

Molecular basis of *Streptomyces* ECF σ^{ShbA} factors transcribing principal σ^{HrdB} genes

Guiyang Liu¹, Xu Yang¹, Wenjin Yan¹, Yiqun Wang¹, Feng Yu¹, Jianting Zheng^{1,2,*}

¹State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

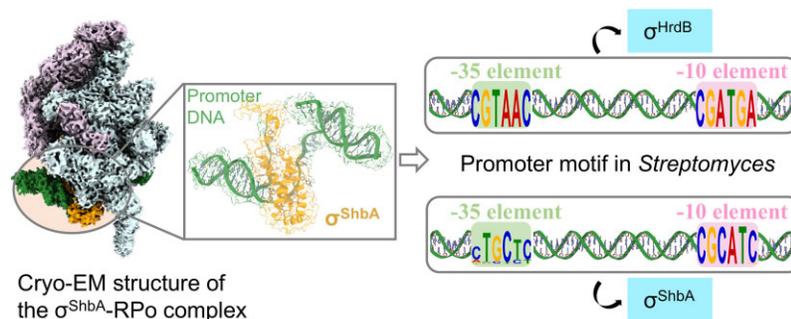
²Joint International Research Laboratory of Metabolic & Developmental Sciences, Shanghai Jiao Tong University, Shanghai 200240, China

*To whom correspondence should be addressed. Email: jtzheng@sjtu.edu.cn

Abstract

In bacteria, principal σ factors (σ^{70} or σ^A) transcribe housekeeping genes required for cell viability. Although most principal σ genes are transcribed by the RNA polymerase (RNAP) holoenzyme containing the principal σ factor itself, an extracytoplasmic function (ECF) σ factor (σ^{ShbA}) governs transcription of the principal σ factor gene (*hrdB*) in two model *Streptomyces*. Here, we employed a combination of cryo-electron microscopy (cryo-EM) and bioinformatics to decipher how σ^{ShbA} -RNAP holoenzymes govern the transcription of *hrdB* genes in *Streptomyces*. A cryo-EM structure of *Streptomyces coelicolor* σ^{ShbA} -RNAP-promoter open (RPO) complex was solved at 2.97 Å resolution. In combination with *in vitro* transcription assays, we demonstrate the unique structural features used by the σ^{ShbA} to recognize the *hrdB* promoter and form a transcription bubble. All *Streptomyces* genomes (603) tagged as ‘reference’ were retrieved from NCBI Datasets. The conserved protein sequences and genomic neighborhoods, as well as the promoter consensus sequences of σ^{ShbA} and σ^{HrdB} homologs, support that the principal σ^{HrdB} being governed by the ECF σ^{ShbA} is a common feature in *Streptomyces*. Overall, these results provide detailed molecular insights into the transcription of the principal σ^{HrdB} gene and pave the way for globally modulating *Streptomyces* cell viability.

Graphical abstract



Introduction

The RNA polymerase (RNAP) holoenzyme, which consists of an RNAP core enzyme and a dissociable σ factor, is the transcription machine of bacteria [1, 2]. The core enzyme composed of five subunits ($\alpha_2\beta\beta'\omega$) is responsible for binding to template DNA to synthesize RNA, while the σ factor recognizes the promoter sequence to initiate promoter-specific transcription [3]. Most bacteria contain a principal σ factor (σ^{70} in *Escherichia coli* and σ^A in other bacteria) to govern transcription that is essential for cell viability [4], as well as a variety of alternative σ factors to initiate transcription-specific genes involved in diverse functions under particular conditions, such as stress responses, morphogenesis, etc [5]. The σ^{70} -type factors are classified into four groups based on their conserved regions [6]. Group I σ factors (or principal

σ factors) contain all conserved regions ($\sigma\text{R1.1}$, $\sigma\text{R1.2}$, σR2 , σR3 , and σR4) and an additional nonconserved region (NCR) (Supplementary Fig. S1). Groups II σ factors are closely related to group I but lack $\sigma\text{R1.1}$. Groups III σ factors contain σR2 , σR3 , and σR4 while group IV σ factors, which are also referred to as “extracytoplasmic σ factors” (ECF σ factors), only contain σR2 and σR4 [6, 7].

The σR2 interacts with the -10 element while the σR4 binds to the -35 element of the promoter. The consensus sequences are very diverse for the promoter -10 element and -35 element. For example, *E. coli* principal σ^{70} recognizes promoters with “TATAAT” (-10 element) and “TTGACA” (-35 element) consensus sequences while *Clostridium thermocellum* ECF σ^1 recognizes promoters with CGWA (-10 element) and A-tract (-35 element) consensus sequences [8, 9].

Received: October 9, 2024. Revised: March 26, 2025. Editorial Decision: April 6, 2025. Accepted: April 22, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Although structures of RNAP in complex with a variety of principal (σ^{70} and σ^A) and ECF (σ^L , σ^H , σ^E and σ^I) σ factors have been reported [8, 10–17], mechanisms by which bacterial σ factors specifically recognize promoters to initiate transcription remains incompletely understood.

Streptomyces are important industrial microorganisms capable of producing numerous antibiotics and other bioactive natural products with important applications in medicine and agriculture [18]. They are multicellular bacteria with complex developmental cycles. A large number of σ factors are encoded in streptomyces genomes [19]. For example, *Streptomyces coelicolor* and *Streptomyces griseus* have 64 and 52 σ factors, respectively [20, 21]. Most of these σ factors belong to the ECF group. A few of them have been characterized in *S. coelicolor*, including σ^E responding to cell envelope stress [22], σ^R responding to thiol-perturbing signals [23], and σ^{BldN} involved in aerial mycelium formation [24]. σ^L , a group III σ factor, directly regulates lincomycin biosynthesis in *Streptomyces lincolnensis* [25]. *Streptomyces* σ^{HrdB} is a functional homolog of *E. coli* σ^{70} , governing the expression of housekeeping genes whose products are essential for cell viability [26]. For many bacteria, the principal σ factor gene is transcribed by RNAP holoenzyme containing the principal σ factor itself [27–30]. However, an alternative σ factor (σ^{ShbA}) is the major determinant of transcription of the σ^{HrdB} gene (*hrdB*) in *S. coelicolor* and *S. griseus* [20, 31]. *hrdB* promoter sequences (CGTAAC-N₁₆-CGATGA) of these two species have little similarity to the consensus sequence [Y][S][W][M][S]-N_{16–18}-[Y]ANNNT (W = A/T; M = A/C; S = C/G; Y = C/T; R = A/G) of σ^{HrdB} -dependent promoters, especially the –10 element [20, 32]. Besides *hrdB* gene, *shbA* gene itself is the only other target of σ^{ShbA} identified by RIViT-seq in *S. coelicolor* [31]. Control of the principal σ^{HrdB} by the alternative ECF σ^{ShbA} is proposed to enable *Streptomyces* to control mycelial growth more sensitively in response to environmental changes.

Despite the importance of σ^{ShbA} in controlling the principal σ^{HrdB} , how the σ^{ShbA} specifically recognizes the *hrdB* promoters and efficiently initiates transcription is still unknown. Here, we solved the cryo-electron microscopy (cryo-EM) structures of *S. coelicolor* RNAP-promoter open complex comprising σ^{ShbA} (σ^{ShbA} -RPO). The interactions between RNAP holoenzyme and promoter DNA observed in the RPO structure reveal the molecular basis for the promoter specificity of σ^{ShbA} . Bioinformatic analysis of *Streptomyces* reference genomes in the NCBI Datasets suggests that the principal σ^{HrdB} is governed by the σ^{ShbA} in all these *Streptomyces*.

Materials and methods

Purification of *Sc* RNAP core enzyme, σ^{HrdB} , rbpA, and CarD

The *Sc* RNAP core enzyme, σ^{HrdB} , RbpA and CarD were purified using previously established methods [33, 34]. For the purification of RNAP core, *E. coli* BL21(DE3) cells harboring the pET28a-RNAP core plasmid were grown at 37°C until reaching an OD₆₀₀ of ~0.6, followed by induction with 0.3 mM isopropyl- β -D-thiogalactopyranoside (IPTG) at 16°C for 16 h. The cells were harvested, washed, resuspended in buffer A [50 mM Tris (pH 8.0), 500 mM NaCl, 5 mM MgCl₂, and 10% glycerol], then lysed via sonication. The supernatant was loaded onto a Ni-NTA agarose column, washed with buffer A supplemented with 20 mM imidazole, and eluted with buffer

B [20 mM Tris (pH 8.0), 300 mM NaCl, 5 mM MgCl₂, 10% glycerol (v/v), and 250 mM imidazole]. The eluted fractions were pooled and applied to a HiTrap Heparin HP column pre-equilibrated with buffer C (20 mM Tris, pH 8.0, 5 mM MgCl₂, 10% glycerol). The proteins were eluted using a linear gradient of buffer D (20 mM Tris, pH 8.0, 1 M NaCl, 5 mM MgCl₂, 10% glycerol) from 0% to 100% over 20 ml, with buffer C as the starting buffer. RNAP core enzyme fractions were concentrated and polished by gel filtration on a Superose 6 column equilibrated with buffer E [20 mM Tris (pH 8.0), 150 mM NaCl, 5 mM MgCl₂, 3 mM dithiothreitol (DTT), and 10% glycerol]. Purity was confirmed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Supplementary Fig. S2A), and the RNAP core was concentrated to 8 mg/ml.

For σ^{HrdB} purification, a similar protocol was followed, with the exception that a Superdex 200 gel filtration column was used for final polishing (Supplementary Fig. S2B), and the protein was concentrated to 5 mg/ml. RbpA and CarD were purified using a procedure analogous to σ^{HrdB} (Supplementary Fig. S2C and D), omitting the HiTrap Heparin step, and both were concentrated to 5 mg/ml. All purified proteins were stored at –80°C.

Purification of σ^{ShbA} and its mutants

The *S. coelicolor* σ^{ShbA} gene was amplified from strain M145 and cloned into pET28a vector using primer pairs ShbA F and ShbA R. Expression was carried out in *E. coli* strain BL21(DE3). Transformed cells were grown in 1L LB broth supplemented with 100 μ g/ml kanamycin at 37°C until reaching an OD₆₀₀ of 0.6. Protein expression was induced with 0.3 mM IPTG, followed by overnight incubation at 16°C. Cells were harvested by centrifugation (5000 \times g; 7 min at 4°C), resuspended in 50 ml of buffer A, and lysed. The supernatant was obtained by centrifugation (10 500 \times g; 40 min at 4°C) and loaded onto a 5 ml Ni-NTA agarose column. The column was washed by 25 ml buffer A with 20 mM imidazole, followed by elution with 4 ml of buffer B. The eluted σ^{ShbA} fractions were loaded onto a Superose 6 column equilibrated with buffer E. The purified σ^{ShbA} were concentrated to 5 mg/ml and stored at –80°C (Supplementary Fig. S2E).

A truncated variant, $\sigma^{\text{ShbA}}_{\Delta R4}$ (M1 ~ L142) was cloned from *S. coelicolor* M145 genomic DNA and inserted into the pET28a using primers ShbA $_{\Delta R4}$ F and ShbA $_{\Delta R4}$ R. Purification of the truncated variant followed the same protocol as that for the wild-type protein (Supplementary Fig. S3).

Site-directed mutagenesis of σ^{ShbA} was performed using a modified protocol based on the GeneTailor™ Site-Directed Mutagenesis System. Briefly, the pET28a- σ^{ShbA} plasmid was used as a template for inverse PCR with overlapping primers, one of which contained the desired mutation. The reaction products were digested with *DpnI* to remove the template DNA and transformed into *E. coli*. Plasmids carrying the desired mutation were isolated and verified by sequencing. Both the forward and reverse primers were approximately 35 nucleotides in length, with a 16–18 nucleotide overlap at their 5' ends to facilitate efficient joining of the mutagenized product. The mutation site was located downstream of the overlapping region in the forward primer. The mutagenic primers included at least 15 nucleotides downstream of the mutation site to ensure efficient annealing. The primers used are listed in Supplementary Table S1. Purification of the σ^{ShbA} mutants

followed the same protocol as that for the wild-type protein (Supplementary Fig. S3).

Assembly of the σ^{HrdB} -RNAP and σ^{ShbA} -RNAP holoenzyme

σ^{HrdB} -RNAP was assembled as previously described [33]. In brief, *S. coelicolor* RNAP core and σ^{HrdB} were incubated at room temperature for 30 min in a 1:4 ratio. The fractions containing the *S. coelicolor* σ^{HrdB} -RNAP holoenzyme were purified using a Superose 6 gel filtration column with buffer E. The peak fraction containing the σ^{HrdB} -RNAP holoenzyme complex was collected and concentrated to 5 mg/ml (Supplementary Fig. S4A). Assembly of σ^{ShbA} -RNAP followed the same protocol and concentrated to 10 mg/ml (Supplementary Fig. S4B).

In vitro transcription

Transcription templates containing the *hrdB*, *shbA*, and *SCO1145* promoter sequences, along with the MangoIII sequence were synthesized by PCR amplification and purified using the SanPrep Column PCR Product Purification Kit (Sangon Biotech). The *hrdBp1*, *shbAp1*, *shbAp2*, and *SCO1145p* promoter regions was inserted between the EcoRI and BamHI sites of the pUC19 vector, respectively, and the 47 bp MangoIII sequence was inserted between the BamHI and HindIII sites. Using the resulting plasmids, transcription templates and mutations of the promoters were generated by PCR amplification using corresponding primers (Supplementary Table S1).

Transcription activities were evaluated using the MangoIII-based transcription assay, as previously described [33]. Transcription reactions were carried out with 100 nM of σ^{ShbA} -RNAP (or σ^{HrdB} -RNAP) and 1 mM of NTPs (0.25 mM each) in a transcription buffer consisting of 20 mM Tris-HCl (pH 8.0), 100 mM KCl, 5 mM MgCl₂, 1 mM DTT, 4 U RNaseIn, 1 μ M TO1-PEG-biotin, and 5% glycerol. Reactions were incubated at 30°C for 10 min and terminated by adding heparin to a final concentration of 0.5 ng/ μ l. A control reaction without RNAP was included. Fluorescence intensities were measured using a Tecan Spark multi-detection microplate reader with excitation at 510 nm and emission at 535 nm. Statistical analysis was performed using a two-tailed unpaired *t*-test assuming Gaussian data distribution in GraphPad Prism 8. The test statistic was calculated using the formula $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$, where \bar{x} is the sample mean; μ is the hypothesized population mean; s is the sample standard deviation; and n is the sample size. The corresponding *P*-value was determined based on the *t*-distribution with $n-1$ degrees of freedom. Differences were considered significant at *P*-value was $<.05$. Significance levels are indicated as follows: $*P < .05$, $**P < .01$, $***P < .001$, and $****P < .0001$, while nonsignificant results ($P > .05$) are labeled as “ns”.

Assembly of the σ^{ShbA} -RPO complex

Synthetic oligos (*hrdBp1*-NT-6 bubble/*hrdBp1*-T-6 bubble and *shbAp1*-NT-6 bubble/*shbAp1*-T-6 bubble, listed in Supplementary Table S1) were annealed to prepare double-stranded nucleic acid scaffolds for the cryo-EM study of the σ^{ShbA} -RPO complexes. The annealing process involved heating the DNA at 95°C for 5 min, followed by a slow cooling to room temperature in an annealing buffer containing 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 10 mM MgCl₂.

For cryo-EM sample preparation, the purified *S. coelicolor* σ^{ShbA} -RNAP holoenzyme was mixed with the annealed nucleic-acid scaffold (*hrdBp1* or *shbAp1*) at a 1:1.4 molar ratio and incubated at 4°C overnight. The reconstituted σ^{ShbA} -RPO complex was further purified using a Superose 6 column equilibrated with buffer F (20 mM Tris-HCl, pH 8.0, and 75 mM NaCl, 5 mM MgCl₂, 1% glycerol, 3 mM DTT). The peak fraction of the σ^{ShbA} -RPO complex was collected at a concentration of 0.8 mg/ml for cryo-EM sample preparation. The composition of the RPO complex was verified by SDS-PAGE (Supplementary Fig. S4C and D).

Cryo-EM data acquisition and processing

The σ^{ShbA} -RPO complex was applied to freshly glow-discharged Quantifoil R1.2/1.3 Au 300 mesh grids and plunge-frozen in liquid ethane using a Vitrobot (FEI) at 16°C with 100% chamber humidity and 2 s blotting time. Imaging was performed on a Titan Krios 300 kV microscopes equipped with a K3 detector. Data were collected at a defocus range of -0.8 to -1.4 μ m, a magnification of $\times 130\,000$ in counting mode, and a total dose of 40 $e^-/\text{\AA}^2$, with 32 frames per movie.

Data processing was carried out using CryoSPARC suite v4.2.1 [35]. The workflow included motion correction, patch CTF estimation, manual exposure curation, and template picking based on selected 2D classes from blob picking, followed by 2D classification. Particles from well-defined 2D classes were subjected to ab-initio reconstruction, and the best 3D class was selected for further refinement to generate the final map. For the σ^{ShbA} -RPO-*hrdBp1*, a total of 5793 movies were collected. The best 3D class (335 850 particles, 42%) was selected for local resolution estimation and filtering to produce the final map (Supplementary Fig. S5). For the σ^{ShbA} -RPO-*shbAp1*, a total of 3 172 movies were collected at a defocus range of -1.0 to -1.6 μ m. The best 3D class (85 184 particles, 36%) was selected for local resolution estimation and filtering to produce the final map (Supplementary Fig. S6).

Cryo-EM model building and refinement

Initial models for the *S. coelicolor* RNAP core enzyme were derived from the previous structure (PDB:8HVR) [34]. The σ^{ShbA} atomic model was predicted using AlphaFold 3 [36], while the promoter DNA model was manually constructed in Coot [37]. These models were integrated into the cryo-EM density maps utilizing ChimeraX [38]. Refinement of the models was performed in Coot and Phenix, applying secondary structure, rotamer, and Ramachandran restraints [39]. Validation was executed using MolProbity [40]. The Fourier Shell Correlation between the map and model was generated by Phenix. Comprehensive statistics on the cryo-EM data processing and refinement are provided in Supplementary Table S2.

Bioinformatics analysis

603 *Streptomyces* genomes tagged as ‘reference’ were downloaded from NCBI Datasets (<https://www.ncbi.nlm.nih.gov/datasets/genome/>) in May 2024. The ‘reference’ subset was specifically chosen to minimize bias towards overrepresented species. A total of 604 genomes, including *S. coelicolor*, were included in the bioinformatics analysis. To identify σ^{ShbA} and σ^{HrdB} homologs across *Streptomyces* genomes, we used Biopython [41]. Specifically, the “SeqIO” module was used to input the fasta-formatted *S. coelicolor* σ^{ShbA} (UniProt entry Q9L0I8)

or σ^{HrdB} (UniProt entry P18183), along with genome data in gbff format. PairwiseAligner, a class within the Bio.Align module, was used for global and local alignments. This tool leverages the Needleman–Wunsch, Smith–Waterman, Gotoh (three-state), Waterman–Smith–Beyer global and local pairwise alignment algorithms, and the Fast Optimal Global Alignment Algorithm (FOGSAA) [42]. The protein with the highest sequence identity to σ^{ShbA} or σ^{HrdB} in each genome was identified as the respective homolog (Supplementary Tables S3 and S4). Multiple-sequence alignments were prepared using ClustalX.

To investigate functional domains within the identified proteins, we utilized InterPro (<https://www.ebi.ac.uk/interpro/search/>) and HMMERScan (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>). Additionally, sequence similarity calculations were performed using a local EMBOSS Water tool (Supplementary Tables S3 and S4).

For genomic context analysis, the upstream and downstream regions of *shbA* homologs and *hrdB* homologs were examined (excluding 2 and 5 incomplete genomic sequences, respectively). Promoter regions (300 nucleotides upstream of the start codons) were extracted for *hrdB* and *shbA* genes and analyzed using the MEME motif discovery tool within the MEME Suite [43]. The MEME analysis parameters were set as follows: distribution (zero or one), width (3~60, allowing identification of either the complete promoter or a segment of the bipartite target promoter), and an optional parameter (searching the given strand only).

Results

Transcription activity of σ^{ShbA} on *hrdB* and *shbA* promoters *in vitro*

To evaluate the transcription activity of *S. coelicolor* σ^{ShbA} -RNAP *in vitro*, the RNAP core enzyme was expressed and purified as described previously (Supplementary Fig. S2A) [33]. The gene encoding σ^{ShbA} was cloned into the pET28a and expressed in *E. coli* (Supplementary Fig. S2E). The σ^{ShbA} -RNAP holoenzyme was prepared by mixing RNAP core enzyme with σ^{ShbA} at a 1:4 molar ratio, and excess σ^{ShbA} was removed by size exclusion chromatography (SEC) (Supplementary Fig. S4B). The *hrdB* gene in *S. coelicolor* has two promoters: p1, the primary promoter, and p2, which is transcribed by the σ^{R} factor under oxidative stress (Supplementary Fig. S7) [23]. As shown in Fig. 1A, the σ^{ShbA} -RNAP holo-enzyme transcribes *hrdBp1* more efficiently than the σ^{HrdB} -RNAP holo-enzyme, which preferentially transcribes the *SCO1145* promoter, a member of the σ^{HrdB} regulon [32]. ECF σ factors usually autoregulate their own transcription [44]. *In vitro* transcription-sequencing (RIViT-seq) confirms that the *shbA* promoter is a target of σ^{ShbA} in *S. coelicolor* [31]. Additionally, *S. coelicolor* dRNA-seq data show that *shbA* has two promoters, with p1 being the primary promoter (Supplementary Fig. S8) [21]. Similarly, in *S. griseus*, *shbA* is shown to possess two promoters based on S1 nuclease mapping [20]. Our *in vitro* assays demonstrate that the σ^{ShbA} -RNAP holo-enzyme transcribes *shbAp1* more efficiently than *shbAp2* (Supplementary Fig. S9).

Overall structure of *S. coelicolor* σ^{ShbA} -RPO

We synthesized a 56-bp DNA scaffold derived from *S. coelicolor* *hrdBp1* promoter, comprising a 35-bp (−41 to −7)

upstream double-stranded DNA (dsDNA), a 6-bp (−6 to −1) noncomplementary bubble and a 15-bp (+1 to +15) downstream ds-DNA (Fig. 1B). The σ^{ShbA} -RPO-*hrdBp1* complex was prepared by mixing the σ^{ShbA} -RNAP holo-enzyme with the DNA scaffold, followed by purification using SEC (Supplementary Fig. S4C). The structure of σ^{ShbA} -RPO-*hrdBp1* was determined by single-particle cryo-EM (Supplementary Fig. S5). A total of 335 850 single particles were selected to reconstruct the cryo-EM map, which was refined to a nominal resolution of 2.97 Å, with a resolution of approximately 2.7 Å at the center of RNAP and 5.5 Å at the peripheral regions.

The primary channel formed by β and β' subunits adopts a narrow conformation in σ^{ShbA} -RPO-*hrdBp1* (Supplementary Fig. S10), resembling previously reported *S. coelicolor* σ^{HrdB} -RPO and other bacterial RNAP structures [33, 34, 45]. The RNAP core enzyme was resolved unambiguously in the cryo-EM map, except for the ω subunit, which was invisible (Fig. 1C and D). Notably, the purified σ^{ShbA} -RPO-*hrdBp1* contains the ω subunit (Supplementary Fig. S4C). In σ^{HrdB} -RPO structures [33, 46], the last 13 residues of the β' subunit fold into a β strand and an α helix to stabilize the ω subunit (Supplementary Fig. S11). The disorder of these residues in the σ^{ShbA} -RPO-*hrdBp1* structure likely explains the absence of the ω subunit in the cryo-EM map. The σ R2, the linker connecting σ R2 and σ R4, and the −23 to +13 of the DNA scaffolds were reliably modeled (Fig. 2A,B), whereas the σ R4 and the corresponding −35 element of the DNA scaffolds cannot be traced in the cryo-EM map, suggesting that they are highly flexible in the σ^{ShbA} -RPO-*hrdBp1* complex. Compared to the wild-type σ^{ShbA} -RNAP, the RNAP holoenzyme assembled from a truncated protein lacking σ^{ShbA} R4 showed no significant difference when transcribing the *hrdBp1* *in vitro* (Supplementary Fig. S12A).

The σ R4 is known to form hydrophobic interactions with the flap-tip helix (β FTH) of the β subunit in many bacterial RNAP structures [47]. The disorder of σ^{ShbA} R4 may explain the $\sim 45^\circ$ rotation of β FTH in the σ^{ShbA} -RPO-*hrdBp1* structure compared to the σ^{HrdB} -RPO (Supplementary Fig. S13A). Additionally, the duplex DNA upstream of −10 element moves away from the β' clamp domain in the σ^{ShbA} -RPO-*hrdBp1* structure (Supplementary Fig. S13B), likely due to the disorder of the σ^{ShbA} R4. Similar to σ^{HrdB} R2, σ^{ShbA} R2 folds into three helices ($\alpha 2$ – $\alpha 4$) (Fig. 2B). Helix $\alpha 3$ binds to the β' subunit coiled-coil region (β' CC) at the RNAP clamp domain (Supplementary Fig. S14A). Residues D60 and E64 form electrostatic interactions with R350, R353, and R356 of β' CC, while Q63 forms a H-bond with E370 of β' CC. These residues are conserved in σ^{HrdB} , consistent with the similar binding of the two σ factors to β' CC (Supplementary Fig. S14B and C). The linker connecting σ^{ShbA} R2 and σ^{ShbA} R4 adopts an extended conformation, interacting with template-strand single-stranded DNA (ssDNA) and the RNA-exit channel, similar to σ^{HrdB} R3.2 (Supplementary Fig. S15A–C).

To further investigate the interactions between σ^{ShbA} and different promoters, we assembled σ^{ShbA} -RPO-*shbAp1* (Supplementary Fig. S16A). The cryo-EM map was refined to a resolution of 3.61 Å (Supplementary Fig. S6). Similar to the σ^{ShbA} -RPO-*hrdBp1* structure, the density of σ^{ShbA} R4, the ω subunit, and the upstream double-stranded DNA (dsDNA) was invisible (Supplementary Fig. S16B). Additionally, the density of β FTH was weak in the σ^{ShbA} -RPO-*shbAp1* structure. The resolution of σ^{ShbA} R2 and the −10 element DNA of

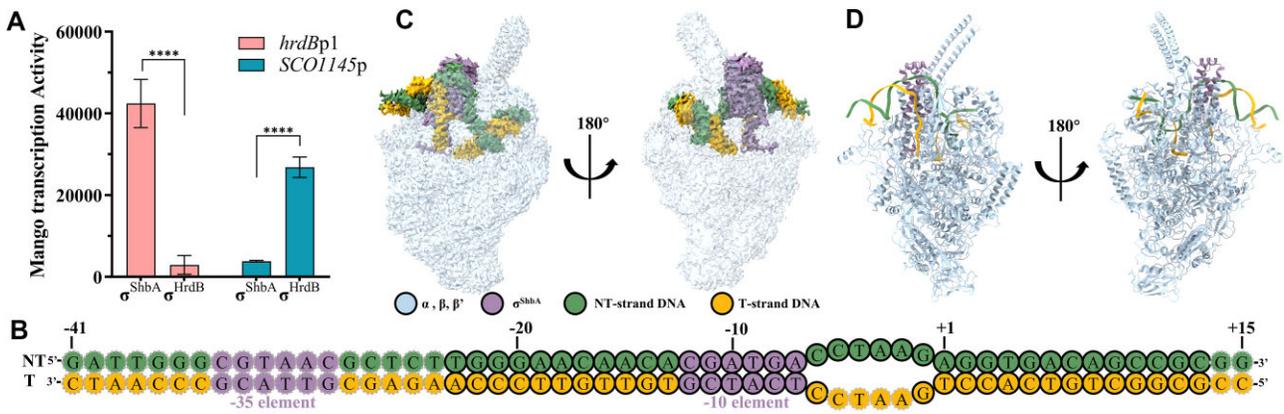


Figure 1. The cryo-EM structure of *S. coelicolor* σ^{ShbA} -RPO-*hrdBp1*. **(A)** *In vitro* transcription assay verified the transcriptional activities of σ^{ShbA} and σ^{HrdB} on *hrdBp1* and *SCO1145p*, respectively. Mean fold-change values were compared using Student's *t*-test. Error bars represent mean \pm standard deviation (SD) of three independent experiments. **** denotes $P < .0001$. **(B)** The *hrdB* promoter fragment used for σ^{ShbA} -RPO-*hrdBp1* assembly. The -10 element, -35 element, and TSS are labeled. The dotted circles were used to show nucleotides that could not be modeled in the structure due to their weak density. **(C)** The cryo-EM map of *S. coelicolor* σ^{ShbA} -RPO-*hrdBp1*. Each subunit and DNA strand are color-coded as shown in the key. **(D)** Cartoon model of *S. coelicolor* σ^{ShbA} -RPO-*hrdBp1* highlighting the overall architecture of the complex.

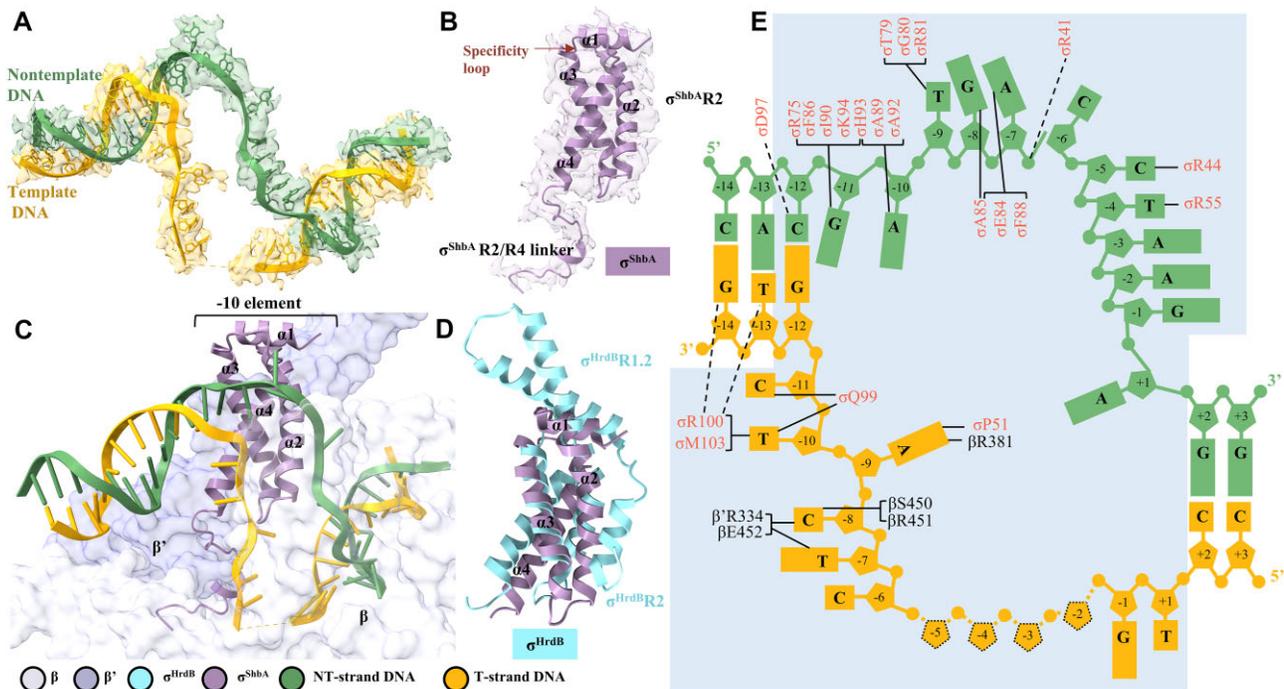


Figure 2. The interactions between σ^{ShbA} R2 and DNA. **(A)** Density map and model of nucleic acids. **(B)** Ribbon representation and density map of σ^{ShbA} R2/R4 linker. **(C)** σ^{ShbA} R2 binds to the -10 element in the major groove of DNA. **(D)** Structural comparison of σ^{ShbA} R2 and σ^{HrdB} R2. Structures are colored according to the color key. **(E)** Schematic contacts between σ^{ShbA} R2 and DNA. Residues making specific contacts with DNA are labeled. Hydrogen bonds are shown as dashed lines.

the *shbAp1* promoter was approximately 6 Å (Supplementary Fig. S16C). The structure of σ^{ShbA} and promoter DNA in the σ^{ShbA} -RPO-*shbAp1* complex closely resembles that in the σ^{ShbA} -RPO-*hrdBp1* complex (Supplementary Fig. S17). Consequently, subsequent analysis of the σ^{ShbA} -RPO structure focused on the higher-quality map of the σ^{ShbA} -RPO-*hrdBp1* complex.

Interactions between σ^{ShbA} R2 and -10 element
 σ^{ShbA} R2 binds within the DNA major groove via helices $\alpha 2$ and $\alpha 4$, engaging the -10 element and the transcriptional

bubble (Fig. 2C). Helix $\alpha 1$ of σ^{ShbA} R2 structurally resembles the second helix of σ^{HrdB} R1.2, occupying a surface pocket of σ^{ShbA} R2 (Fig. 2D). Helices $\alpha 1$ and $\alpha 2$ of σ^{ShbA} R2 are significantly shorter than their counterparts in σ^{HrdB} R2 (Fig. 2D). In σ^{HrdB} R2, the NCR between $\sigma 1.2$ and $\sigma 2$ forms additional helices that serve as a docking platform for RbpA, a transcription activator critical in *Actinobacteria* [48, 49]. NCR is missing in ECF σ factors, consistent with the observation that RbpA has a negligible effect on *in vitro* transcription of σ^{ShbA} -RNAP (Supplementary Fig. S12B). CarD, another transcription activator of principal σ factors that interact with the β subunit and upstream double-stranded/single-

stranded (ds/ss) junction of the transcription bubble [50, 51], also has no obvious effect on the transcription of σ^{ShbA} -RNAP (Supplementary Fig. S12B), further highlighting the intrinsic difference between σ^{ShbA} and principal σ factors in transcription.

In the σ^{ShbA} -RPO structure, σ^{ShbA} R2 unwinds the -10 element DNA duplex to form a 12 bp transcription bubble (from -11 to $+1$), which including the pre-melted 6-bp region (Fig. 2C and E). The bubble size is consistent with other ECF σ factors (12 bp) [14], but differs from principal σ factors, which typically form larger bubbles (13–15 bp) [33]. The conservation of the -10 element is structurally explained by specific interactions between σ^{ShbA} R2 and the promoter. Helix $\alpha 4$ and the specificity loop connecting $\alpha 3$ and $\alpha 4$ primarily recognize the six nucleotides of the -10 element (Fig. 2C). Specifically, the -12 C:G pair forms interactions with D97 (Fig. 3A). H93 serves as a wedge to disrupt the -11 G:C pair, initiating bubble formation (Fig. 3C). Consequently, the base moiety of $G_{-11}(\text{nt})$ is displaced from the DNA major groove into a shallow pocket formed by F86, I90, H93, and K94 (Fig. 3A and C). The base moiety of $A_{-10}(\text{nt})$ remains in the DNA major groove (Fig. 3C), forming a hydrogen bond with the sugar moiety of $C_{-12}(\text{nt})$ and Van der Waal interactions with H93 (Fig. 3A). The next three nucleotides (-9 to -7) are flipped out and positioned in a surface pocket of σ^{ShbA} (Fig. 3B). $T_{-9}(\text{nt})$ interacts with T79, G80, and R81 located at the C-terminal of the “specificity loop”. $A_{-7}(\text{nt})$ forms interactions with E84, A85, and F88 (Fig. 3A). $G_{-8}(\text{nt})$ stacks between $A_{-7}(\text{nt})$ and A85. Moreover, extensive electrostatic interactions are formed between backbone phosphate groups of the -10 element and the “specificity loop” residues of σ^{ShbA} , including R75 and R81 (Fig. 3F). Helix $\alpha 2$ contacts $C_{-6}(\text{nt})$, $C_{-5}(\text{nt})$ via its R41 and R44 while R55 of helix $\alpha 3$ makes contacts with $C_{-5}(\text{nt})$ and $T_{-4}(\text{nt})$ of the “Core recognition element” (CRE) DNA sequence (Fig. 3A).

σ^{ShbA} R2 directs the template ssDNA into the RNAP active center cleft, which is formed by β and β' subunits (Fig. 3D). Residues Q99, R100, and M103 create a pocket to accommodate the base of $C_{-11}(\text{nt})$ and $T_{-10}(\text{nt})$ (Fig. 3E). R100 also form hydrogen bonds with $G_{-14}(\text{nt})$ and $T_{-13}(\text{nt})$ in the upstream double-stranded DNA. $A_{-9}(\text{nt})$ engages in stacking interactions with P51 of σ^{ShbA} and R381 of the β subunit (Fig. 3E). $C_{-8}(\text{nt})$ and $T_{-7}(\text{nt})$ form extensive contacts with the β subunit and R334 of the β' subunit. Alanine substitution of these σ^{ShbA} residues contacting -10 -element nucleotides resulted in modest or severe activity loss (Fig. 3G), confirming the functional significance of the observed interactions.

Conserved genomic context of σ^{ShbA} and σ^{HrdB}

To investigate if the principal σ^{HrdB} is governed by σ^{ShbA} across *Streptomyces* species, we analyzed 603 reference genomes from the NCBI Datasets (Supplementary Table S3). σ^{ShbA} homologs were identified with an average length of 196 aa (median: 195 aa) (Supplementary Figs S18A and S19A), exhibiting 70%–100% sequence identity and 88%–100% sequence similarity (Supplementary Table S3). Pfam profile analysis confirmed that all belong to ECF σ factors, containing σ R2 and σ R4 (Fig. 4A), with average flanking lengths of 32 aa (N-terminal) and 4 aa (C-terminal). ECF σ factors are recently expanded and re-classified, resulting in 157 phylogenetic ECF groups [52]. According to the ECF classification process, these proteins belong to subgroup 2 of the ECF40

group (ECF40s2). Interestingly, all members of ECF40 are from *Actinobacteria*. Genomic context analysis didn't identify putative anti- σ factors (Supplementary Table S5). Instead, 92% of *shbA* genes are genomically associated with a LysR family transcriptional regulator, a WhiB family transcriptional regulator (WhiD in *S. coelicolor*), a two-component response regulator with N-terminal NarL-like receiver and C-terminal luxR_HTH domains (BldM in *S. coelicolor*), an IMP dehydrogenase and an IMP dehydrogenase-associated GuaB3 family protein (Fig. 4C and Supplementary Fig. S19B). In *S. coelicolor*, WhiD (SCO4767) have been shown to interact with σ^{HrdB} R4 [53]; BldM (SCO4768) is essential for aerial mycelium formation [54]; and the IMP dehydrogenase (SCO4770) is a member of σ^{HrdB} regulon [32]. Using Operon-mapper, an online tool that predicts operons in bacterial genomes based on intergenic distances and functional connections between genes [55], we analyzed these adjacent genes in *S. coelicolor* and found that none of them shares an operon with σ^{ShbA} .

Next, we identified σ^{HrdB} homologs from the genome data, exhibiting sequence identities ranging from 57.6% to 100% and similarities from 64.5% to 100% (Supplementary Fig. S18B and Supplementary Table S4). These proteins have an average length of 516 amino acids (aa), with a median of 514 aa (Supplementary Fig. S19A). Using Hmmscan to analyze Pfam profiles, we confirmed the presence of all conserved regions, including σ R1.1, σ R1.2, σ R2, σ R3, and σ R4 (Fig. 4B). The average lengths before the σ R1.1 (N-terminal) and after the σ R4 (C-terminal) are 72 aa and 14 aa, respectively. Genomic context analysis revealed that over 95% of *hrdB* genes are associated with a FadR/GntR family transcriptional regulator (WhiH in *S. coelicolor*), a serine protease, a DNA topoisomerase IV subunit B (ParE in *S. coelicolor*), and an integral membrane protein (Fig. 4D, Supplementary Fig. S19B, and Supplementary Table S6). WhiH (SCO5819) is involved in spore formation, cell differentiation, segregation, and secondary metabolite production [56], while ParE is essential for proper chromosome segregation in *Streptomyces* [57]. Operon-mapper analysis of *S. coelicolor* confirmed that none of these adjacent genes reside in the same operon as *hrdB* [55].

Consensus sequences of σ^{HrdB} and σ^{ShbA} promoters

To identify the σ^{HrdB} and σ^{ShbA} promoters, we extracted 300 nt regions upstream of *hrdB* (Supplementary Table S7) and *shbA* homologs (Supplementary Table S8) and screened them for over-represented sequence motifs using MEME [43]. ECF σ factors typically autoregulate their own transcription [44], and in *S. coelicolor*, *shbA* has been experimentally confirmed as a target of σ^{ShbA} [31]. We identified p1 and p2 promoters for both *hrdB* and *shbA* homologs. In the *hrdB*p1 regions, an enriched motif, CGTAAC-N₁₆-CGATGA, was detected (Fig. 5A), which bears little resemblance to the consensus sequence of σ^{HrdB} -dependent promoters [32]. *In vitro* transcriptional assays confirmed that *S. coelicolor* *hrdB*p1 is σ^{ShbA} -dependent (Fig. 1A), and this motif is also present in the experimentally identified *S. griseus* *hrdB*p1 (Supplementary Fig. S7). For *hrdB*p2, a σ^{R} -dependent sequence, GGGAAT-N₁₈-GCTGT, was identified (Supplementary Fig. S20A) [23]. In *shbA*p1 regions, an over-represented motif, CTGCTC-N₁₆-CGCATC, was found (Fig. 5B). *In vitro* assays demonstrated that the σ^{ShbA} -RNAP holoenzyme transcribes *S. coelicolor* *shbA*p1

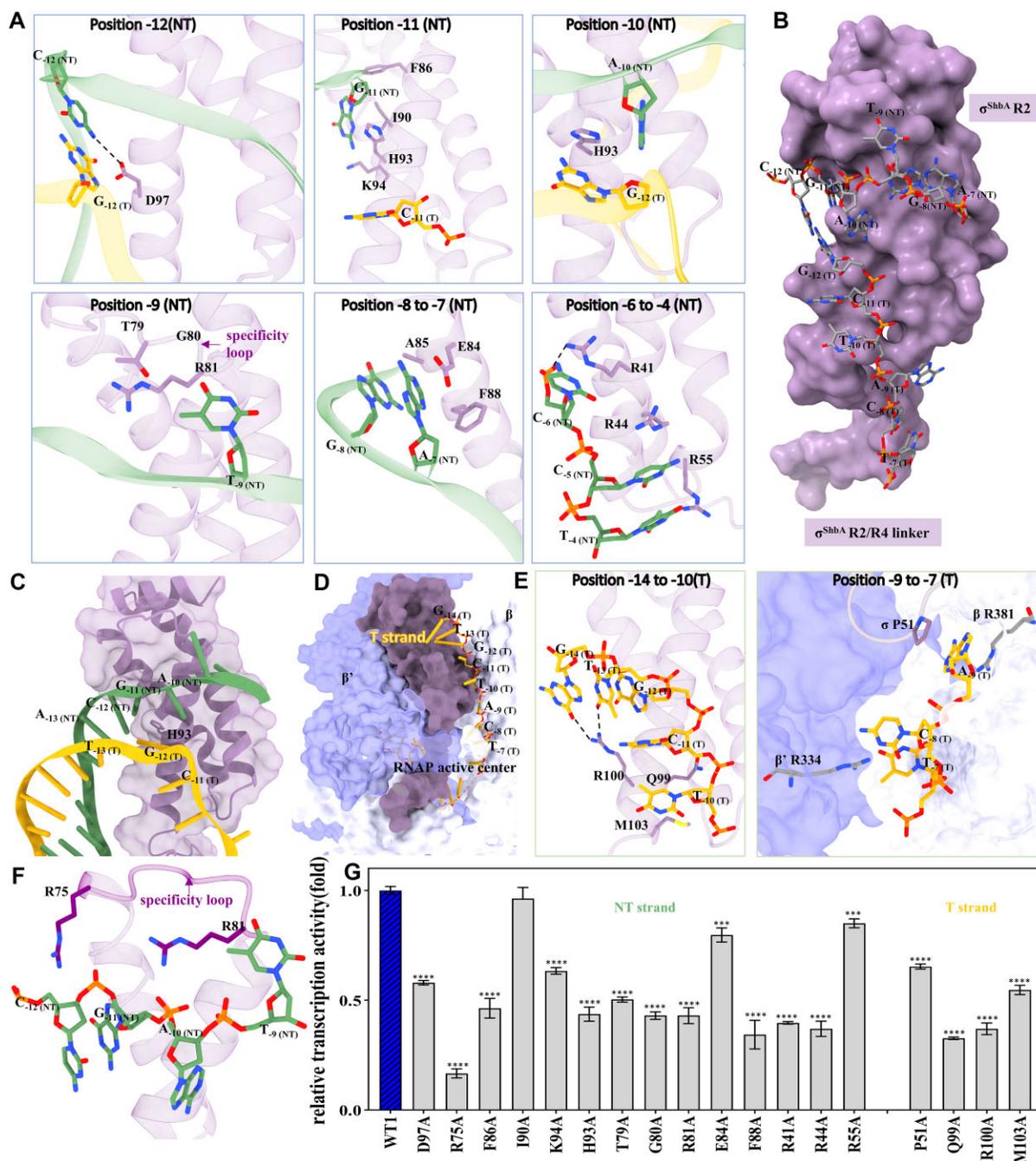


Figure 3. Structural basis of $\sigma^{\text{ShbA R2}}$ recognizing the -10 DNA element. (A) Detailed interactions of $\sigma^{\text{ShbA R2}}$ with the nontemplate strand at positions -12 to -4 . $\sigma^{\text{ShbA R2}}$ is shown in purple, nontemplate DNA (NT) in green, and template DNA (T) in yellow. Hydrogen bonds are shown as dashed lines. (B) Surface representation of $\sigma^{\text{ShbA R2}}$ -DNA interactions. (C) Ribbon representation of $\sigma^{\text{ShbA R2}}$ -DNA interactions. (D) Interactions of $\sigma^{\text{ShbA R2}}$ with template DNA in the RNAP active center cleft formed by β and β' subunits. (E) Interactions with the template strand (-14 to -7 positions). (F) The residues R75 and R81 in the σ^{ShbA} specificity loop interact with the DNA phosphate backbone. (G) *In vitro* transcriptional analysis showed that mutations of interacting residues affect σ^{ShbA} transcriptional activity. Mean fold-change values were compared using Student's *t*-test. The error bars represent the mean \pm SD of three independent experiments. **** indicates $P < .0001$, *** indicates $P < .001$. The I90A mutation displayed no significant difference, suggesting that this site may not be crucial for the activity of σ^{ShbA} .

(Supplementary Fig. S9), which contains both the -35 (CTGCTC) and -10 (CGCATC) elements. In contrast, *S. griseus* *shbAp1* retains only the conserved -10 element (CGCATC) (Supplementary Fig. S7). In *shbAp2* regions, the predicted target promoter motif contains TGCC in -35 element and CACCAT in -10 element (Supplementary Fig. S20B). Both *brdBp1* and *shbAp1* promoters share a conserved $C_{-12}G_{-11}$ motif in

their -10 elements (Fig. 5A and B). Additionally, *brdBp1* promoters feature a conserved G_{-13} in their extended -10 element, while *shbAp1* promoters have a conserved $T_{-15}G_{-14}$. The extended -10 element compensates for the absence of a conserved -35 element [58], a common feature of ECF σ factors, where the -10 element primarily determines promoter specificity [59]. Downstream of the -35 elements in *shbAp1*

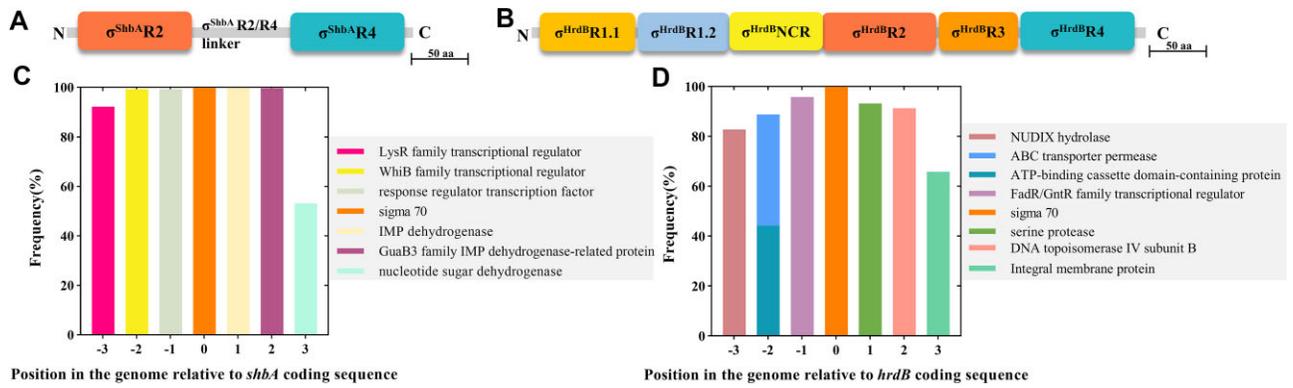


Figure 4. Genetic context of analysis of σ^{ShbA} and σ^{HrdB} . Regions of σ^{ShbA} (A) and σ^{HrdB} (B). Regions are shown in different colors. (C) Location-dependent frequency of genes encoding characteristic protein domains in the genetic background of σ^{ShbA} homologs from the ‘reference’ *Streptomyces*. (D) Location-dependent frequency of proteins with specific protein domains in the genetic background of σ^{HrdB} homologs from ‘reference’ *Streptomyces*. Only genes present in over 50% at a given position are displayed.

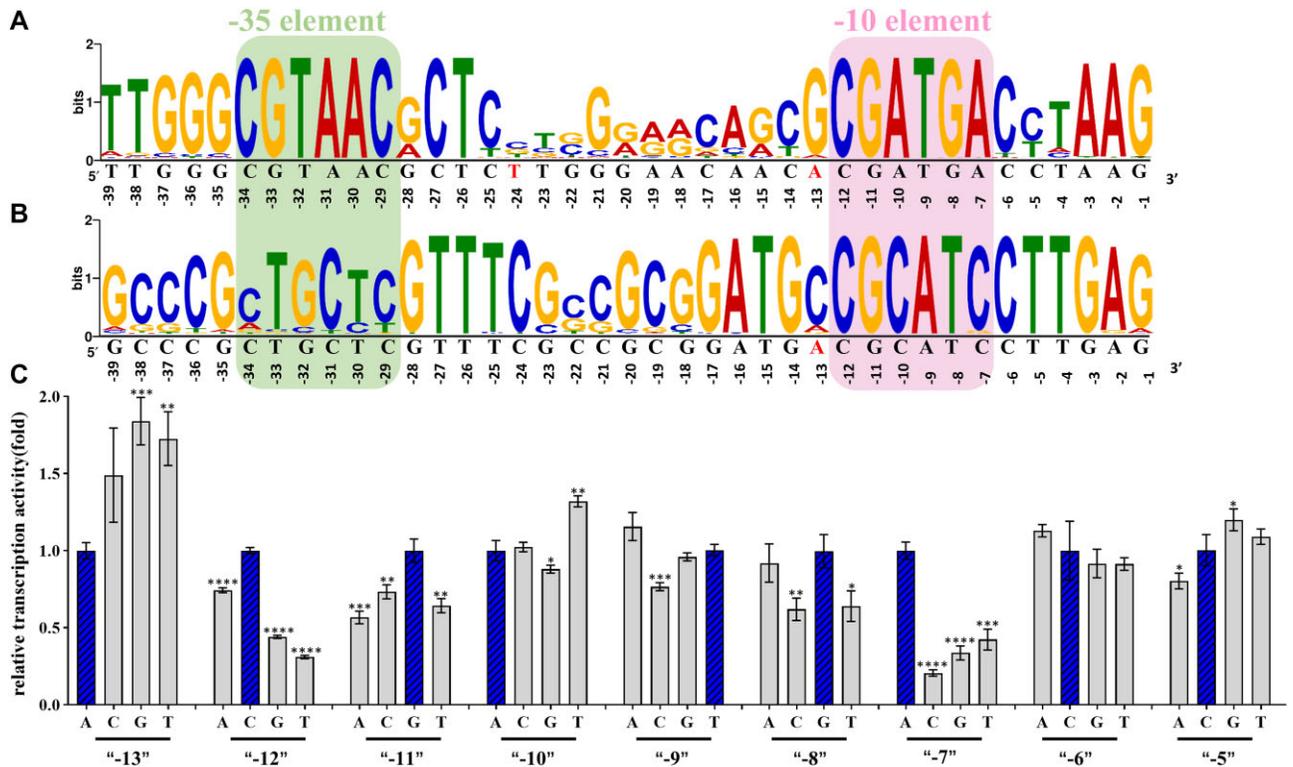


Figure 5. Identification of σ^{ShbA} -regulated promoters. (A) Promoter motifs in *hrdBp1* regions predicted by MEME in the ‘reference’ *Streptomyces*. The 300 nt upstream of the start codons from 599 *hrdB* homologs were analyzed. An enriched motif, CGTAAC₋₁₆-CGATGA (97.8%) was identified. The putative -10 and -35 elements are displayed on a green and pink background, respectively. The nucleotide sequence of *S. coelicolor hrdBp1* is shown at the bottom. All promoter sequences are in [Supplementary Table S7](#). (B) Promoter motifs in *shbAp1* regions predicted by MEME in the ‘reference’ *Streptomyces*. The 300 nt upstream of the start codons from 602 *shbA* homologs were analyzed. An enriched motif, CTGCTC₋₁₆-CGCATC (49.3%) was identified in *shbAp1* regions. The nucleotide sequence of *S. coelicolor shbAp1* is shown at the bottom. All promoter sequences are in [Supplementary Table S8](#). Position weight matrices illustrate the sequence conservation of the promoter motifs. (C) *In vitro* transcription assays validating the conservation of the promoter recognized by σ^{ShbA} (-13 to -5 region). Blue bars represent transcriptions of the wild-type *hrdBp1* promoter. Mean fold-change values were compared using Student’s *t*-test. Error bars indicate mean \pm SD of three experiments. **** represents $P < .0001$; *** represents $P < .001$; ** represents $P < .01$; * represents $P < .05$.

promoters, a conserved T-tract motif was observed (Fig. 5B), which alters DNA curvature and facilitates holoenzyme binding [60].

The structure of σ^{ShbA} -RPO details the interactions between the σ^{ShbA} and the -10 element. Bioinformatics-derived sequence preferences were validated through comprehensive

mutagenesis studies, examining all possible single-base pair mutations at positions -13 to -5 within the *S. coelicolor hrdBp1* promoter (Fig. 5C). Substitutions at positions “-12”, “-11”, and “-7” significantly reduced activity by 30%–70%, 30%–50%, and 60%–80%, respectively, confirming the critical role of these positions in sequence preference. Mutations at

“-9” and “-10” had a minor impact, while G/A was favored at position “-8.” These findings indicate that σ^{ShbA} preferentially recognizes a -10 element with the sequence C₋₁₂-G₋₁₁-N₋₁₀-N₋₉-G/A₋₈-A₋₇. Notably, *S. coelicolor* *shbAp1* features a C₋₇ instead of A₋₇ (Fig. 5B). Additionally, G is strongly preferred at position -13 in the extended -10 element, aligning with bioinformatic predictions, although *S. coelicolor* *hrdBp1* contains an A₋₁₃ (Fig. 5B and C). While previous studies have shown that G₋₆ and G₋₅ in the promoter discriminator influence transcription by the *E. coli* σ^{70} -RNAP holoenzyme [61–63], these positions have only a minor effect on σ^{ShbA} -dependent transcription of the *hrdBp1* promoter (Fig. 5C).

Discussion

In many bacteria, principal σ factors transcribe housekeeping genes, including the gene encoding the principal σ factor itself [27]. However, an alternative ECF σ^{ShbA} governs the principal σ factors in two model *Streptomyces*, *S. coelicolor* and *S. griseus* [20]. In this study, the cryo-EM structures of *S. coelicolor* σ^{ShbA} -RPO demonstrates the molecular basis of the σ^{ShbA} recognizing the *hrdB* promoter. The specific interactions of bases primarily occur at positions -12, -11, and -7 within the -10 element, and these contacts further corroborate the sequence preferences observed in bioinformatics and biochemical experiments. All *Streptomyces* reference genomes in the NCBI Datasets were analyzed. Both protein sequences and genomic neighborhoods of σ^{ShbA} and σ^{HrdB} homologs are highly conserved (Fig. 4C and D, and Supplementary Fig. S19B). Additionally, MEME identified similar over-represented sequence motifs in the promoter regions upstream of σ^{ShbA} and σ^{HrdB} genes (Fig. 5A and B). These results suggest that the principal σ^{HrdB} genes of all these *Streptomyces* are controlled by the ECF σ^{ShbA} factors.

The differences between the *hrdB* and *shbA* promoter regions likely reflect their distinct roles in *Streptomyces*. As the principal σ factor, σ^{HrdB} is required by all housekeeping genes. Therefore, the *hrdB* promoters are strong in *Streptomyces* [64]. Consistent with the -10 element of C₋₁₂-G₋₁₁-N₋₁₀-N₋₉-G/A₋₈-A₋₇ preferred by σ^{ShbA} , the *hrdBp1* promoter has an A₋₇. In contrast, σ^{ShbA} is specifically required for the transcription of *hrdB* and *shbA* genes. Previous transcriptional analysis of *S. griseus* *shbA* by S1 nuclease mapping reveals a low abundance of *shbA* transcripts [20], suggesting that the *shbA* promoters are relatively weak. According to a dynamic transcriptional landscape of *S. coelicolor*, p1 is the primary promoter of *shbA* gene [21]. This may be the reason that *shbAp1* has a C₋₇ rather than an A₋₇. *In vitro* transcription of *hrdBp1* by σ^{ShbA} -RNAP holo-enzyme is slightly more efficient than that of *shbAp1* (Supplementary Fig. S9). *In vivo*, other unknown factors may further enhance transcription of *hrdBp1*.

The deletion of σ^{ShbA} R4 has little effect on the transcriptional activity of *hrdBp1* (Supplementary Fig. S12A), based on which we hypothesize that the transcriptional activity of σ^{ShbA} mainly relies on σ^{ShbA} R2. Studies on *E. coli* σ^{70} show that the deletion of region 4.2 does not affect transcription initiation from promoters containing extended -10 element [58]. Regions 1.2–3.1 of *Thermus aquaticus* σ^A contain all the essential elements required for the core functions of σ factor to guide the RNAP holoenzyme to specific DNA sequences, to melt the double-stranded DNA around the transcription start

site, and to initiate RNA synthesis [65]. Furthermore, regions 2 and 3 of σ^{70} , along with the N terminus of the β' subunit (amino acids 1–314), melt a promoter with an extended -10 element in a reaction similar to that of an intact RNAP holoenzyme [66].

The -10 element recognition mechanism of σ^{ShbA} differs from other known σ^{70} family members (Supplementary Fig. S21). Principal σ factors use a conserved W-dyad of principal σ factors (W332/W333 of *S. coelicolor* σ^{HrdB}) to maintain ds/ss DNA junctions at the upstream edge of the -10 element (Supplementary Fig. S21A), and a conserved pocket in σ_2 to capture the flipped bases of A₋₁₁(nt) [33]. The promoter recognized by σ^{HrdB} is AT-rich at the -12 and -11 positions [32], which triggers the separation of the DNA double-strand, whereas σ^{ShbA} requires two G-C base pairs to efficiently initiate transcription. This difference in base content helps to increase the specificity of promoter recognition by σ^{ShbA} and reduces the functional overlap of promoters. In σ^{ShbA} -RPO, due to the lack of the W-dyad, the conserved H93 functions as a wedge to open the transcription bubble at position -11 (Supplementary Fig. S21B). Research on protein-DNA interactions has indicated that residues H and R have a strong preference for interacting with base G [67, 68]. This further validates the contacts between G₋₁₁ and residues H93 and R75 in our structural model. The base of T₋₉ (nt) flips out from the DNA minor groove and interact with the specificity loop (Fig. 3B and F), although substitutions at position “-9” only have slight effect on *in vitro* transcription of σ^{ShbA} -RNAP (Fig. 5C). Several RNAP structures comprising ECF σ factors are reported, demonstrating the molecular basis that σ_2 mediate -10 element melting and specifically recognize the ssDNA. In *M. tuberculosis* σ^H , N88 opens the transcription bubble at position -12, and the base of T₋₁₁ (nt) flips out from the DNA minor groove and is recognized by the specificity loop (Supplementary Fig. S21C). In *M. tuberculosis* σ^L , W68 opens the transcription bubble at the -11 position (Supplementary Fig. S21D). In *E. coli* σ^S -RPO, transcription bubbles are turned on at positions -11, and Q63 specifically recognizes A₋₁₁(nt) (Supplementary Fig. S21E). The DXXR motif serves to recognize the CG sequence within the -10 element of σ^S . However, σ^{ShbA} recognizes promoters with CG in the upstream boundary of the -10 element associated with the H₉₃R₉₄XXD₉₇ motif. In *C. thermocellum* σ^I , an arginine (R101 in SigI1) functions similarly to H93 of σ^{ShbA} , opening the transcription bubble at position -12 (Supplementary Fig. S21F). The A₋₁₂(nt) is trapped in a pocket formed by $\alpha 3$ and $\alpha 4$, while the bases from -11 to -7 are flipped and accommodated by the specificity loop and $\alpha 1(8)$. These results demonstrate the unique structural feature adopted by each ECF σ factor to recognize the corresponding promoters.

According to the recent ECF grouping, *Streptomyces* σ^{ShbA} belongs to the ECF40 group, whose members are all from the phylum *Actinobacteria*. All σ^{ShbA} homologs contain conserved σ R2 and σ R4 but do not have extended N-terminal or C-terminal structural domains that affect the activity of σ factor transfer in group ECF41 [69]. Two well-studied ECF40 members [70, 71], *Mycobacterium tuberculosis* σ^D and *Corynebacterium glutamicum* σ^D , shared 48% and 45.6% sequence identity with *S. coelicolor* σ^{ShbA} , respectively (Supplementary Fig. S22A). CHIP microarray analysis reveals that they recognize promoters with a 5'-GTAAC-N₁₇₋₁₈-CGAT-3' consensus sequence [20], which is similar to that of *Streptomyces* σ^{ShbA}

factors (Fig. 5A,B). However, multiple sequence alignments of the ECF40 σ factors reveal that the σ^{ShbA} residues (P51, T79, E84, and Q99) interacting with the promoter -10 element are the least conserved region in other ECF40 members, suggesting the specificity of σ^{ShbA} in recognizing the corresponding -10 element (Supplementary Fig. S22B). The function of *M. tuberculosis* σ^{D} is regulated by an anti-sigma factor RsdA [70]. However, an anti-sigma factor is absent in the genomic context of σ^{ShbA} . Consequently, the residues involved in σ^{D} -RsdA interactions are not conserved in σ^{ShbA} homologs (Supplementary Fig. S22A).

In summary, we provide detailed structural and bioinformatic evidence supporting the unique mechanism of the ECF σ^{ShbA} governing the housekeeping σ^{HrdB} factor in *Streptomyces*, extending the previous model that bacterial housekeeping σ factors are primarily self-regulated.

Acknowledgements

This work was supported by National Key Research and Development Program of China (2020YFA0907900), and National Natural Science Foundation of China (32370071). We thank the staff members of the Electron Microscopy System at the National Facility for Protein Science in Shanghai (NFPS), Shanghai Advanced Research Institute, Chinese Academy of Sciences, China for providing technical support and assistance in data collection. We thank Jing Liu, Xinqiu Guo, and Mengyu Yan at the Instrument Analysis Center (IAC) of Shanghai Jiao Tong University for providing assistance in data collection.

Author contributions: Conceptualization: Jianting Zheng; Data curation: Guiyang Liu, Xu Yang; Formal analysis: Guiyang Liu; Funding acquisition: Jianting Zheng; Investigation: Guiyang Liu, Xu Yang, Wenjin Yan, Yiqun Wang, Feng Yu; Methodology: Guiyang Liu; Project administration: Guiyang Liu; Software: Xu Yang, Jianting Zheng; Supervision: Jianting Zheng; Validation: Guiyang Liu; Visualization: Guiyang Liu; Writing—original draft: Guiyang Liu, Jianting Zheng; Writing—review & editing: Guiyang Liu, Jianting Zheng.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by National Key Research and Development Program of China (2020YFA0907900), and National Natural Science Foundation of China (32370071). Funding to pay the Open Access publication charges for this article was provided by National Key Research and Development Program of China (2020YFA0907900), and National Natural Science Foundation of China (32370071).

Data availability

The cryo-EM density maps and corresponding model from this study have been deposited in the Electron Microscopy

Data Bank with accession codes EMD-60839 and EMD-63701 and in the Protein Data Bank with accession codes 9ISN and 9M84. Additional structural data used for analysis (8HVR, 5ZX2, 6DV9, 6PMJ, 8I23, 7MKE, 6JBQ) are accessible from the Protein Data Bank. The protein sequences analyzed in this study can be retrieved from Uniprot under the accession code Q9L0I8 (σ^{ShbA}) [<https://www.uniprot.org/uniprotkb/Q9L0I8/entry>]. All source data are included in this paper.

References

- Liu B, Hong C, Huang RK *et al.* Structural basis of bacterial transcription activation. *Science* 2017;358:947–51. <https://doi.org/10.1126/science.aao1923>
- Zhang Y, Feng Y, Chatterjee S *et al.* Structural basis of transcription initiation. *Science* 2012;338:1076–80. <https://doi.org/10.1126/science.1227786>
- Ring BZ, Yarnell WS, Roberts JW. Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell* 1996;86:485–93. [https://doi.org/10.1016/S0092-8674\(00\)80121-X](https://doi.org/10.1016/S0092-8674(00)80121-X)
- Lonetto M, Gribskov M, Gross CA. The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* 1992;174:3843–9. <https://doi.org/10.1128/jb.174.12.3843-3849.1992>
- Mascher T. Signaling diversity and evolution of extracytoplasmic function (ECF) σ factors. *Curr Opin Microbiol* 2013;16:148–55. <https://doi.org/10.1016/j.mib.2013.02.001>
- Paget MS. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 2015;5:1245–65. <https://doi.org/10.3390/biom5031245>
- Sun D, Liu C, Zhu J *et al.* Connecting metabolic pathways: sigma factors in streptomyces spp. *Front Microbiol* 2017;8:2546. <https://doi.org/10.3389/fmicb.2017.02546>
- Li J, Zhang H, Li D *et al.* Structure of the transcription open complex of distinct sigma(I) factors. *Nat Commun* 2023;14:6455. <https://doi.org/10.1038/s41467-023-41796-4>
- Feklistov A, Darst SA. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* 2011;147:1257–69. <https://doi.org/10.1016/j.cell.2011.10.041>
- Mekler V, Pavlova O, Severinov K. Interaction of *Escherichiacoli* RNA polymerase $\sigma 70$ subunit with promoter elements in the context of free $\sigma 70$, RNA polymerase holoenzyme, and the β' - $\sigma 70$ complex. *J Biol Chem* 2011;286:270–9. <https://doi.org/10.1074/jbc.M110.174102>
- Zhang G, Campbell EA, Minakhin L *et al.* Crystal structure of thermus aquaticus core RNA polymerase at 3.3 Å resolution. *Cell* 1999;98:811–24. [https://doi.org/10.1016/S0092-8674\(00\)81515-9](https://doi.org/10.1016/S0092-8674(00)81515-9)
- Lin W, Mandal S, Degen D *et al.* Structural basis of ECF- σ -factor-dependent transcription initiation. *Nat Commun* 2019;10:710. <https://doi.org/10.1038/s41467-019-08443-3>
- Li L, Fang C, Zhuang N *et al.* Structural basis for transcription initiation by bacterial ECF σ factors. *Nat Commun* 2019;10:1153. <https://doi.org/10.1038/s41467-019-09096-y>
- Fang C, Li L, Shen L *et al.* Structures and mechanism of transcription initiation by bacterial ECF factors. *Nucleic Acids Res* 2019;47:7094–104. <https://doi.org/10.1093/nar/gkz470>
- Shen L, Lai G, You L *et al.* An Si3- σ arch stabilizes cyanobacteria transcription initiation complex. *Proc Natl Acad Sci USA* 2023;120:e2219290120. <https://doi.org/10.1073/pnas.2219290120>
- Qayyum MZ, Imashimizu M, Leanca M *et al.* Structure and function of the Si3 insertion integrated into the trigger loop/helix of cyanobacterial RNA polymerase. *Proc Natl Acad Sci USA* 2024;121:e2311480121. <https://doi.org/10.1073/pnas.2311480121>

17. Han SJ, Jiang YL, You LL *et al.* DNA looping mediates cooperative transcription activation. *Nat Struct Mol Biol* 2024;31:293–9. <https://doi.org/10.1038/s41594-023-01149-7>
18. van der Meij A, Worsley SF, Hutchings MI *et al.* Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiol Rev* 2017;41:392–416. <https://doi.org/10.1093/femsre/fux005>
19. Flårdh K, Buttner MJ. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat Rev Micro* 2009;7:36–49. <https://doi.org/10.1038/nrmicro1968>
20. Otani H, Higo A, Nanamiya H *et al.* An alternative sigma factor governs the principal sigma factor in *Streptomyces griseus*. *Mol Microbiol* 2013;87:1223–36. <https://doi.org/10.1111/mmi.12160>
21. Jeong Y, Kim JN, Kim MW *et al.* The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat Commun* 2016;7:11605. <https://doi.org/10.1038/ncomms11605>
22. Tran NT, Huang X, Hong HJ *et al.* Defining the regulon of genes controlled by $\sigma(E)$, a key regulator of the cell envelope stress response in *Streptomyces coelicolor*. *Mol Microbiol* 2019;112:461–81. <https://doi.org/10.1111/mmi.14250>
23. Kim MS, Dufour YS, Yoo JS *et al.* Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes. *Mol Microbiol* 2012;85:326–44. <https://doi.org/10.1111/j.1365-2958.2012.08115.x>
24. Bibb MJ, Molle V, Buttner MJ. sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* 2000;182:4606–16. <https://doi.org/10.1128/JB.182.16.4606-4616.2000>
25. Tu B, Mao Y, Wang R *et al.* An alternative σ factor $\sigma(L)$ (sl) regulates lincomycin production in *Streptomyces lincolnensis*. *J Basic Microbiol* 2023;63:190–9. <https://doi.org/10.1002/jobm.202200485>
26. Zhuo Y, Zhang W, Chen D *et al.* Reverse biological engineering of hrdB to enhance the production of avermectins in an industrial strain of *Streptomyces avermitilis*. *Proc Natl Acad Sci USA* 2010;107:11250–4. <https://doi.org/10.1073/pnas.1006085107>
27. Štěpánek V, Dostálová H, Busche T *et al.* Sigma regulatory network in *Rhodococcus erythropolis* CCM2595. *FEMS Microbiol Lett* 2022;369:fnac014. <https://doi.org/10.1093/femsle/fnac014>
28. Chauhan R, Ravi J, Datta P *et al.* Reconstruction and topological characterization of the sigma factor regulatory network of mycobacterium tuberculosis. *Nat Commun* 2016;7:11062. <https://doi.org/10.1038/ncomms11062>
29. Burton ZF, Gross CA, Watanabe KK *et al.* The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in *E. coli* K12. *Cell* 1983;32:335–49. [https://doi.org/10.1016/0092-8674\(83\)90453-1](https://doi.org/10.1016/0092-8674(83)90453-1)
30. Wang LF, Doi RH. Promoter switching during development and the termination site of the sigma 43 operon of *Bacillus subtilis*. *Mol Gen Genet* 1987;207:114–9. <https://doi.org/10.1007/BF00331498>
31. Otani H, Mouncey NJ. RIViT-seq enables systematic identification of regulons of transcriptional machineries. *Nat Commun* 2022;13:3502. <https://doi.org/10.1038/s41467-022-31191-w>
32. Smidova K, Zikova A, Pospisil J *et al.* DNA mapping and kinetic modeling of the HrdB regulon in *Streptomyces coelicolor*. *Nucleic Acids Res* 2019;47:621–33. <https://doi.org/10.1093/nar/gky1018>
33. Yang X, Wang Y, Liu G *et al.* Structural basis of streptomycetes transcription activation by zinc uptake regulator. *Nucleic Acids Res* 2022;50:8363–76. <https://doi.org/10.1093/nar/gkac627>
34. Wang Y, Yang X, Yu F *et al.* Structural and functional characterization of AfsR, an SARP family transcriptional activator of antibiotic biosynthesis in streptomycetes. *PLoS Biol* 2024;22:e3002528. <https://doi.org/10.1371/journal.pbio.3002528>
35. Punjani A, Rubinstein JL, Fleet DJ *et al.* cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 2017;14:290–6. <https://doi.org/10.1038/nmeth.4169>
36. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. <https://doi.org/10.1038/s41586-024-07487-w>
37. Casañal A, Lohkamp B, Emsley P. Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci* 2020;29:1055–64. <https://doi.org/10.1002/pro.3791>
38. Pettersen EF, Goddard TD, Huang CC *et al.* UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* 2021;30:70–82. <https://doi.org/10.1002/pro.3943>
39. Liebschner D, Afonine PV, Baker ML *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr D Struct Biol* 2019;75:861–77. <https://doi.org/10.1107/S2059798319011471>
40. Williams CJ, Headd JJ, Moriarty NW *et al.* MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci* 2018;27:293–315. <https://doi.org/10.1002/pro.3330>
41. Cock PJ, Antao T, Chang JT *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009;25:1422–3. <https://doi.org/10.1093/bioinformatics/btp163>
42. Durbin R, Eddy SR, Krogh A *et al.* *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge: Cambridge University Press, 1998.
43. Bailey TL, Boden M, Buske FA *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 2009;37:W202–8. <https://doi.org/10.1093/nar/gkp335>
44. Staron A, Sofia HJ, Dietrich S *et al.* The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* 2009;74:557–81. <https://doi.org/10.1111/j.1365-2958.2009.06870.x>
45. Shi W, Zhou W, Zhang B *et al.* Structural basis of bacterial σ^{28} -mediated transcription reveals roles of the RNA polymerase zinc-binding domain. *EMBO J* 2020;39:e104389. <https://doi.org/10.15252/embj.2020104389>
46. Lilic M, Holmes NA, Bush MJ *et al.* Structural basis of dual activation of cell division by the actinobacterial transcription factors WhiA and WhiB. *Proc. Natl. Acad. Sci. USA* 2023;120:e2220785120. <https://doi.org/10.1073/pnas.2220785120>
47. Lu Q, Chen T, Wang J *et al.* Structural insight into the mechanism of σ_{32} -mediated transcription initiation of bacterial RNA polymerase. *Biomolecules* 2023;13:738. <https://doi.org/10.3390/biom13050738>
48. Hubin EA, Fay A, Xu C *et al.* Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *eLife* 2017;6:e22520. <https://doi.org/doi:10.7554/eLife.22520>
49. Hubin EA, TabibSalazar A, Humphrey LJ *et al.* Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. *Proc Natl Acad Sci USA* 2015;112:7171–6. <https://doi.org/10.1073/pnas.1504942112>
50. Zhu DX, Stallings CL. Transcription regulation by CarD in mycobacteria is guided by basal promoter kinetics. *J Biol Chem* 2023;299:104724. <https://doi.org/10.1016/j.jbc.2023.104724>
51. Rammohan J, Ruiz Manzano A, Garner AL *et al.* Cooperative stabilization of mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD. *Nucleic Acids Res* 2016;44:7304–13.
52. Casas-Pastor D, Muller RR, Jaenicke S *et al.* Expansion and re-classification of the extracytoplasmic function (ECF) sigma factor family. *Nucleic Acids Res* 2021;49:986–1005. <https://doi.org/10.1093/nar/gkaa1229>
53. Stewart MYY, Bush MJ, Crack JC *et al.* Interaction of the *Streptomyces* wbl protein WhiD with the principal sigma factor sigma(HrdB) depends on the WhiD [4Fe-4S] cluster. *J Biol Chem* 2020;295:9752–65. <https://doi.org/10.1074/jbc.RA120.012708>

54. Al-Bassam MM, Bibb MJ, Bush MJ *et al.* Response regulator heterodimer formation controls a key stage in streptomyces development. *PLoS Genet* 2014;**10**:e1004554. <https://doi.org/10.1371/journal.pgen.1004554>
55. Taboada B, Estrada K, Ciria R *et al.* Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics* 2018;**34**:4118–20. <https://doi.org/10.1093/bioinformatics/bty496>
56. Persson J, Chater KF, Flårdh K. Molecular and cytological analysis of the expression of *Streptomyces* sporulation regulatory gene whiH. *FEMS Microbiol Lett* 2013;**341**:96–105. <https://doi.org/10.1111/1574-6968.12099>
57. Huang TW, Hsu CC, Yang HY *et al.* Topoisomerase IV is required for partitioning of circular chromosomes but not linear chromosomes in *Streptomyces*. *Nucleic Acids Res* 2013;**41**:10403–13. <https://doi.org/10.1093/nar/gkt757>
58. Kumar A, Malloch RA, Fujita N *et al.* The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an “extended minus 10” promoter. *J Mol Biol* 1993;**232**:406–18. <https://doi.org/10.1006/jmbi.1993.1400>
59. Swingle B, Thete D, Moll M *et al.* Characterization of the PvdS-regulated promoter motif in *Pseudomonas syringae* pv. tomato DC3000 reveals regulon members and insights regarding PvdS function in other pseudomonads. *Mol Microbiol* 2008;**68**:871–89. <https://doi.org/10.1111/j.1365-2958.2008.06209.x>
60. Gaballa A, Guariglia-Oropeza V, Dürr F *et al.* Modulation of extracytoplasmic function (ECF) sigma factor promoter selectivity by spacer region sequence. *Nucleic Acids Res* 2018;**46**:134–45. <https://doi.org/10.1093/nar/gkx953>
61. Saecker RM, Chen J, Chiu CE *et al.* Structural origins of *Escherichia coli* RNA polymerase open promoter complex stability. *Proc Natl Acad Sci USA* 2021;**118**:e2112877118. <https://doi.org/10.1073/pnas.2112877118>
62. Winkelman JT, Chandrangu P, Ross W *et al.* Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of *E. coli* ribosomal RNA promoters. *Proc Natl Acad Sci USA* 2016;**113**:E1787–95. <https://doi.org/10.1073/pnas.1522159113>
63. Haugen SP, Berkmen MB, Ross W *et al.* rRNA promoter regulation by nonoptimal binding of sigma region 1.2: An additional recognition element for RNA polymerase. *Cell* 2006;**125**:1069–82. <https://doi.org/10.1016/j.cell.2006.04.034>
64. Du D, Zhu Y, Wei J *et al.* Improvement of gougerotin and nikkomycin production by engineering their biosynthetic gene clusters. *Appl Microbiol Biotechnol* 2013;**97**:6383–96. <https://doi.org/10.1007/s00253-013-4836-7>
65. Campbell EA, Muzzin O, Chlenov M *et al.* Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* 2002;**9**:527–39. [https://doi.org/10.1016/S1097-2765\(02\)00470-7](https://doi.org/10.1016/S1097-2765(02)00470-7)
66. Young BA, Gruber TM, Gross CA. Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science* 2004;**303**:1382–4. <https://doi.org/10.1126/science.1092462>
67. Hoffman MM, Khrapov MA, Cox JC *et al.* AANT: The Amino Acid–Nucleotide Interaction Database. *Nucleic Acids Res* 2004;**32**:174D–181. <https://doi.org/10.1093/nar/gkh128>
68. Luscombe NM, Laskowski RA, Thornton JM. Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res* 2001;**29**:2860–74. <https://doi.org/10.1093/nar/29.13.2860>
69. Wecke T, Halang P, Starón A *et al.* Extracytoplasmic function σ factors of the widely distributed group ECF41 contain a fused regulatory domain. *MicrobiologyOpen* 2012;**1**:194–213. <https://doi.org/10.1002/mbo3.22>
70. Jaiswal RK, Prabha TS, Manjeera G *et al.* *Mycobacterium tuberculosis* RsdA provides a conformational rationale for selective regulation of σ -factor activity by proteolysis. *Nucleic Acids Res* 2013;**41**:3414–23. <https://doi.org/10.1093/nar/gks1468>
71. Dostálová H, Busche T, Holátko J *et al.* Overlap of promoter recognition specificity of stress response sigma factors SigD and SigH in *Corynebacterium glutamicum* ATCC 13032. *Front Microbiol* 2018;**9**:3287. <https://doi.org/10.3389/fmicb.2018.03287>