The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase

Aline Tabib-Salazar¹, Bing Liu², Philip Doughty¹, Richard A. Lewis¹, Somadri Ghosh¹, Marie-Laure Parsy², Peter J. Simpson², Kathleen O’Dwyer², Steve J. Matthews² and Mark S. Paget¹,*

¹School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK and ²Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

Received February 19, 2013; Revised March 26, 2013; Accepted March 28, 2013

ABSTRACT

RbpA is a small non–DNA-binding transcription factor that associates with RNA polymerase holoenzyme and stimulates transcription in actinobacteria, including Streptomyces coelicolor and Mycobacterium tuberculosis. RbpA seems to show specificity for the vegetative form of RNA polymerase as opposed to alternative forms of the enzyme. Here, we explain the basis of this specificity by showing that RbpA binds directly to the principal σ subunit in these organisms, but not to more diverged alternative σ factors. Nuclear magnetic resonance spectroscopy revealed that, although differing in their requirement for structural zinc, the RbpA orthologues from S. coelicolor and M. tuberculosis share a common structural core domain, with extensive, apparently disordered, N- and C-terminal regions. The RbpA–σ interaction is mediated by the C-terminal region of RbpA and σ domain 2, and S. coelicolor RbpA mutants that are defective in binding σ are unable to stimulate transcription in vitro and are inactive in vivo. Given that RbpA is essential in M. tuberculosis and critical for growth in S. coelicolor, these data support a model in which RbpA plays a key role in the σ cycle in actinobacteria.

INTRODUCTION

The high G+C, Gram-positive actinobacteria are of immense medical and industrial importance. For example, although the Streptomyces genus is a key source of bioactive compounds, including antibiotics, immunosuppressants and anti-cancer drugs, Mycobacterium tuberculosis has maintained its global notoriety, currently causing 9 million cases and 1.5 million tuberculosis deaths per year (1). As well as being a major target for anti-mycobacterial agents, such as rifampicin, RNA polymerase (RNAP) also has a major influence on the control of antibiotic biosynthesis in Streptomyces, and mutations in this enzyme complex can both positively and negatively impact production (2,3). Therefore, a deeper understanding of RNAP and its regulation in the Actinobacteria phylum can contribute to both the production and development of new medicines.

Transcription initiation in bacteria occurs in several steps. First, an RNAP holoenzyme binds to a promoter element to form an initial closed complex in which the DNA remains double-stranded. This triggers a process of isomerization in which the DNA unwinds, and localized DNA melting reveals the template strand to the active site of the enzyme, giving rise to an open complex. Subsequently, short abortive transcripts are produced in a cyclical manner before the RNAP escapes into the elongation phase (4). The RNAP holoenzyme can be divided into core (five subunits, 2σ, 2β, β′ and ω), which is catalytically active for elongation, and a dissociable σ factor subunit that is required for promoter recognition. The σ subunit provides critical DNA-binding determinants in both closed and open complexes, but it stochastically dissociates from elongating RNAP soon after promoter escape (5,6). This scheme gives rise to a ‘σ cycle’ whereby the dissociated σ factor enters a pool of σ factors that compete for binding to core RNAP (7). Bacteria are, therefore, able to tune gene expression by controlling the cellular level and composition of σ factors (8). All bacteria seem to contain a single essential σ factor that is orthologous to σ⁰ of Escherichia coli and directs the transcription of most ‘housekeeping’ genes during exponential growth, such as those involved in ribosome production and central metabolism. Structural studies on the principal σ factors of E. coli and Thermus...
sp., either alone or in the context of holoenzyme, revealed four domains, σ1,1, σ2, σ3 and σ4, that are interconnected by flexible linkers and correspond to conserved regions 1.1, 1.2–2.4, 3.0–3.1 and 4.1–4.2, respectively (9–12). The σ2, σ3 and σ4 domains each comprise key contact points for both core enzyme and promoter DNA, whereas σ1,1 is thought to maintain free σ in a compact non–DNA-binding form through interaction with σ2 and σ3 (12).

In addition to a principal and essential σ factor, bacteria usually contain one or more alternative σ factors that can redirect the transcription machinery to specific regulons, allowing the cell to respond to a wide variety of environmental, physiological or developmental cues (8). Most of these are related to σ70 and have been classified into four main groups (Groups 1–4) based on phylogeny and biological function (13): Group 1 σ factors are the σ70 orthologues; Group 2 σ factors are closely related to Group 1 σ factors but have specialized non-essential functions; Group 3 σ factors are structurally and functionally diverse but include all three globular domains (σ2-σ3); Group 4 comprises the highly diverse extracytoplasmic function (ECF) σ factors that lack both Region 1 and the σ3 globular domain. The control of alternative σ factor activity is particularly well established and often occurs at the post-translational level through the action of anti-σ factors that bind to σ and prevent its association with core RNAP (14). The σ factor is released after the inactivation of the anti-σ factor through competition with an alternative binding partner (partner-switching), regulated proteolysis or an allosteric change in response to the direct sensing of a signal. Although the σ subunit can act as an essential contact point for DNA-binding activators, proteins that bind to σ and activate free σ factors are rare. One example, however, is the small protein Crl, which stimulates the activity of the stationary phase regulator σS in E. coli and Salmonella (15) through direct interaction, both with free σ and as part of the holoenzyme (16,17). The role of Crl seems to be to improve the association of σS with core RNAP through interaction with the σ2 globular domain (18–21).

RbpA is an RNAP-binding protein that is confined to, and widespread within, the actinobacteria, and evidence is accumulating that it plays a crucial role in transcription initiation. RbpA was discovered in the antibiotic-producing bacterium Streptomyces coelicolor where the 124 amino acid protein was shown to confer basal levels of resistance to rifampicin in vivo and stimulate transcription in vitro (22,23). The S. coelicolor RbpA (designated RbpAsc for clarity) co-purifies as a major component of RNAP holoenzyme, suggesting a broad role in transcription, although stimulation of transcription was previously only demonstrated using the ribosomal RNA promoter rrrDp3 (22). The rrrDp3 promoter is dependent on the principal σ factor in S. coelicolor, σHrdB, and RbpAsc failed to stimulate transcription from a promoter dependent on the alternative σ factor σR (22). Although S. coelicolor mutants that lack RbpAsc grow slowly, they are viable. However, it was recently demonstrated that RbpA (Rv2050; designated RbpAMt for clarity) is essential in M. tuberculosis (24). Furthermore, RbpAMt was shown to stabilize the principal σA-RNAP holoenzyme and stimulate transcription from wide range of σA-dependent promoters in vitro, suggesting a general role in transcription initiation (25). Two biochemical studies have proposed that the primary contact point for RbpAMt is the β subunit. However, although one study proposed binding close to the active-site cleft near the rifampicin-binding site (26), the other study proposed that RbpAMt contacts the sandwich-barrel hybrid motif located on the surface of the β subunit (25). In either study, the apparent specificity of RbpAMt for the principal σ factor was not explained, although it was proposed that RbpAMt might allosterically modify core RNAP to improve its interaction with σ (25). Here, we show that RbpA binds specifically to the principal and certain related σ factors of S. coelicolor and M. tuberculosis, and that RbpAsc is present at a σHrdB-RNAP transcription initiation complex in vivo. The orthologous RbpA proteins from these organisms comprise a core domain, for which we present the high-resolution solution structures, together with extensive N- and C-terminal regions that are disordered in the free proteins. Finally, we map the interaction between S. coelicolor σHrdB and RbpAsc, and we show that binding occurs primarily between the σ2 domain of σHrdB and the C-terminal region of RbpAsc. Our data support a model in which RbpA plays a key role in the σ cycle of the vegetative RNAP in actinobacteria, possibly acting as a chaperone-like protein to aid the formation of active holoenzyme, which might be particularly important during stress.

**MATERIALS AND METHODS**

**Strains, plasmids and growth conditions**

Strains and plasmids are listed in Supplementary Table S1, and oligonucleotides used for the construction of plasmids are listed in Supplementary Table S2. E. coli K-12 strain DH5α was used for general cloning and plasmid propagation, and E. coli ET12567 (pUZ8002) was used to conjugate plasmids into S. coelicolor (27). S. coelicolor A3(2) strains were cultivated on mannitol–soya (MS) agar or in a modified yeast extract–malt extract (YEME, 10% sucrose and 0.01% antifoam) broth as previously described (27). To construct S. coelicolor S129, a ΔrbpA:apr allele was introduced into the S. coelicolor J1981 (rpoC-His6) genome by double crossover recombination as described previously (22). To construct pSX190, the S. coelicolor rbpA gene and promoter region were amplified using oligonucleotide primers rbpA_H3_rev and DT1a (22), thereby introducing a HindIII site immediately upstream of the stop codon, then fused to a triple FLAG tag sequence in the integrative vector pSET152. Site-directed mutagenesis of the C-terminal region of rbpA was performed in pSX512, and the rbpA derivatives were cloned along with the native promoter into pSETΩ as a BglIII fragment.

**Protein overexpression and purification**

Core RNAP was prepared from late-exponential YEME cultures of S. coelicolor S129. Harvested
mycelium was re-suspended in binding buffer (20 mM Tris–HCl, pH 7.5, 500 mM NaCl, 20 mM imidazole, 3 mM 2-mercaptoethanol, 5% glycerol and 25 μg/ml phenylmethylsulphonyl fluoride) plus complete ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor (Roche). Mycelium was lysed by passing twice through a French press at 10 000 kPa before centrifugation to remove the cell debris. The lysate was immediately loaded on a Ni-IDA sepharose column and washed with 10 column volumes of binding buffer followed by 5 column volumes buffer A (20 mM Tris–HCl, pH 7.5, 3 mM 2-mercaptoethanol, 5% glycerol and 150 mM NaCl) containing 20 mM imidazole, then eluted with buffer A containing 250 mM imidazole. The eluate was immediately applied to a HiTrap Heparin HP column (GE Healthcare) equilibrated with buffer A, and protein was dialysed against buffer A containing 250 mM imidazole, then eluted with 20 mM Tris–HCl, pH 7.5, 500 mM NaCl, 20 mM imidazole, and subsequently cloned into pET15b to give pSX505. For NMR analysis, the full-length HrdB domain fragments were removed by treatment with thrombin, and the protein was purified by gel filtration. For expression of truncated HrdB domain fragments (σ2, σ2-σ4, σ3-σ4 and σ4) with N-terminal His6-tags, polymerase chain reaction (PCR) was used to amplify the corresponding hrdB regions with primers incorporating N-terminal NdeI and C-terminal XbaI sites and cloned into pET15b. Expression was achieved in E. coli BL21λDE3 (pLysS), and fragments were purified using Ni-NTA spin columns (Qiagen) as recommended by the manufacturer. For NMR analysis, the full-length M. tuberculosis rbpA gene (Rv2050) was amplified by PCR incorporating an N-terminal NdeI site and C-terminal BamHI site and cloned into pET20b to give pSX500. To solve the structure of the core domain of RbpA Sc, a truncated gene encoding the first 75 amino acids was chemically synthesized with codon optimization (Eurofins) for expression in E. coli and subsequently cloned into pET15b to give pSX505. E. coli BL21λDE3 (pLysS) containing pSX505 or pSX500 was grown in L-broth to OD600 of 3–5, then transferred to optimized high-cell density isopropyl-β-D-thiogalactopyranoside (IPTG)-induction minimal media (28,29) containing 0.1% 15NH4Cl and 1% 13C-glucose, grown for 2 h at 37°C, before induction with 1 mM IPTG at 30°C for 3 h. Oligonucleotides used to amplify each gene or domain are listed in Supplementary Table S2.

**In vitro transcription reactions**

*In vitro* transcription assays were carried out using *S. coelicolor* S129 core RNAP and purified native HrdB and RbpA Sc proteins, essentially as described previously (30). Templates containing promoters were amplified by PCR using primers listed in Supplementary Table S2. Assays (10 μl) were performed at 30°C, and they contained 5 nM DNA template in reaction buffer (50 mM Tris–HCl, pH 8, 10 mM MgCl2, 200 mM KCl, 0.1 mg/ml of bovine serum albumin and 10 mM β-mecaptoethanol), σHrdB, RbpA Sc and core RNAP were mixed in RNAP dilution buffer (10 mM Tris–HCl, pH 8, 10 mM KCl, 0.4 mg/ml of bovine serum albumin, 0.1 mM Na2EDTA, 0.1% Triton X-100 and 10 mM β-mecaptoethanol) (30) for 15 min on ice, and reactions were initiated by the addition of nucleotide mix (1 mM adenosine triphosphate, GTP, CTP and 50 μM UTP), including 5 μCi [γ-32P] UTP (800 Ci/mmol, GE Healthcare), and incubated for 15 min. Reactions were halted by addition of an equal volume of *in vitro* loading dye [80% (w/v) formamide, 0.01% (w/v) xylene cyanol and 0.01% (w/v) bromophenol blue], separated on denaturing 6% polyacrylamide gels and quantified by phosphorimaging.

**Bacterial two-hybrid analysis**
The complete rbpA genes from *S. coelicolor* and *M. tuberculosis*, as well as truncated derivatives as indicated in figure legends, were PCR amplified then cloned into pUT18, generating rbpA-T18 translational fusions to the T18 catalytic domain of *Bordetella pertussis* adenylate cyclase. The σ factor genes from *S. coelicolor* (hrdA, hrdB, hrdC, hrdD, sigB, sigE, sigR and whiG) and *M. tuberculosis* (sigA and sigB), as well as truncated derivatives of hrdB, were cloned into pKT25, generating T25-σ translational fusions to the T25 catalytic domain. The amino acids included in each truncated hybrid construct are indicated in the figure legends, and the oligonucleotides used for PCR amplification are indicated in Supplementary Table S2. Functional complementation between chimeric proteins was evaluated in *E. coli* BTH101 by performing β-galactosidase measurements on permeabilized overnight cultures grown in L-broth containing 0.5 mM IPTG. One unit of β-galactosidase activity corresponds to 1 nmol of o-nitrophenyl-β-D-galactoside (ONPG) hydrolysed per min at 28°C. Results were presented either as Miller units (31) or as % Miller units relative to a control.

**In vitro pull-down assays**
Experiments were performed using Dynabeads® His-Tag Isolation and Pulldown beads (Invitrogen) essentially as recommended by the manufacturer. Test proteins were mixed on ice in binding buffer (50 mM sodium phosphate, pH 8, 300 mM NaCl and 0.01% Tween20), then added to 20 μl magnetic Dynabeads and left to incubate for 30 min at ambient temperature with shaking. Samples were washed four times, and then eluted in 100 μl binding buffer (binding buffer plus 300 mM imidazole) for 15 min.

**Chromatin immunoprecipitation–quantitativePCR**
Strains were grown in modified YEME to mid-late exponential phase (OD450 2.0), treated with rifampicin (140 μg/ml) for 20 min, then formaldehyde [1.1% (v/v); Sigma] for a further 20 min, with continued shaking at 200 rpm. After quenching with 0.5 M glycine for 5 min at 30°C, mycelium was collected by centrifugation, washed
twice in phosphate-buffered saline and stored at −20°C. Chromatin extraction and immunoprecipitation protocols were essentially as described previously (32). Mycelial pellets were subjected to 3 × 90 s of cryogenic grinding (Retsch MM301, 30 Hz) then resuspended in 9 ml of IP buffer (32), and 3-ml aliquots were sonicated (35% amplitude) for 4, 8 and 12 cycles of 15 s ON and 45 s OFF. After clearing by centrifugation, 70 µl of samples were incubated with proteinase K (Roche), de-cross-linked and then analysed by gel electrophoresis (32). Samples comprising chromatin fragments of ~0.5 kb were pre-cleared by incubation (rotation for 4 h at 4°C) with either protein A or protein G magnetic beads (NEB) before the addition of antibodies. Immunoprecipitations were conducted with M2 mouse monoclonal anti-FLAG antibody (Sigma), monoclonal 8RB13 anti-E. coli RNAP β subunit antibody (Abcam) or a custom rabbit polyclonal antiserum raised against σ^{HrdB} (Eurogentec). Corresponding mock immunoprecipitations (IPs) with no added antibody were performed as negative control. After incubation with antibodies (rotation for 12 h at 4°C), immunoprecipitations were performed using protein G magnetic beads (RbpASc and RNAP; NEB) or protein A magnetic beads (σ^{HrdB}, NEB). Subsequent steps to dissociate cross-links and purify DNA were as described previously (32), and DNA was resuspended in 25 µl of ultrapure water. Quantitative real-time PCR was performed using the standard curve absolute quantification approach using serial dilutions of sonicated S. coelicolor M145 genomic DNA to generate standard curves for each of the four primer pairs used (Supplementary Table S2). qPCR reactions were set-up in 0.1 ml of MicroAmp FAST optical 96-well reaction plates (Applied Biosystems) using GoTaq qPCR reaction mix (Promega), 10 ng of template DNA and each primer at 0.15 pmol/µl. The data were analysed using a modified method to that previously described (33). The mock immunoprecipitation (IP) signal was subtracted from the signal obtained using the corresponding immunoprecipitated sample DNA as input to correct for background signal. To allow visual comparison of the relative enrichments obtained using the three different antibodies, the background corrected signals obtained using a particular antibody were expressed relative to the highest corrected signal obtained using either of the four primer pairs.

**NMR and solution structure determination**

Samples were buffer-exchanged into 100 mM MES buffer, pH 6.0, with 150 mM KCl. NMR experiments were performed on >200 µM samples of uniformly ^{15}N- and ^{13}C-labelled protein. The backbone resonance assignments of RbpASc and RbpAMt were completed by using the standard triple-resonance assignment approach (34), and data were analysed using in-house algorithms with the program NMRView (35). Hz and Hβ assignments were obtained using HBHA(CBCACO)NH. The side-chain assignments of the two proteins were extended using HCCH TOCSY and (H)CC(CO)NH TOCSY. The distance restraints were obtained from ^{1}H-^{15}N/^{13}C nuclear Overhauser effect spectroscopy (NOESY)–heteronuclear single- or multiple-quantum coherence (HSQC/HMQC) experiments (mixing time, 100 ms at 800 MHz). Data were collected on Bruker AvanceIII (600 MHz) and AvanceII (800 MHz) spectrometers equipped with TCI and TXI cryoprobes, respectively. The ARIA protocol was used for completion of the NOE assignment and structure calculation (36). Dihedral angle restraints derived from TALOS were also incorporated in the calculation (37). The frequency window tolerances for assigning NOEs were ±0.05 ppm for direct proton dimensions and ±0.05 ppm for indirect proton dimensions, and ±0.5 ppm for nitrogen dimensions and ±1.3 ppm for carbon dimensions. One hundred structures were calculated in the final iteration, and these had no NOE violations >0.5 Å and no dihedral angle violations >5°. The 10 lowest energy structures were selected for further analysis and figure generation.

**RESULTS**

RbpASc activates a range of σ^{HrdB}-dependent promoters

Previously, using total RNAP holoenzyme preparations, we showed that RbpASc stimulated transcription in vitro from the σ^{HrdB}-dependent promoter rrnDp3, but not from a σ^{R}-dependent promoter, rbpAp (22). Initially, we extended this work by showing that RbpA increased the maximal activity of σ^{HrdB}-RNAP at the rrnDp3 promoter by titrating increasing concentrations of σ^{HrdB} in the presence and absence of excess RbpASc (Figure 1A). We then tested whether the RbpASc activity was specific to rrnDp3 or whether it played a more general role in activating σ^{HrdB}-dependent promoters. Transcription run-off assays were performed with a selection of promoters that we or others have predicted to be dependent on σ^{HrdB}, based on the putative −10 promoter elements: sacAp (38), tuf3p (39), relAp2 (40), atiplp (unpublished observations) and rpllp (41). In each case, RbpASc stimulated transcription in assays consisting of core RNAP, σ^{HrdB} and linear DNA templates (Figure 1B). RbpASc also activated transcription initiation in single-round transcription assays, which suggests that the stimulatory effect is not a consequence of increased recycling or elongation (data not shown). Together with the previous demonstration that RbpAMt stimulates a range of σ^{A}-dependent promoters with no apparent specificity (25),

**Metal analysis**

Native (un-tagged) RbpA proteins were produced in E. coli and purified by ion-exchange chromatography and gel filtration, before concentration as described previously (22). Approximately 0.5 mg of protein was vacuum-dried then re-suspended in 86 µl of 70% nitric acid (metal analysis grade) and heated for 10 min at 105°C. Water (metal analysis grade) was added to give a final concentration of 2% nitric acid, and inductively coupled plasma mass spectrometry was performed using an Agilent 7500 series instrument. A solution containing 1 ppm of each metal to be analysed (Mn, Fe, Ni, Cu and Zn) was used as standard, and values were normalized against protein storage buffer alone. Experiments were performed in duplicate.
core RNAP (75 nM) reactions using the dependent promoters in vitro or 2, or 3, in the presence (grey bars) or absence (black bars) of (ChIP–qPCR). The presence of RbpASc, immunoprecipitation combined with quantitative PCR therefore, decided to test this using chromatin

tation reactions contained core RNAP (75 nM) RNAP have residual levels of relative to the highest signal. Note that these preparations of core

S129, indicating that the epitope tag did not impede

struct, pSX190, fully restored normal growth rate to

ment site using the vector pSET152. The resulting con-

tor of the principal RNAP holoenzyme in actinobacteria.

Although it is known that RbpASc is present in purified RNAP preparations, it has not yet been shown to be present in RNAP–promoter complexes in vivo. We, therefore, decided to test this using chromatin immunoprecipitation combined with quantitative PCR (ChIP–qPCR). The presence of RbpASc, σ^HrdB and the β subunit of RNAP was analysed at the rplJ promoter. Immunoprecipitation of β and σ^HrdB was performed using a commercially available monoclonal antibody (42) and a polyclonal antibody, respectively. To immunoprecipitate RbpASc, the rbpA gene was C-terminally tagged with tandem FLAG epitopes and integrated into the genome of the S. coelicolor rbpA mutant S129 at the φC31 attachment site using the vector pSET152. The resulting construct, pSX190, fully restored normal growth rate to S129, indicating that the epitope tag did not impede

normal RbpASc function (data not shown). S. coelicolor S129 (pSX190) was grown to mid-exponential phase before treatment with rifampicin to inhibit global RNA synthesis by trapping RNAP at promoters (43). As expected, the rplJ promoter region (qPCR product 2; Figure 2), centred −79 bp upstream from the transcription start point, was highly enriched for both σ^HrdB and the β subunit, with much lower enrichment seen for control regions centred −653 bp upstream, or +235 bp and +608 bp downstream, of the transcription initiation site, suggesting that σ^HrdB–RNAP was indeed trapped at the rplJ promoter (Figure 2). The enrichment patterns seen using the anti-FLAG antibody were similar to those obtained with the σ^HrdB and the β antibodies, indicating that RbpASc is present in these transcription initiation complexes, although it should be noted that differences in antibody–epitope affinities prevent the relative proportion of initiation complexes that contain RbpASc from being assessed. Control anti-FLAG immunoprecipitation experiments using an equivalent strain in which RbpASc was untagged showed no enrichment above that seen with the ‘no antibody’ control, confirming that this signal is specific to RbpASc (data not shown).

**RbpA binds specifically to principal and related σ factors**

We considered the possibility that the apparent specificity of RbpA activity towards promoters recognized by the principal σ factors might be determined by a direct inter-action with the σ factor itself. To test this idea, we applied
a bacterial two-hybrid (BACTH) assay based on the
B. pertussis adenylate cyclase (44). The S. coelicolor rbpA
gene was fused to the gene that encodes the T18 domain of
adenylate cyclase such that RbpASc was at the N-terminus
of the fusion protein. A range of σ factor genes were fused
to the gene that encodes the T25 domain, such that the σ
factors were located at the C-terminus of the fusion
proteins. Initially we tested for interaction between
RbpASc and the following S. coelicolor σ factors: σHrdB
(Group 1); σHrdA, σHrdC and σHrdD (Group 2); σWhiG
and σB (Group 3); and σE and σR (Group 4) (Figure 3A).
The Group 1 and Group 2 σ factors were N-terminally truncated such that they included only
conserved regions 1.2–4.2, which comprise structural
domains σ2, σ3 and σ4. This was necessary because full-
length T25–hrdB fusions seemed to be toxic to E. coli.
A positive result was detected for σHrdB and σHrdA,
whereas none of the remaining σ factors bound to
RbpASc. This suggests that RbpASc interacts only with
principal and closely related σ factors. To broaden this
idea, we tested for interaction between RbpAMt and
M. tuberculosis σ factors σA (Group 1) and σB (Group
2). Binding was detected in both cases (Figure 3A),
which suggests that the interaction with principal and related σ
factors is a general feature of RbpA proteins in
actinobacteria. In principal, the interactions observed
during in vivo two-hybrid analysis could involve E. coli
RNAP in hybrid higher order complexes. To confirm
that the interaction between RbpASc and σHrdB does not
involve any other factors, the two proteins were analysed,
separately or mixed, by gel filtration chromatography
(Figure 3B). σHrdB (55.9 kDa) and RbpASc (14.1 kDa)
ebulated at different volumes with apparent molecular
masses of 183.2 and 29.9 kDa, respectively, suggesting
trimeric and dimeric complexes. When mixed together,
a complex of the two proteins eluted close to the σHrdB
peak with an apparent mass of 211.4 kDa. Although it is clear
that the two proteins interact, the difference in size between
σHrdB and RbpASc, and our suspicion that σHrdB might run
anomalously during gel filtration, makes it difficult to
predict precise stoichiometry at this point.

RbpA orthologues differ in their zinc content

Protein sequence alignments suggested that the RbpA
orthologues from S. coelicolor and M. tuberculosis might
differ in their metal content (Figure 4). S. coelicolor
RbpASc contains three cysteine residues, Cys35, Cys56
and Cys59, and a histidine His38, each of which are
widely conserved (e.g. in Bifidobacterium sp. and Frankia
sp.) and arranged as a putative C(H/C)CC zinc-ribbon
motif. Surprisingly, RbpAMt, along with RbpA orthologues
from Corynebacterium sp. and Nocardioid sp., lacks
all of these residues apart from the Cys56-equivalent
cysteine. To investigate metal content, we isolated
RbpASc and RbpAMt from overproducing E. coli strains
in their native (untagged) forms. Inductively coupled
plasma mass spectrometry (ICP-MS), set to detect Mn,
Fe, Ni, Cu and Zn, revealed zinc as the only metal
present at significant quantities in RbpASc, at a stoichiometry
of 0.83 (±0.04) mol zinc mol⁻¹ protein. Conversely,
no metal was detected at significant levels in RbpAMt.
To test the importance of the presumed cysteine zinc
ligands in RbpASc, each was mutated to alanine, and the
resultant genes were introduced into S. coelicolor S101
(ΔrbpA::hyg) using the integrative expression vector
pJ6902 to test for complementation. In each case, the
mutant genes failed to restore normal growth rate
(Supplementary Figure S1), suggesting that each residue
is crucial for the function or stability of the protein. Each

![Figure 3. RbpA binds to the principal and closely related σ factors.](http://nar.oxfordjournals.org/ at University of Sussex on June 18, 2014)
of the three histidine residues in RbpASc was also changed to alanine, although in each case the mutant protein complemented the S101 strain (Supplementary Figure S1), suggesting that the protein retains function even when the functional requirement of zinc is conferred solely by the cysteine ligands.

The solution structure of the core domain of RbpA from *M. tuberculosis* and *S. coelicolor*

To investigate why RbpASc and RbpAMt differ in their zinc content and to dissect the structure–function relationship of RbpA, we used NMR spectroscopy to determine the solution structure of each protein. The dispersion of resonances in proton NMR spectra recorded on both RbpA proteins clearly showed that the proteins were at least partially structured (data not shown), although the *M. tuberculosis* protein gave better spectra and was, therefore, our initial focus. To determine the structured portions of the protein, attempts were made to assign backbone amides. Despite efforts to prevent proteolysis, the samples degraded rapidly, and complete assignments were obtained only for the core domain. Resonance line widths and chemical shift dispersion indicate that regions outside these domains are highly disordered. The solution structure of the core domain was then determined using standard multidimensional NMR spectroscopy methods. After this, we determined the structure of the RbpASc core domain using a truncated clone that included RbpASc residues 1–75. The structures of the RbpA orthologues (Figure 5; see Supplementary Table S3 for NMR and refinement statistics) reveal a compact core domain (residues 1–75). The structures of the RbpA orthologues (Figure 5; see Supplementary Table S3 for NMR and refinement statistics) reveal a compact core domain (residues 1–75). The structures of the RbpA orthologues (Figure 5; see Supplementary Table S3 for NMR and refinement statistics) reveal a compact core domain (residues 1–75). The structures of the RbpA orthologues (Figure 5; see Supplementary Table S3 for NMR and refinement statistics) reveal a compact core domain (residues 1–75).

The C-terminal region of RbpASc is required for interaction with $\sigma^{HrdB}_{E. coli}$

Although the NMR structures of the RbpA core domain did not include the C-terminal region, and indeed suggested that this region was disordered, sequence-based structural predictions using the Phyre 2 package (50) suggested the presence of two $\alpha$-helices (designated $\alpha_1$ and $\alpha_2$), the second of which would be particularly amphipathic with a negatively charged face (Figure 4). To identify the region of RbpASc that interacts with $\sigma^{HrdB}_{E. coli}$, we used BACTH to test three truncated *S. coelicolor* rbpA-T18 fusions for interaction with
T25-σ^HrdB (Figure 6): RbpASc lacking both predicted C-terminal helices (residues 1–72, RbpA1-72); RbpASc lacking only α2 (RbpA1-90); and RbpASc lacking the N-terminal region and RCD domain but including α1 and α2 (RbpA73–124). In contrast to RbpA1-72 and RbpA1-90 for which an interaction with σ^HrdB was not detected, RbpA73–124 did interact, indicating that the C-terminal helical region of RbpASc is necessary and sufficient for binding. To identify critical residues involved in this interaction, we performed alanine-scanning mutagenesis, focusing on highly conserved residues in the C-terminal region of RbpASc. The RbpA^Sc fragments and σ^HrdB were fused to the T18 and T25 domains of B. pertussis adenylate cyclase, respectively. (A) A schematic diagram indicating the regions of RbpASc tested for interaction with σ^HrdB using BACTH analysis. The RbpA^Sc fragments and σ^HrdB were fused to the T18 and T25 domains of B. pertussis adenylate cyclase, respectively. (B) BACTH analysis of alanine point mutations in conserved residues in the C-terminal region of RbpASc. C, control strain containing pKT25–hrdB and pUT18. For (A) and (B), experiments were performed in triplicate (standard deviations indicated), and data are presented as % Miller units relative to results obtained with the full-length wild-type RbpASc. Control strains with pKT25–hrdB and pUT18 exhibited <1% the activity of the full-length RbpASc interaction. (C) Growth of S129 (ΔrbpA), S129 containing pSETV::rbpA (WT), or equivalent constructs with the rbpA mutations as indicated. Strains were streaked to MS agar and photographed after 4 days incubation at 30°C. (D) In vitro transcription from the rplJ promoter in the presence of RbpASc or RbpASc R89A/R90A mutant proteins. Multi-round in vitro transcription reactions contained core RNAP (50 nM) σ^HrdB (250 nM), RbpASc (250 nM or 1.25 μM) and a DNA template generated by PCR using primers listed in Supplementary Table S2. Transcript levels were quantified by phosphorimaging, and data are presented relative to reactions performed in the presence of σ^HrdB but the absence of RbpA.
mutant was completely defective in $\sigma^{HrdB}$ binding. To test the importance of this interaction in the function of RbpA$^{Sc}$, we cloned the mutant genes into the integrative plasmid pSET$\Omega$ and tested for complementation of S. coelicolor S129 ($\Delta$rbpA:acr). Although the introduction of wild-type rbpA and the E88A mutant (data not shown) completely restored normal growth to S129, the rbpA R89A/R90A mutant did not. The individual R89A and R90A mutants also failed to completely restore normal growth, each exhibiting a slight, but reproducible, reduction in colony size and a delay in sporulation timing. As the inability of the R89A/R90A double mutant to complement S. coelicolor S129 might be due to protein instability in vivo, we overexpressed the protein in E. coli and tested whether the pure protein would stimulate transcription in vitro. Using the rplJp promoter as template, the RbpA$^{Sc}$ R89A/R90A mutant failed to stimulate transcription, which suggests that a direct interaction with $\sigma^{HrdB}$ is required for the action of RbpA$^{Sc}$.

RbpA binds to the $\sigma_2$ domain of the principal sigma factor

The different domains of $\sigma$ have distinct functions; therefore, to better understand the role of RbpA in transcription initiation, we sought to localize the interaction to individual domain(s) of $\sigma^{HrdB}$. The $\sigma_2$, $\sigma_3$, and $\sigma_4$ structural domains of $\sigma^{HrdB}$ were predicted based on the structures of $\sigma^A$ from Thermus sp. (11,51), and five overlapping T25–hrdB fusions were constructed and tested for interaction with the rbpA–T18 fusion in BACTH analysis (Figure 7). No interaction was detected for T25–hrdB ($\sigma_3$–$\sigma_4$) or T25–hrdB ($\sigma_4$). However, all fragments that included $\sigma_2$ interacted, including T25–hrdB ($\sigma_2$) (amino acid residues 211–347), which comprises conserved regions 1.2–2.4. To confirm these data, each of the hrdB constructs used in the BACTH analysis was overexpressed with an N-terminal His$_6$-tag, and the corresponding proteins purified, then tested for interaction with RbpA$^{Sc}$ using an in vitro pull-down assay (apart from $\sigma_1$–$\sigma_2$, which was insoluble). The pure $\sigma^{HrdB}$ fragments (domains $\sigma_2$, $\sigma_2$–$\sigma_4$, $\sigma_3$–$\sigma_4$ and $\sigma_4$) were mixed with native RbpA$^{Sc}$ for 15 min and then isolated using magnetic Ni-affinity beads. RbpA$^{Sc}$ was co-isolated with the $\sigma_2$ and $\sigma_2$–$\sigma_4$ fragments of $\sigma^{HrdB}$ but not with $\sigma_3$–$\sigma_4$ and $\sigma_4$, confirming a direct and specific interaction with the $\sigma_2$ domain. Using both BACTH and in vitro pull downs, we also detected equivalent interactions between the $\sigma_2$ domain of $\sigma^A$ and the RbpA$^{Mt}$ (data not shown). Furthermore, when equivalent $\sigma_2$ and RbpA proteins from S. coelicolor and M. tuberculosis are co-expressed in E. coli, with one partner His$_6$-tagged, they co-elute as a complex during Ni-affinity chromatography (data not shown). Taken together, these data suggest that the $\sigma_2$–RbpA interaction involving the principal $\sigma$ factor is conserved across the actinobacteria.

DISCUSSION

Here, we provide an explanation for why the RNA polymerase-binding protein RbpA stimulates transcription from promoters that are dependent on the principal $\sigma$ factor in S. coelicolor and M. tuberculosis ($\sigma^{HrdB}$ and $\sigma^A$), but not from promoters dependent on alternative $\sigma$ factors. Although RbpA$^{Sc}$ was originally discovered as a component of RNAP holoenzyme, and it is demonstrated here to be present at transcription initiation complexes in S. coelicolor, we found that RbpA$^{Sc}$ can also form a complex with the free principal $\sigma$ factor. We failed to detect RbpA$^{Sc}$ interactions with selected Group 3 ($\sigma^B$ and $\sigma^{WhgC}$) and Group 4 ($\sigma^E$ and $\sigma^F$) $\sigma$ factors in S. coelicolor, although we found that RbpA$^{Sc}$ and RbpA$^{Mt}$ can bind to the Group 2 $\sigma$ factors, $\sigma^{HrdA}$ in S. coelicolor and $\sigma^B$ in M. tuberculosis, respectively. In S. coelicolor, the three Group 2 $\sigma$ factors ($\sigma^{HrdA}$, $\sigma^{HrdC}$ and $\sigma^{HrdD}$) are collectively non-essential (52), as is $\sigma^B$ in M. tuberculosis (53), which implies that the interaction with, and activation of, the essential primary $\sigma$ factor underlies the interaction.
critical importance of RbpA for growth. Consistent with this, an RbpA<sup>Sc</sup> mutant that failed to bind σ<sup>HbdR</sup> (RbpA<sup>HbdR/ScSc</sup>) did not stimulate transcription <i>in vitro</i> and could not complement an <i>S. coelicolor</i> rbpA null mutant. Given that σ<sup>HbdR</sup>-RbpA<sup>Sc</sup> complexes are stably maintained during gel filtration and Ni-affinity chromatography, we speculate that the two proteins are able to form an initial binary complex before the formation of holoenzyme. One role of RbpA might, therefore, be to act in a chaperone-like fashion to remodel σ for optimal binding to core RNAP, and thereby be an integral component of the ‘σ cycle’ in actinobacteria. This is consistent with the observation that RbpA<sup>Mt</sup> stabilizes <i>M. tuberculosis</i> σ<sup>A</sup>-RNAP holoenzyme, but it argues for a more direct role rather than a proposed allosteric effect through changing the conformation of the β lobes (25). Such a role might account for the induction of <i>rbpA</i> during stress, as also proposed by others (25). Although <i>rbpA</i> is continually expressed in unstarved <i>S. coelicolor</i> cultures, consistent with its general importance for growth (23), transcription is also induced in response to disulphide stress, directed by the alternative ECF σ factor σ<sup>R</sup> (23). In <i>M. tuberculosis</i>, the <i>rbpA</i> gene is also induced by stress through the σ<sup>H</sup> orthologue, σ<sup>H</sup> (54).

The induction of <i>rbpA</i> might, therefore, ensure that the increased cellular levels of free alternative σ factors during stress do not adversely affect σ<sup>HbdR</sup>/σ<sup>A</sup>-dependent ‘housekeeping’ gene expression, and it may also help the cell to reset the normal state once the stress has been resolved. Interestingly, it was recently discovered that the gene that encodes σ<sup>HbdR</sup> is also controlled by two ECF σ factors: σ<sup>ShbA</sup> during normal growth and σ<sup>R</sup> during disulphide stress, which would also help to maintain σ<sup>HbdR</sup> activity when stress-induced alternative σ factors accumulate in the cell (55,56). Furthermore, the Group 2 σ factors σ<sup>HbdR</sup> in <i>S. coelicolor</i> and σ<sup>R</sup> in <i>M. tuberculosis</i> are induced by stress (23,57), which might also contribute to maintaining the ‘housekeeping’ duties of RNAP during stress.

Although RbpA might play an important role in promoting the formation of fully active σ<sup>HbdR</sup> σ<sup>A</sup>-holoenzyme, we show that RbpA is also poised to play additional roles in the subsequent stages of transcription initiation, such as promoter binding, and the formation and stability of open complexes. Our ChIP-qPCR data suggest that RbpA<sup>Sc</sup> is present at RNAP-promoter complexes <i>in vivo</i> that have been trapped using rifampicin. As rifampicin does not seem to influence promoter binding or open complex formation, but rather blocks transcription when the nascent transcript is ~3 nt in length (58), RbpA might play a role in any of the preceding stages of initiation. We mapped the primary binding site of RbpA<sup>Sc</sup> to the σ<sup>2</sup> domain of σ<sup>HbdR</sup> (conserved regions 1.2 and 2.1–2.4), although we cannot rule out additional weaker or transient contacts elsewhere on the protein. The σ<sup>2</sup> domain has several key roles in transcription initiation. As well as forming the largest interface between σ and RNAP (10), it confers promoter specificity for the −10 element and provides aromatic side chains that nucleate and maintain the strand-separated state during the formation of the open complex (6,59). Furthermore, variation in the sequence of the discriminator element that lies between the −10 element and the +1 transcription start point can influence stabilizing interactions with σ region 1.2, which is exploited by factors such as ppGpp that further destabilize open complexes (60). An understanding of the mechanism of action of RbpA is an important future goal.

In a similar manner to RbpA, the Crl protein of enteric bacteria interacts with the σ<sub>2</sub> domain of the stationary phase σ factor σ<sup>2</sup> and stimulates holoenzyme formation, thereby contributing to the ability of σ<sup>2</sup> to compete with σ<sup>70</sup> during stationary phase or under stress, despite its lower abundance (19,21). Crl seems to be specific for σ<sup>2</sup>, and although <i>crl</i> mutants were originally discovered by their defect in curli fimbriae formation, they are in fact defective in the expression of a large subset of σ<sup>2</sup>-dependent genes. Nonetheless, it was found that Crl also improved the activity of σ<sup>70</sup>-RNAP when mixed with σ<sup>70</sup> before formation of holoenzyme (18), suggesting that Crl might also play a non-essential role in formation of the vegetative holoenzyme. Although Crl and RbpA are not obviously related in protein sequence, the two proteins might play an analogous role in transcription initiation, for example, through modulating the interaction of σ factors with RNAP and promoting the formation and activation of certain holoenzyme complexes.

How do we reconcile our observations that RbpA interacts with the σ subunit with previous proposals that RbpA<sup>Mt</sup> principally interacts with the β-subunit of RNAP (25,26)? First, it should be noted that previous studies involved attaching chemical probes to the single cysteine residue in RbpA<sup>Mt</sup> located in the RCD, which does not interact with σ. Moreover, the cysteine is likely to be important in proper folding of the RCD (Figure 5), which might have caused anomalous results, although evidence was provided that the labelled proteins were functional. Alternatively, and perhaps more likely, RbpA might interact with both σ and β, through distinct regions of RbpA. For example, although the σ<sup>2</sup> and the β sandwich-barrel hybrid motif are distantly located in the holoenzyme structure, the disordered regions of RbpA<sup>Mt</sup> would have significant conformational freedom to bridge between these regions. In the structurally related ribosomal protein L27, the N-terminal 1–19 residues are unstructured in its free form, but extend to a distance of >35 Å from the core domain when assembled with the ribonucleoprotein complex.

The RCD domains of <i>S. coelicolor</i> and <i>M. tuberculosis</i> have virtually identical structures, but differ in their ability to bind zinc. This correlates with the replacement of three of the likely zinc ligands in RbpA<sup>Sc</sup> with hydrophilic residues in RbpA<sup>Mt</sup> that stabilize the domain through a hydrogen bond network. As the zinc-coordinating ligands are only absent from the <i>Corynebacterium– Mycobacterium–Nocardia</i> (CMN) group of actinobacteria, but are present in certain deeply rooted actinobacteria, such as <i>Bifidobacterium</i> sp. (61), it is likely that zinc-free RbpA evolved soon after the emergence of the CMN group. The RbpA proteins, therefore, present an example of how the evolution of a structural domain can contribute to a reduction in zinc requirements, which has also been described for alternative
zinc-free ribosomal proteins (62,63). However, it is unlikely that this adaptation is linked to pathogenesis because the majority of CMN group comprises free-living non-pathogenic strains. Although mutations of the three zinc-coordinating cysteine residues seemed to inactivate or destabilize RbpA, the RCD domain does not seem to be required for σ interaction. BACTH analysis and alanine scanning mutagenesis of RbpA revealed that the C-terminal predicted helical region was required for this interaction, and that the highly conserved amino acid residues R89 and R90 play a critical role, although it is not yet clear whether they contact σ directly, or whether their mutation disrupts the overall conformation of this region. Although the C-terminal region is predicted to be helical, NMR analysis of the full-length RbpA proteins from both S. coelicolor and M. tuberculosis suggested that this region in each protein is unstructured, and despite extensive attempts to prevent it, it is readily proteolysed to the RCD during storage. The binding to σ, therefore, likely promotes folding of the RbpA C-terminal region, thereby stabilizing the complex. Clearly, future structural characterization of RbpA either in a binary complex with σ or in higher order complexes with RNAP will be highly valuable in understanding further mechanistic aspects of the function of this protein. This is likely to be of particular importance in the development of new approaches to target the process of transcription in M. tuberculosis and other pathogenic actinobacteria.

ACCESSION NUMBERS

Atomic coordinates for RbpA Sc and RbpA Mt have been deposited with the Protein Data Bank under accession codes 2m6o and 2m6p, respectively. NMR chemical shifts for RbpA Sc and RbpA Mt have been deposited with the Biological Magnetic Resonance DataBank under accession codes 19148 and 19149, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3, Supplementary Figures 1 and 2 and Supplementary References [64–73].

ACKNOWLEDGEMENTS

The authors are grateful to Chris Dadswell, Sophie Buitenhuys and Yichao Zhou for contributions to experiments.

FUNDING

A.T.-S. B.L. and R.J.L. were funded by BBSRC grant [BB/1003045 to M.S.P. and S.J.M.]; P.D. was funded by a BBSRC grant [BB/D018293/1 to M.S.P.]; A.T.-S. was funded by the University of Sussex and the Overseas Research Students Awards Scheme. Funding for open access charge: BBSRC grant [BB/1003045].

Conflict of interest statement. None declared.

REFERENCES

of rifampicin resistance on Streptomyces coelicolor.


