterized monomeric versus dimeric regulators in the tuberculosis causative agent *Mycobacterium tuberculosis*.

Of the >160 transcriptional regulators annotated in *M. tuberculosis*, 154 transcriptional regulators were tested, 22% probably act as monomers and most are annotated as hypothetical regulators. Notably, all members of the WhiB-like protein family are classified as monomers. To further investigate mechanisms of monomeric regulators, we analysed the actions of these WhiB proteins and found that the majority interact with the principal sigma factor σ^A^, which is also a monomeric protein within the RNA polymerase holoenzyme. Taken together, our study for the first time globally classified monomeric regulators in *M. tuberculosis* and suggested a mechanism for monomeric regulators in controlling gene transcription through interacting with monomeric sigma factors.

INTRODUCTION

Transcriptional regulators coordinately function as monomers or dimers in controlling bacterial responses to environmental changes. After infection, *Mycobacterium tuberculosis* can either cause disease immediately or persist in host tissues for years in a dormant state (Gengenbacher & Kaufmann, 2012). Transcriptional regulation of gene expression is a critical feature that allows *M. tuberculosis* to adapt to the hostile environments inside host cells and also control the transitioning between different infection states (Manganelli et al., 2004; Gengenbacher & Kaufmann, 2012). Complete genome sequencing of *M. tuberculosis* has led to important advances in understanding modes of transcriptional regulation in this deadly pathogen (Cole et al., 1998; Guo et al., 2009; Peterson et al., 2014). However, most of the >160 transcriptional regulators annotated in the *M. tuberculosis* genome are largely uncharacterized at the functional level.

Dimerization of transcriptional regulators plays important roles in transcriptional regulation. For example, the two-component regulators PhoP and DevR form homodimers in binding with DNA upstream of the target promoter region to activate gene transcription in *M. tuberculosis* (Florczyk et al., 2003; He & Wang, 2014). Transcriptional repressors also mostly bind to DNA as dimers, and the binding region is always located around the RNA polymerase binding region (van Hijum et al., 2009). In these cases, the binding of a repressor therefore abolishes the binding or mobility of RNA polymerase on template DNA. For example, the repressor MogR in *Listeria monocytogenes* was characterized as binding with an inverse repeated AT-rich motif (T^4^-N^2^-A^4^) that overlapped with the −35 and −10 regions on the flaA promoter, and this repressed gene transcription (Shen et al., 2009). Based on the fact that dimeric regulators bind to either inverted or direct repeat DNA sequences, bioinformatics analysis of promoter sequence has been widely used to predict the locations of conserved binding sequences targeted by potential transcriptional regulators (Zhang et al., 2010).
In contrast to the vast numbers of known dimeric regulators, some transcriptional regulators bind to DNA lacking a characteristic repeated motif, suggesting that they may act as monomers. One example is WhiB7 in Mycobacteria, which is one of the WhiB-like iron–sulfur (Fe–S) cluster regulators. WhiB7 has been demonstrated to bind to an AT-rich motif upstream of the −35 region in its own promoter, and this motif does not carry any repeated sequence (Burian et al., 2012, 2013). Owing to limited information regarding the WhiB7 binding sequence, the prediction of other direct binding targets remains a challenge. Similarly, another WhiB family regulator, WhiB4, was demonstrated to bind to a GC-rich region of DNA with no obvious repeat motif, located upstream of the targeted promoter (Chawla et al., 2012). In the absence of a repeated binding sequence, it seems credible that these two regulators act as monomers rather than as dimers. M. tuberculosis contains seven WhiB proteins (WhiB1–WhiB7), which are expressed under different conditions (Geiman et al., 2006; Larsson et al., 2012); however, whether these proteins all act as monomers in regulating gene expression is unknown. Until now, the action of monomeric regulators in gene transcriptional regulation has remained unexplored.

In this study, a detection system was first developed to distinguish monomeric and dimeric proteins in M. tuberculosis, and the majority of monomeric regulators were globally characterized. About 22% of regulators were classified as monomers, including all of the WhiB proteins. As examples, the mechanisms of WhiB proteins in transcriptional regulation were further studied to explore the actions of monomeric regulators in controlling gene transcription. Six WhiB proteins were confirmed to be monomers (WhiB1–WhiB7), which are expressed under different conditions expressed under different conditions (Geiman et al., 2006; Larsson et al., 2012); however, whether these proteins all act as monomers in regulating gene expression is unknown. Until now, the action of monomeric regulators in gene transcriptional regulation has remained unexplored.

In this study, a detection system was first developed to distinguish monomeric and dimeric proteins in M. tuberculosis, and the majority of monomeric regulators were globally characterized. About 22% of regulators were classified as monomers, including all of the WhiB proteins. As examples, the mechanisms of WhiB proteins in transcriptional regulation were further studied to explore the actions of monomeric regulators in controlling gene transcription. Six WhiB proteins were confirmed to be monomers (WhiB1–WhiB7), which are expressed under different conditions (Geiman et al., 2006; Larsson et al., 2012); however, whether these proteins all act as monomers in regulating gene expression is unknown. Until now, the action of monomeric regulators in gene transcriptional regulation has remained unexplored.

**METHODS**

**Bacterial strains, growth media and oligonucleotides.** Bacterial strains and plasmids used in this study are summarized in Table 1. Escherichia coli cultures were routinely grown in Luria–Bertani (LB) medium (1% tryptone, 0.5% yeast extract and 1% NaCl) or on LB agar plates (LB medium with 1% agar) at 37 or 30 °C. Sequences of all oligonucleotides used in this study are listed in Table S1 (available in the online Supplementary Material).

**Plasmid constructions.** The N-terminal fragment of the mogR gene was amplified from the genomic DNA of L. monocytogenes and ligated onto an overexpression plasmid, pOVR200 (Peterson et al., 2014). The resulting plasmid was designated pMonT, and this contains multiple cloning sites upstream of the mogR insert for the cloning of additional DNA fragments. Each of the annotated transcriptional regulators was PCR amplified from M. tuberculosis H37Rv genomic DNA using the specific primer pairs listed in Table S1. The amplified DNA was digested with appropriate enzymes and ligated with T4 ligase into identically digested pMonT DNA.

**E. coli background strain construction.** To eliminate background β-galactosidase activity, the lacI2 gene from E. coli K-12 strain MG1655 was deleted by homologous recombination-dependent allelic exchange using the suicide plasmid pDM4-lacZm (Peterson et al., 2014). A plasmid named pZT-Rp, which contains the synthetic promoter carrying the MogR binding sites overlapping with the −10 box in the promoter region (sequence shown in Fig. 1b) upstream of the lacI2 gene, was transformed into the lacI2-deleted E. coli to obtain a strain named E. coli dZMT, which was used as the host strain for analysing the function of protein monomers.

**β-Galactosidase assay.** Constructs of pMonT were transformed into E. coli dZMT and spread onto LB agar plates with ampicillin (50 μg ml⁻¹) and kanamycin (50 μg ml⁻¹). After overnight culture at 30 °C, three colonies from each plate were separately cultured and transferred at 1% into fresh medium. After incubation at 30 °C for 4–5 h until the OD₆₀₀ reached ~0.8, the β-galactosidase assay was performed and enzyme activity was calculated as described previously (Hu et al., 2009). All assays were carried out at least in triplicate. Student’s t-test was used to calculate the degrees of significant difference between each group.

**RNA extraction and semiquantitative reverse transcriptase PCR (RT-PCR) analysis.** Two micro litres of bacterial cultures under conditions described for β-galactosidase assays were collected by centrifugation at 6000 g for 5 min. The pellets were ground in liquid nitrogen and RNA was extracted with TRIzol (Invitrogen) according to the manufacturer’s protocol. After treatment with RNase-free DNase I (Promega), 2 μg of each RNA sample was used in reverse transcription to obtain the cDNA template using Random 9-mer primers (TaKaRa). Semiquantitative PCR was performed using DNA polymerase with appropriate primers and cDNA template.

**Protein co-purification.** To overexpress non-tagged σ², the sigA gene containing the stop codon from M. tuberculosis H37Rv was amplified and inserted between the pET21a plasmid. whiB genes were separately cloned into the pET28a plasmid between the NdeI and HindIII sites in the pET21a plasmid. whiB genes were separately cloned into the pET28a plasmid between the NdeI and XhoI/SalI sites to ensure that WhiB proteins could be expressed with N-terminal histidine tags. For protein co-expression, pET21a carrying sigA was co-transformed with pET28a (used as a control) or pET28a carrying whiB genes into E. coli BL21(DE3). Protein purification was performed as described previously (Hu et al., 2014), with minor modifications. Briefly, protein expression was induced by addition of 1 mM IPTG at OD₆₀₀ ~ 0.6 and grown at 24 °C overnight. Cell pellets were collected by centrifugation and resuspended in a lysis buffer (20 mM Tris/HCl pH 7.9, 300 mM NaCl, 5% v/v glycerol, 0.2 mg ml⁻¹ lysozyme). Cells were lysed by ultrasonication, and the lysate was cleared by centrifugation (10 000 g, 30 min, 4 °C) and then loaded onto 1 ml Ni-NTA resin equilibrated with 20 mM Tris/HCl pH 7.9, 300 mM NaCl, 5% v/v glycerol. After washing with 10 ml 30 mM imidazole buffer (20 mM Tris/HCl pH 7.9, 300 mM NaCl, 5% glycerol, 30 mM imidazole), proteins on the resin were eluted with this buffer containing 300 mM imidazole.

**Bacterial adenylate cyclase-based two-hybrid (BACTH) assays.** The interactions between sigma factors and WhiB proteins were monitored by the BACTH system as described elsewhere (Karimova et al., 1998). In brief, the sigA gene (encoding σ²) or individual whiB genes were cloned into the pUT18 plasmid expressing T18-fused protein to generate T18 fusions, and individual whiB genes were cloned into the pKT25 plasmid expressing T25-fused protein to obtain various T25-WhiB family fusions. A mutated sigA gene was constructed using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene) following the manufacturer’s instructions. Pairs of pKT25 and pUT18 plasmids were co-transformed into the reporter strain E. coli BTH101. After 2 days of incubation at 30 °C, three
RESULTS AND DISCUSSION

Monomeric transcription regulators in M. tuberculosis

A system for identifying monomeric regulators

Testing for protein self-interaction is an efficient way to distinguish between monomeric and dimeric regulators. Similar to a previous system based on the repressive effects of IclR protein on its own promoter (Furuta et al., 2005), we developed a system for screening homodimerization status of proteins based on the repressive effects of dimeric MogR protein on flaA transcription (Shen et al., 2009). MogR is composed of two discrete domains connected by a flexible linker: the N-terminal DNA binding domain (aa 1–162) and a C-terminal zipper dimerization domain (Zip; aa 218–263) (Fig. 1a). MogR1–220 lacking the dimerization domain binds well to DNA containing the TTTT-N5-AAAA sequence, but it could not repress the transcription from a promoter containing this sequence (Shen et al., 2009). Hence, we synthesized a promoter containing the −35 and −10 boxes of the lac promoter and a TTTTTAAAAAAA region overlapping with the −10 box (Fig. 1b) upstream of the lacZ gene in the background strain E. coli dZMT, and constructed a plasmid, designated pMonT, with the purpose of overexpressing target proteins appended to the C terminus of MogR1–220 fragment. Thus, we proposed that, if the fused protein contained a dimerization domain, active MogR would be reconstituted if the fused protein functioned as a monomer, the lac promoter activity would not be repressed (Fig. 1c). On the other hand, if the fused protein functioned as a monomer, the lac promoter activity would not be repressed (Fig. 1c).

To validate the effectiveness of this system, we first concentrated on the well-characterized PhoP protein from M. tuberculosis as an example, which contains clearly defined modular N- and C-terminal domains, with only the N-terminal domain able to dimerize (Menon & Wang, 2011). We fused full-length PhoP or the two individual domains to the C terminus of MogR1–220. Consistent with previous data, full-length and N-terminal PhoP both contributed to the MogR-mediated repression of transcription of the lacZ reporter gene in the E. coli dZMT strain.

Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli dZMT</td>
<td>Background strain used for analysing monomeric proteins</td>
<td>This study</td>
</tr>
<tr>
<td>E. coli BTH101</td>
<td>Strain used for BACTH assay</td>
<td>Karimova et al. (1998)</td>
</tr>
<tr>
<td>E. coli DH5x</td>
<td>Strain used for constructing clones</td>
<td>Lab collection</td>
</tr>
<tr>
<td>E. coli BL21(DE3)</td>
<td>Strain used for protein overexpression</td>
<td>Novagen</td>
</tr>
<tr>
<td>M. smegnatis mc²155</td>
<td>Host strain used for MPFC assays</td>
<td>Lab collection</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pOV200</td>
<td>Gene overexpression plasmid, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Peterson et al. (2014)</td>
</tr>
<tr>
<td>pMonT</td>
<td>Plasmid for expressing MogR&lt;sub&gt;1–220&lt;/sub&gt; fused proteins, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pZT-Rp</td>
<td>Plasmid carrying MogR repressed promoter</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>upstream of lacZ gene, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pET21a</td>
<td>Protein expression plasmid, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET28a</td>
<td>Protein expression plasmid, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Novagen</td>
</tr>
<tr>
<td>pKT25</td>
<td>Plasmid for BACTH assay, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Karimova et al. (1998)</td>
</tr>
<tr>
<td>pUT18</td>
<td>Plasmid for BACTH assay, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Karimova et al. (1998)</td>
</tr>
<tr>
<td>pKT25-Zip</td>
<td>Control plasmid for BACTH assay, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Karimova et al. (1998)</td>
</tr>
<tr>
<td>pUT18-Zip</td>
<td>Control plasmid for BACTH assay, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Karimova et al. (1998)</td>
</tr>
<tr>
<td>pUAB100</td>
<td>oriM oriE hisp60-GCN4-Gly&lt;sub&gt;10&lt;/sub&gt;-mDHFR&lt;sub&gt;F1,2&lt;/sub&gt;, Hyg&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Singh et al. (2006)</td>
</tr>
<tr>
<td>pUAB200</td>
<td>oriE attP int hisp60-GCN4-Gly&lt;sub&gt;10&lt;/sub&gt;-mDHFR&lt;sub&gt;F3&lt;/sub&gt;, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Singh et al. (2006)</td>
</tr>
<tr>
<td>pUAB300</td>
<td>oriM oriE hisp60-mDHFR&lt;sub&gt;F1,2&lt;/sub&gt;-Gly&lt;sub&gt;10&lt;/sub&gt;-MCS, Hyg&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Singh et al. (2006)</td>
</tr>
<tr>
<td>pUAB400</td>
<td>oriE attP int hisp60-mDHFR&lt;sub&gt;F3&lt;/sub&gt;-Gly&lt;sub&gt;10&lt;/sub&gt;-MCS, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Singh et al. (2006)</td>
</tr>
</tbody>
</table>

*Amp<sup>R</sup>, Kan<sup>R</sup> and Hyg<sup>R</sup> represent resistance to ampicillin, kanamycin and hygromycin, respectively.

colonies were cultured in LB medium with 0.5 mM IPTG at 30 °C for 15 h. The β-galactosidase assays were performed as described above.

Mycobacterial protein fragment complementation (MPFC) assays. The interactions between WhiB proteins and σ<sup>E</sup> in Mycobacteria were assessed using MPFC (Singh et al., 2006). Briefly, the coding region of whiB genes was amplified separately and cloned into pUAB300 by the One Step Cloning kit (Vazyme), and the coding region of the WT or R515H-mutated sigA gene was amplified and cloned, respectively, into pUAB400 by the One Step Cloning kit. Plasmid pairs of pUAB300 and pUAB400 derivatives were respectively co-transformed into Mycobacterium smegmatis by electroporation. The plasmid pair pUAB100/pUAB200 was used as a positive control, and the pUAB300/pUAB400 pair was used as a negative control. The growth of bacteria in 7H9 medium in the presence of kanamycin (10 μg ml<sup>−1</sup>), hygromycin (25 μg ml<sup>−1</sup>) and trimethoprim (6.25 μg ml<sup>−1</sup>) was determined to analyse the interaction between WhiB proteins and σ<sup>E</sup> as described by others (Tao et al., 2013).
On the other hand, production of the C-terminal PhoP domain fusion did not support MogR-mediated repression, reflecting its monomeric state (Fig. 1d). These data support the successful design and construction of a MogR-based assay system that is suitable for monomeric protein detection. Critically, the AT-rich nature of the MogR binding box is expected to be quite rare in the M. tuberculosis genome, which boasts a GC content in the vicinity of 65.6%, suggesting that this system would be particularly useful for performing a global analysis of transcriptional regulators in this bacterium.

Distinguishing monomeric from dimeric regulators in M. tuberculosis

We next tried to analyse all transcriptional regulators in M. tuberculosis. Of the 160 annotated transcriptional regulators, 154 were successfully cloned into the MogR1–220 fusion expression vector. We compared the relative β-galactosidase activities of strains expressing these fused proteins with the strain expressing MogR1-220 fused with the C terminus of PhoP, which we established above as a positive control for monomeric protein status. As shown in Fig. 2, about 70% (109/154) of the clonable regulators repressed β-galactosidase activity by more than twofold (P<0.01) when expressed as a fusion with MogR1–220.

We believe an imposed threshold of twofold repression of β-galactosidase activity is an appropriate cut-off value (Fig. 2, Table S2), for it applies to the mycobacterial regulators KmsR (Campbell et al., 2007), ArgP (Zhou et al., 2010), Blal (Sala et al., 2009) and EspR (Gangwar et al., 2014), which have all independently been confirmed to form homodimers in vitro.

A small number of regulators (32/154) we characterized as active in monomeric form, since their production did not obviously inhibit the β-galactosidase activity (P>0.1). Most of these proposed monomeric regulators are annotated as hypothetical proteins with uncharacterized functions in gene regulation. Interestingly, all seven members of the WhiB protein family (WhiB1–WhiB7) are classified as monomeric proteins in our in vivo tests. Although previous work suggested some of the WhiB family regulators (such as WhiB2) aggregate in vitro (Alam et al., 2009), none of the WhiB family proteins has been reported to bind to DNA with a repeat motif, supporting our proposal that they may act as monomers. Critically, we used semiquantitative RT-PCR analysis to confirm that the failure to repress expression of the lacZ reporter by these characterized monomeric proteins was not due to poor transcription of the corresponding gene chimeras (Fig. 3).

To further test the reliability of our assay in predicting monomeric proteins, we next applied the BACTH system...
(Karimova et al., 1998) to test for WhiB protein homodimerization. In this assay, the domains T25 and T18 of adenylate cyclase were individually fused to the WhiB proteins. When fused proteins interact with each other, T25 and T18 domains form the catalytic enzyme to produce cAMP (Karimova et al., 1998), which in turn activates the expression of the lacZ gene in the reporter strain E. coli BHT101 (Fig. 4a). As a representative example of dimeric regulators, the N terminus of PhoP (PhoP-N) was observed to dimerize with itself, as indicated by the level of recorded β-galactosidase activity being comparable to the internal positive control in the form of Zip–Zip homodimerization (Fig. 4b). Next, the potential of WhiB proteins to homodimerize was determined. The WhiB1 and WhiB2 groups showed weak β-galactosidase activities (Fig. 4b), considerably lower than recorded for the PhoP

Fig. 2. Analysis of monomeric and dimeric regulators in M. tuberculosis. β-Galactosidase activities of an E. coli biosensor strain expressing the very many MogR1–220-regulator fusions shown relative to the value for the control strain. Red type indicates WhiB family regulators. Blue type indicates reported dimeric regulators with a repressive effect greater than twofold. Purple type indicates the two-component regulators. See Table S2 for the relative β-galactosidase activities.
and Zip positive controls. All other WhiB proteins showed β-galactosidase activities similar to the negative control group (Fig. 4b). Hence, these data are consistent with our predictions using the MogR1–220 fusion system (Fig. 2), and suggest that the WhiB proteins can be classified as functional monomers.

While the MogR1–220 fusion system cannot exclude the possibility of false classification due to in vivo protein aggregation, heightened protein instability or poor protein synthesis, our analyses still provided pertinent information for subsequent investigations on various specific regulators. Critically, most of our results showing that a large portion of Mycobacteria regulators are dimeric corroborate previous reports that have shown a number of regulators – including KmsR, ArgP, BlaI and EspR as well as the well-characterized two-components regulators – act as dimers in binding to DNA with repeated sequence (Parish, 2014). Moreover, none of monomeric regulators characterized in this study has been previously reported to form dimers. It is evident from a review of the literature that most research has to date focused on the characterization of dimeric transcriptional regulators, meaning that descriptions of the actions of monomeric transcriptional regulators are relatively few. We believe that, by identifying several monomeric regulators, this study can serve as a fillip to initiate investigations of their uncharacterized regulatory mechanisms.

**Interaction of WhiB proteins with the principal sigma factor**

To begin exploring the mechanisms of monomeric regulators, we targeted the WhiB protein family for further analyses. Among these proteins, WhiB3 and WhiB7 were previously reported to interact with the principal sigma factor σ^A in activating gene transcription in *M. tuberculosis* (Steyn et al., 2002; Burian et al., 2013). The sigma factor σ^A binds to RNA polymerase core enzyme as monomer to form the holoenzyme (Murakami et al., 2002). We considered whether all of these WhiB proteins are truly active as monomeric proteins in binding to σ^A. To investigate this, we first tested the interactions between WhiB proteins and σ^A by co-purification as described by others (Burian et al., 2013). For this purpose, we overexpressed histidine-tagged WhiB proteins together with non-tagged σ^A in *E. coli*, and used Ni-NTA to purify proteins from each of the seven WhiB groups as well as a control group only expressing non-tagged σ^A protein. Unexpectedly, the full-length σ^A without histidine-tag binds weakly to Ni-NTA, as evidenced by the detection of some σ^A in the elution fraction even after several washings with 30 mM imidazole buffer in the control group (Fig. 5). WhiB1, WhiB3, WhiB4, WhiB6 and WhiB7 obviously increased the amount of eluted σ^A compared with the group without WhiB protein, suggesting these proteins probably interact with σ^A (Fig. 5). We could not conclude that WhiB2 and WhiB5 interacted with σ^A, for these WhiB proteins contain an iron–sulfur (Fe–S) cluster and are sensitive to oxygen (Smith et al., 2010; Burian et al., 2013), and this makes it difficult to preserve their function during in vitro manipulation under aerobic conditions.
We next applied the BACTH system to confirm the interaction between WhiB proteins and $\sigma^A$. As shown in Fig. 6(a), six of the seven WhiB proteins (except WhiB5) showed an obvious increase in $\beta$-galactosidase activity compared with the negative control expressing the non-fused T25 and T18 fragments, although the $\beta$-galactosidase activities in the WhiB6 and WhiB7 groups were relatively low. These differences may be due to variations in protein–protein affinity among these WhiB proteins or in recombinant protein stability. Corroborating our data, an interaction between WhiB5 and $\sigma^A$ had not been detected (Casonato et al., 2012). Since $M$. tuberculosis contains another principal-like sigma factor, named $\sigma^B$, we also tested the interaction between WhiB5 and $\sigma^B$, but no positive interaction was detected (data not shown). This suggests that WhiB5 may act independently of the sigma factors, which would set it apart mechanistically from the other six WhiB proteins. However, we cannot exclude the possibility that production of T25-WhiB5 is poor because the product is unstable, so that the interaction between WhiB5 and $\sigma^A$ (or $\sigma^B$) is beyond the detection limits of the BACTH system.

As a consequence, we sought to further confirm the interaction between WhiB proteins and $\sigma^A$, by applying the MPFC assay in $M$. smegmatis (Singh et al., 2006). Similar to the BACTH system, the MPFC system is based on the principle that mDHFR complementary fragments $F_{[1,2]}$ and $F_{[3]}$ can be functionally reassembled by two interacting proteins in $M$. smegmatis, which results in bacterial resistance to trimethoprim (Singh et al., 2006). As seen in Fig. 6(b), all $F_{[1,2]}$-fused WhiB proteins except WhiB5 restored the mDHFR activity in combination with $F_{[3]}$-$\sigma^A$ protein in $M$. smegmatis. These observations further demonstrated that WhiB proteins other than WhiB5 interact with $\sigma^A$. Altogether, we conclude that the WhiB proteins act as monomers in transcriptional regulation and most of them probably do so by interacting with the principal $\sigma^A$ protein.

**WhiB proteins have distinct interacting regions with the principal sigma factor**

The arginine 515 on $\sigma^A$ is important for the interaction with WhiB3 (Styn et al., 2002) and WhiB7 (Burian et al., 2013). To explore whether all of the six WhiB family regulators that interact with $\sigma^A$ do so at the same site, we first constructed a mutation in which arginine 515 was replaced by histidine (R515H) within $\sigma^A$. 

---

**Fig. 5.** Co-purification of His$_6$-WhiB proteins with $\sigma^A$ protein. SDS-PAGE of control group ($\sigma^A$ alone) and WhiB groups (His$_6$-WhiB proteins + $\sigma^A$) were stained with Coomassie brilliant blue. Representative fractions of input (I) and eluate (E) from each group are shown.
Then we analysed the interaction between all the WhiB proteins and the mutated $\sigma^A$ variant. Using both BACTH and MPFC systems, we confirmed that mutation of R515H on $\sigma^A$ obviously decreased its interaction with WhiB3 and WhiB7 (Fig. 6) in line with the previous reports (Steyn et al., 2002; Burian et al., 2013). Reduced interaction tendencies with mutated $\sigma^A$ were also observed for the WhiB4 and WhiB6 proteins (Fig. 6). This suggests that WhiB3, WhiB4, WhiB6 and WhiB7 may all bind to $\sigma^A$ using a common interface. However, R515H-mutated $\sigma^A$ still interacted as strongly with WhiB1 and WhiB2 as did the WT sigma factor (Fig. 6). These findings eliminated the possibility that mutation of R515H may influence the stable expression of $\sigma^A$ in the BACTH and MPFC systems and also imply that WhiB1 and WhiB2 target other specific interaction sites on $\sigma^A$.

We next aligned the sequences of all seven WhiB proteins in *M. tuberculosis* using Clustal X and created the phylogenetic tree by ‘Average Distance’ in Jalview software.

As shown in Fig. 7, WhiB5 is clustered into a distinctly separate node, while WhiB1 and WhiB2 are clustered into a node among all the $\sigma^A$-interacting WhiB proteins. These observations are in direct support of our functional data. Since WhiB1 and WhiB2 are essential for *Mycobacteria* (Gomez & Bishai, 2000; Raghunand & Bishai, 2006; Smith et al., 2010), further characterization of novel interacting interfaces of these two proteins with $\sigma^A$ will provide useful information to benefit understanding of the essential regulatory mechanisms involving these regulators.

In summary, we have identified the WhiB family of proteins as major members of a hitherto poorly characterized group of monomeric functioning transcriptional regulators. Moreover, interaction studies revealed that the majority of the WhiB proteins interact with the principal sigma factor, and that this is likely to be a critical mechanistic feature of their ability to control transcription. Based on these data, we hypothesize that the monomeric sigma factor may be a relatively common target for monomeric regulators. In this context, *M. tuberculosis* contains 13 sigma factors (Cole et al., 1998). It would therefore be desirable to examine whether monomeric regulators engage with any of these additional sigma factors as a means to further explore the specific molecular mechanisms used by monomeric proteins in transcriptional regulation.

ACKNOWLEDGEMENTS

We are grateful to Dr Matthew Francis for critically reviewing the manuscript, Dr Adrie J. C. Steyn for providing the plasmids used in the MPFC assays, and Dr Qin Luo for providing the genomic DNA of *L. monocytogenes*. The following reagent was obtained through BEI Resources, NIAID, NIH: genomic DNA from *M. tuberculosis* strain H37Rv, NR-14865. This work was supported by the National Natural Science Foundation of China (numbers 81271797 and 31200111). Y. H. was supported by the Youth Innovation Promotion Association CAS (2013220).
REFERENCES


Edited by: R. Manganelli