Molecular mechanism of mureidomycin biosynthesis activated by introduction of an exogenous regulatory gene *ssaA* into *Streptomyces roseosporus*

Ning Liu[1,2], Hanye Guan[1,2], Guoqing Niu[3], Lingjuan Jiang[1,4], Yue Li[1], Jihui Zhang[1], Jine Li[1] & Huarong Tan[1,2*]

[1]*State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;*
[2]*College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*Biotechnology Research Center, Southwest University, Chongqing 400715, China;*
[4]*Department of Medical Research Center, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China*

Mureidomycins (MRDs), a group of unique uridyl-peptide antibiotics, exhibit antibacterial activity against the highly refractory pathogen *Pseudomonas aeruginosa*. Our previous study showed that the cryptic MRD biosynthetic gene cluster (BGC) *mrd* in *Streptomyces roseosporus* NRRL 15998 could not be activated by its endogenous regulator 02995 but activated by an exogenous activator SsaA from sansanmycin's BGC *ssa* of *Streptomyces* sp. strain SS. Here we report the molecular mechanism for this inexplicable regulation. EMSAs and footprinting experiments revealed that SsaA could directly bind to a 14-nt palindrome sequence of 5′-CTGRCNNNNGTCAG-3′ within six promoter regions of *mrd*. Disruption of three representative target genes (*SSGG-02981*, *SSGG-02987* and *SSGG-02994*) showed that the target genes directly controlled by SsaA were essential for MRD production. The regulatory function was further investigated by replacing six regions of *SSGG-02995* with those of *ssaA*. Surprisingly, only the replacement of 343–450 nt fragment encoding the 115–150 amino acids (AA) of SsaA could activate MRD biosynthesis. Further bioinformatics analysis showed that the 115–150 AA situated between two conserved domains of SsaA. Our findings significantly demonstrate that constitutive expression of a homologous exogenous regulatory gene is an effective strategy to awaken cryptic biosynthetic pathways in *Streptomyces*.

mureidomycin, transcriptional activation, exogenous regulator—SsaA, cryptic gene cluster, *Streptomyces roseosporus*, *Pseudomonas aeruginosa*, antibiotics

INTRODUCTION

Persistent abuse of antibiotics led to the loss of antibiotic efficiency, which has caused the unceasing emergence and prevalence of muiti-/pan- resistant pathogens. However, current antibiotics are far from what is actually needed and

development of new antibiotics has become more urgent. Streptomycetes are the most important producer of industrial, agricultural and clinical applied natural antibiotics (Li et al., 2019a; Liu et al., 2019). However, owing to the lack of available methods, the numbers of newly discovered antibiotics from streptomycetes are decreasing year by year (Katz and Baltz, 2016). Sequencing of multiple *Streptomyces* genomes revealed the presence of a large number of "cryp-

*Corresponding author (email: tanhr@im.ac.cn)

tic" biosynthetic gene clusters (BGCs) for previously unrecognized products. It has therefore become necessary to devise effective methods or strategies to activate these "cryptic" biosynthetic pathways (Liu et al., 2013; Zhang et al., 2019). Usually, structural genes in antibiotic biosynthesis are controlled by cluster-situated regulators (CSRs), such as the most frequently encountered *Streptomyces* antibiotic regulatory proteins (SARPs), large ATP-binding regulators of the LuxR (LALs) and TetR family of regulators (TFRs). Typically, SARPs and LALs are activators, while TFRs are often repressors (Liu et al., 2013; Lu et al., 2017; McLean et al., 2019). The genetic manipulation of CSRs, such as overexpression of activator genes and deletion of repressor genes, has become simple and effective method to wake up cryptic BGCs for the discovery of new antibiotics in recent years. For example, the overexpression of a LAL activated a 150 kb cryptic type I PKS gene cluster from *Streptomyces ambofaciens* ATCC 23877, leading to the discovery of four 51-membered ring macrolides (Laureti et al., 2011); the deletion of *jadR** or *jadR2* in jadomycin gene cluster (*jad*) led to the expression of the cryptic *jad* gene cluster in *Streptomyces venezuelae* ISP5230 (Zhang et al., 2013). Usually, CSRs need to recognize specific *cis* elements to regulate the expression of target genes, therefore, revealing the regulatory mechanism between CSRs and their corresponding *cis* elements within cryptic BGCs could open a window to discover novel antibiotics.

Mureidomycins (MRDs) are a class of structurally unique uridyl peptide antibiotics that were initially isolated from *Streptomyces flavidovirens* (Inukai et al., 1989; Isono et al., 1989a). Other members of the uridyl peptide antibiotics also include napsamycins, pacidamycins and sansanmycins (Chen et al., 1989; Chatterjee et al., 1994; Xie et al., 2007). They share common structural scaffold: an unique 3′-deoxy-4′,5′-enamino-uridine nucleoside connected to a branching N-methyl-2,3-di aminobutyric acid (N-methyl-Daba) within a tetra- or pentapeptide backbone via an enamide bond (Niu and Tan, 2015; Niu et al., 2019). Among them, the structures of MRDs show more resemblance with those of napsamycins. The main differences are components linked to the branch of N-methyl-Daba with an *m*-tyrosine linked in MRDs and a tetrahydroisoquinoline derivative linked in napsamycins (Kaysser et al., 2011). MRDs show specific inhibitory bioactivity against *Pseudomonas aeruginosa*, which is notorious for its broad-spectrum resistance to multiple antibiotics and has therefore become a main source of nosocomial infections worldwide (Breidenstein et al., 2011; Alford et al., 2020). Furthermore, the ED_{50}s for MRD A-D with *P. aeruginosa* infected mice were 69, 75, 50 and >100 mg kg^{-1} in the *in vivo* protection test (Isono et al., 1989b). Like the other uridyl peptide antibiotics, MRDs inhibit *P. aeruginosa*'s peptidoglycan synthesis by targeting the bacterial translocase MraY, which catalyzes the transfer

of phospho-MurNAc-pentapeptide from the hydrophilic substrate uridine diphosphate-MurNAc-pentapeptide to the lipid carrier undecaprenyl phosphate in the presence of Mg^{2+} (Isono and Inukai, 1991; Mashalidis et al., 2019). Since MraY is a promising novel target for the development of new drugs, MRDs have recently emerged as promising candidates for new antibacterial agents (Bugg et al., 2006; Mashalidis et al., 2019).

Streptomyces roseosporus NRRL 15998 (Sr-WT) is a producer of the clinical important antibiotic daptomycin (Du et al., 2015) and its complete genome was sequenced by the Board Institute (UK). Mining of its genome sequences by antiSMASH online server reveals that there are at least 28 BGCs for secondary metabolite biosynthesis (Blin et al., 2019), including a BGC for putative napsamycin biosynthesis. However, uridyl peptide antibiotic napsamycin or its analogue had never been identified previously, implying that the BGC for napsamycin may remain silent. In view of this, putative CSR genes *SSGG-02978* and *SSGG-02995* were knocked out or over-expressed to activate the predicted napsamycin-BGC. Unfortunately, no expected compound was detected. Interestingly, it was reported that SsaA, a 02995 homolog from *Streptomyces* sp. strain SS, positively regulates sansanmycin's production (Li et al., 2013). Due to the sequence similarity between SsaA and 09225 (putative gene product of *SSGG-02995*), SsaA was overexpressed in *S. roseosporus* NRRL 15998 and it activated the production of a series of MRDs (Jiang et al., 2015). Hereby two interesting questions need to be addressed: why an exogenous CSR could activate the MRDs' production but its endogenous homologue failed to do so? And how the MRDs' production was activated? In this work, we report the molecular mechanism for the activation of mureidomycin biosynthesis mediated by SsaA.

RESULTS

Overexpression of *ssaA* activated the transcription of the cryptic cluster *mrd*

Our previous work reported overexpression of an exogenous gene *ssaA* in Sr-WT activated MRD production but overexpression of *ssaA*'s endogenous homologous gene *SSGG-02995* failed to do so (Jiang et al., 2015). This study aimed to unveil the mechanism of this action. The *ssaA* and *SSGG-02995* expression plasmids were constructed (Figure S1A in Supporting Information) and over-expressed in Sr-WT, then the resulting overexpression strains Sr-ssaAOE and Sr-02995OE were obtained. Production of MRDs in Sr-ssaAOE was confirmed by anti-*P. aeruginosa* bioassay, high-performance liquid chromatography (HPLC) and high resolution mass-spectrometry (HRMS) analyses. While no MRD production was found in Sr-WT and Sr-02995OE, being con-

sistent with our previous results (Figure S1B–D in Supporting Information).

In order to know why mureidomycin biosynthetic gene cluster (*mrd*) was activated by over-expression of *ssaA* in *S. roseosporus*, co-transcriptional analyses were carried out to determine the putative operons and promoter regions within the gene cluster. Previously, we reported that *mrd* consists of 28 putative open reading frames (ORFs) and its organization was outlined (Jiang et al., 2015). So, 14 pairs of PCR primers were designed to dissect the genetic organization of *mrd* (Figure 1A). The semi-quantitative reverse transcriptional PCR (RT-PCR) results showed that nine transcriptional units existed in *mrd,* they are *SSGG-02979–02978* (I), *SSGG-02980–02983* (II), *SSGG-02986–02984* (III), *SSGG-02987–02991* (IV), *SSGG-02994–02992* (V), *SSGG-02995* (VI), *SSGG-02997–02996* (VII), *SSGG-02998–03002* (VIII) and *SSGG-03004–03003* (IX) (Figure 1B). Except *SSGG-02995* (VI) transcribed independently, the rest co-transcribed units form 8 operons. Analysis of these operons revealed that there are five potential promoter regions within *mrd*, which are designated as $P_{02979–02980}$ (controlling operons I and II), $P_{02986–02987}$ (controlling operons III and IV), $P_{02994–02995}$ (controlling operons V and VI), $P_{02997–02998}$ (controlling operons VII and VIII) and P_{03004} (controlling operon IX), respectively (Figure 1A).

Based on the organization of operons in the *mrd*, 10 genes were further selected for transcriptional analyses using real-time quantitative PCR (RT-qPCR) to reflect the overall transcriptional profiles of *mrd*, including a N-acetyltransferase coding gene *SSGG-02979*, a phosphatase coding gene *SSGG-02980*, a putative regulatory gene *SSGG-02995*, a transporter coding gene *SSGG-02997*, a methyltransferase coding gene *SSGG-03004* and five non-ribosomal peptide synthetase (NRPS) coding genes *SSGG-02981*, *02986*, *02987*, *02994* and *02998*. Except *SSGG-02980* and *SSGG-02981* were from operon II, the other 8 genes were the first gene in their respective operons. The MRD production of Sr-ssaAOE was discovered from the culture broth of 3rd-day using *P. aeruginosa* as indicator strain and inhibitory zone was clearly observed. Here we can deduce that the transcription of *mrd* should be active within the 3rd and 4th-days. So, cultures from the two days of incubation time were used for transcriptional analyses. The RT-qPCR results indicated that no obvious (or very weak) transcription of these 10 genes was observed in Sr-WT (or Sr-02995OE). But in Sr-ssaAOE, the transcriptional level of *mrd* was significantly upregulated. For example, data from the 3rd day cultures showed that transcriptions of *SSGG-02979*, *02980*, *02981*, *02986*, *02987*, *02994*, *02995*, *02997*, *02998* and *03004* in Sr-ssaAOE increased 35.3-fold, 63.8-fold, 51.7-fold, 39.6-fold, 67.0-fold, 27.2-fold, 2.8-fold, 43.5-fold, 57.8-fold and 127.2-fold, compared with those in Sr-02995OE (Figure 2). It was concluded that only constitutive expression of *ssaA* could result in high-level transcription of the whole *mrd*.
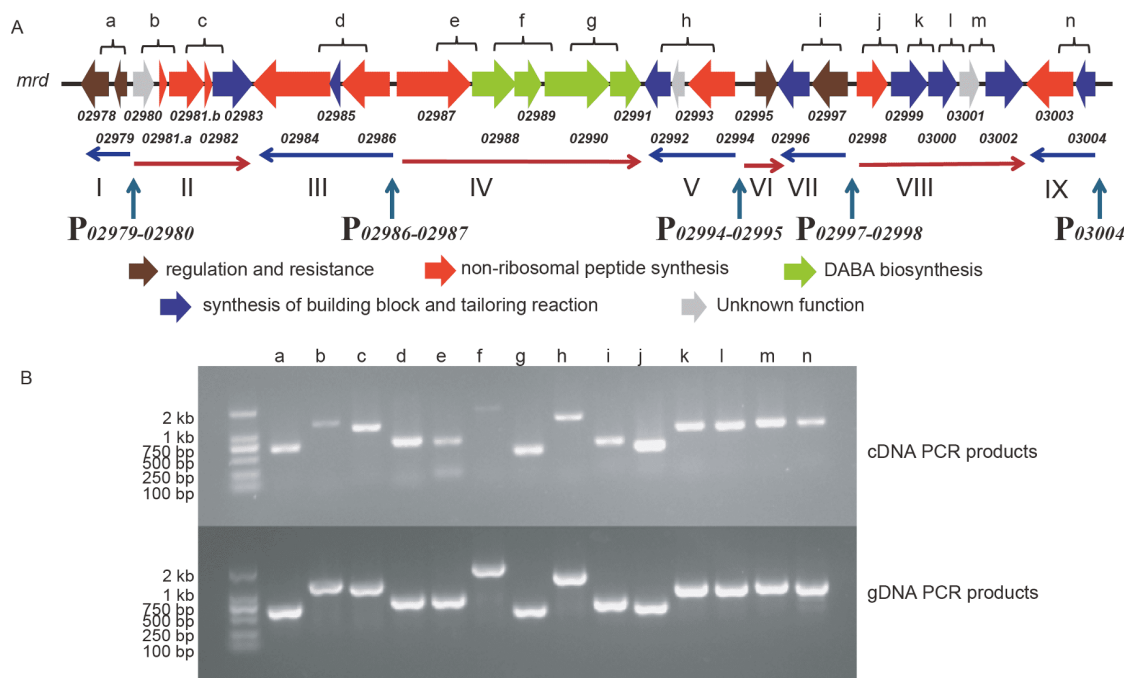

Figure 1 Co-transcriptional analysis of MRD biosynthetic gene cluster (*mrd*) by RT-PCR. A, Gene organization of *mrd*. Fourteen intergenic regions used for PCR amplification are labelled a–n; nine operons (I–IX) and their transcriptional directions are marked by horizontal arrows; deduced promoter regions of $P_{02979–02980}$, $P_{02986–02987}$, $P_{02994–02995}$, $P_{02997–02998}$ and P_{03004} are marked by vertical arrows. B, RT-PCR analysis of co-transcribed genes. Total RNAs of Sr-ssaAOE were isolated after incubation for 72 h and used for synthesizing complementary DNA (cDNA). The genome DNA (gDNA) was used as positive control.
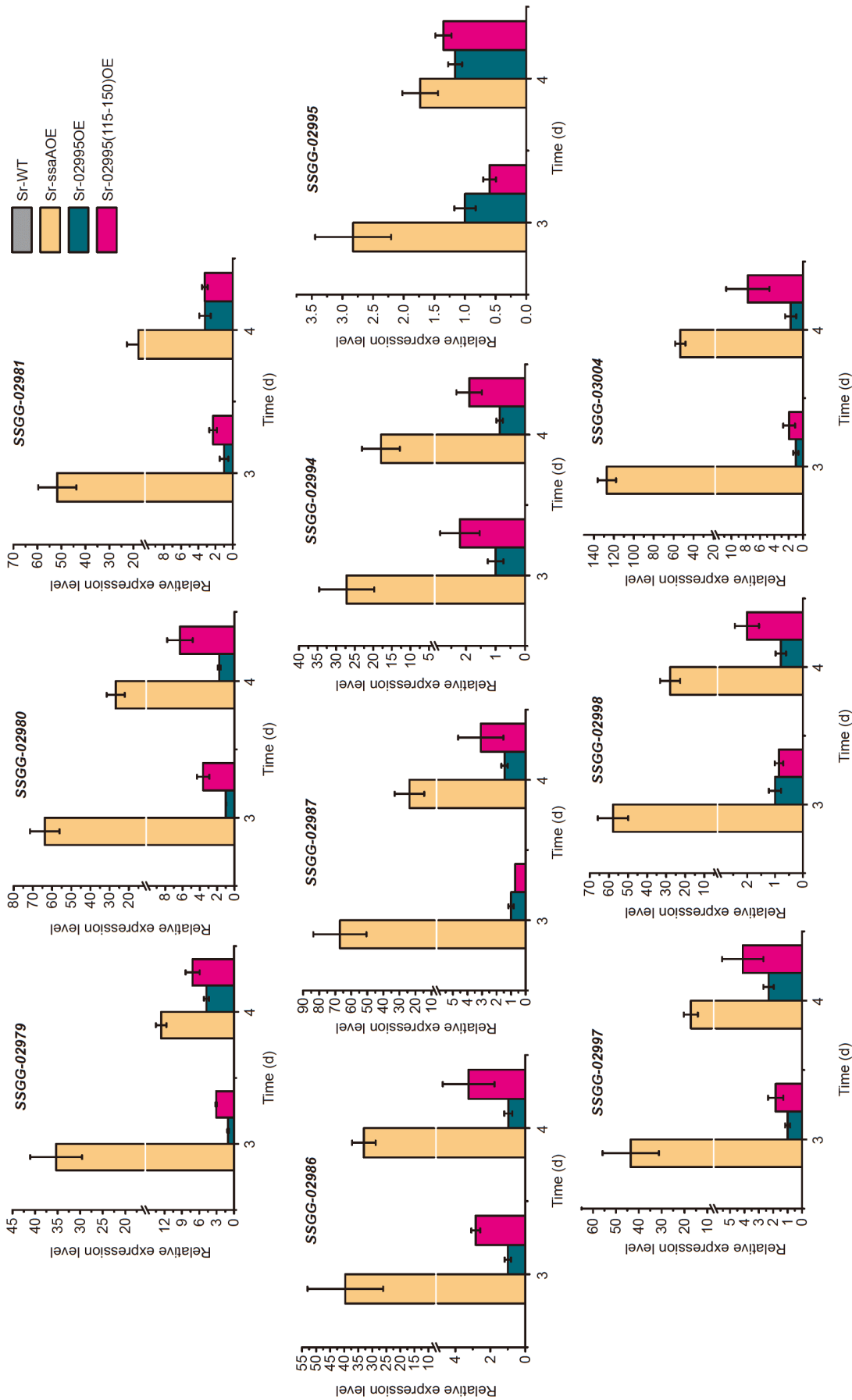
Figure 2   Transcriptional analysis of representative genes in *mrd* from strains Sr-WT, Sr-ssaAOE, Sr-02995OE and Sr-02995(115-150)OE by RT-qPCR. Total RNAs were isolated after incubation for 72 and 96 h. The transcripts of 10 genes (*SSGG-02979, 02980, 02981, 02986, 02987, 02994, 02995, 02997, 02998* and *03004*) were analyzed and relative values were obtained using the transcript of S12 (*rpsL*) as internal reference. Data are presented as means±standard deviations of three replicates. Almost no gene transcriptions were detected from Sr-WT.

SsaA controls the biosynthesis of MRD by binding to its target genes in *mrd*

It was reported that SsaA could bind to multiple promoter regions of *ssaN-P*, *ssaU-A*, *ssaC-D*, *ssaH* and *ssaW* genes in the *ssa* cluster of *Streptomyces* sp. strain SS to upregulate the expression of sansanmycin (Li et al., 2013). Therefore, we speculated that SsaA may also directly bind to the promoter regions of its cognate target genes to activate transcription of *mrd in S. roseosporus*. The main DNA binding sequence of SsaA in the *ssa* cluster is 5′-GTMCTGACANNTGTCA-GKAC-3′ (M stands for A/C, K stands for G/T, and N stands for any base in A/T/G/C). According to the reference sequence, we applied MEME Suite online server (http://meme-suite.org/tools/fimo) (Bailey et al., 2015) to search target genes of SsaA in *mrd*, and similar DNA sequences were found in the 5 promoter regions of target genes. Furthermore, a similar sequence was also found in the intergenic region of

SSGG-02980–02981 (named as P_{02981}), although the two genes are co-transcribed. The six putative promoter/intergenic regions were amplified by PCR, the recombinant protein His6-SsaA was expressed in *E. coli* and purified to homogeneity (Figure S2 in Supporting Information), then the six promoter regions were incubated with purified protein his6-SsaA under appropriate conditions, respectively. Electrophoretic mobility-shift assays (EMSAs) showed that SsaA could selectively bind to six probes to form DNA-protein complexes and differential shift bands at different protein concentrations were clearly observed (Figure 3A). These findings indicated that SsaA was most likely to activate the transcription of cryptic *mrd* by directly binding to the promoters/intergenic regions of $P_{02979–02980}$, P_{02981}, $P_{02986–02987}$, $P_{02994–02995}$, $P_{02997–02998}$ and P_{03004}.

To identify the precise binding sequences of SsaA on its target genes, DNase I footprinting experiments showed that SsaA protected nucleotide (nt) positions from −157 to
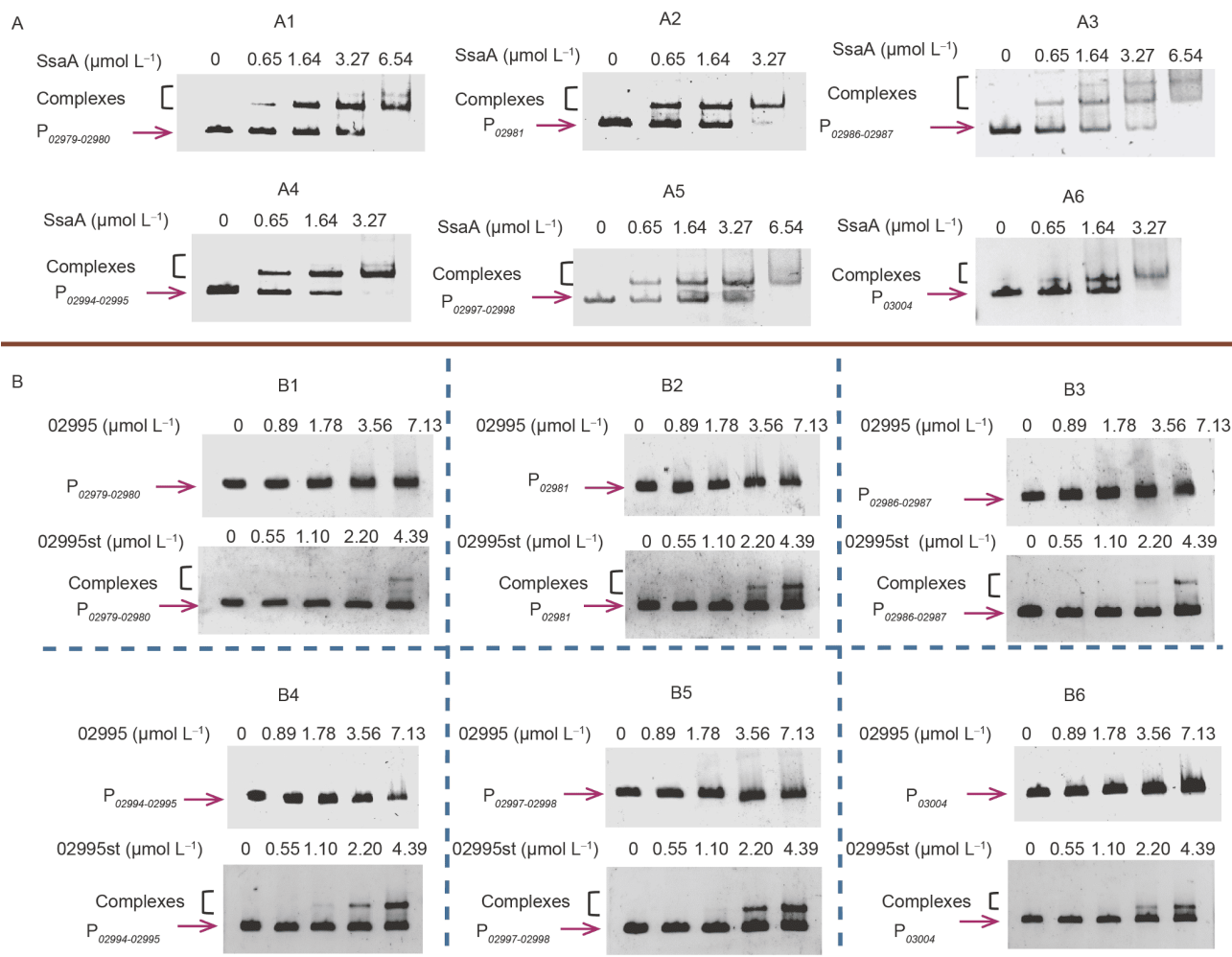

Figure 3   EMSAs of SsaA, 02995 and 02995st. A, EMSAs of different concentrations of SsaA binding to $P_{02979–02980}$ (A1), P_{02981} (A2), $P_{02986–02987}$ (A3), $P_{02994–02995}$ (A4), $P_{02997–02998}$ (A5) and P_{03004} (A6), respectively. B, Difference of binding affinity with different concentrations of proteins 02995 and 02995st to promoter regions of $P_{02979–02980}$ (B1), P_{02981} (B2), $P_{02986–02987}$ (B3), $P_{02994–02995}$ (B4), $P_{02997–02998}$ (B5) and P_{03004} (B6), respectively. Free probes are indicated by purple arrows and SsaA-DNA complexes are indicated by brackets. 02995st is a modified protein of 02995, whose 115–150 AA was substituted by the 115–150 AA of SsaA. Each lane contains 30 ng DNA probes.

−135 nt relative to the *SSGG-02980* translation start point (tsp) for P*02979–02980*, −116 to −73 nt relative to the *SSGG-02981* tsp for P*02981*, −192 to −137 nt relative to the *SSGG-02987* tsp for P*02986–02987*, −133 to −84 nt relative to the *SSGG-02994* tsp for P*02994–02995*, −224 to −192 nt relative to the *SSGG-02998* tsp for P*02997–02998* and −146 to −176 nt relative to the *SSGG-03004* tsp for P*03004* (Figure 4A–F and Figure S3 in Supporting Information). Analysis of the above SsaA protected nucleotide positions using the MEME Suite online server revealed a conservative consensus sequence 5′-YMCTGRCNNNNGTCAGKR-3′ (Y stands for C/T, M stands for A/C, R stands for A/G, K stands for T/G, N stands for any base in A/T/C/G) that is similar to conserved promoter regions in *ssa*, but some variations are as follows: (i) the consensus sequence in *mrd* is shorter than that in *ssa*, and the left and right sides are not strictly conservative complementary bases GTM...KAC, but are YM...KR. (ii) Palindrome in *mrd* is shorter. The 14 bases consensus sequence CTGACANNTGTCAG from the promoter regions of *ssa*

has a typical 6 bases palindrome structure, however, the palindrome structure in sequence CTGRCNNNNGTCAG from *mrd* is only 5 bases. In addition, the inverted repeats appeared in SsaA binding sites of P*ssaC-D* region is distinct from its consensus sequence (Li et al., 2013), and similar atypical inverted repeats also appeared in the protection region of P*02997–02998*. Compared with the inverted repeats of P*ssaC-D*, the sequence of P*02997–02998* has lower reverse repeatability (Figure 4G).

To evaluate the importance of the conserved sequence to the DNA binding affinity of SsaA, the P*02981* and P*02986–02987* were selected for sequence mutation. In P*02981*, its 14-bp palindrome sequence (5′-CTGACTCTTGTCAG-3′) was mutated to sequence (5′-TTGGCGGCGAATGG-3′) to generate a mutated probe P*02981M*. EMSAs were carried out with both P*02981* and P*02981M*. The results indicated that P*02981M* completely lost the ability to form a complex with SsaA (Figure S4A and B in Supporting Information). Moreover, another two 14-nt palindrome sequences (5′-
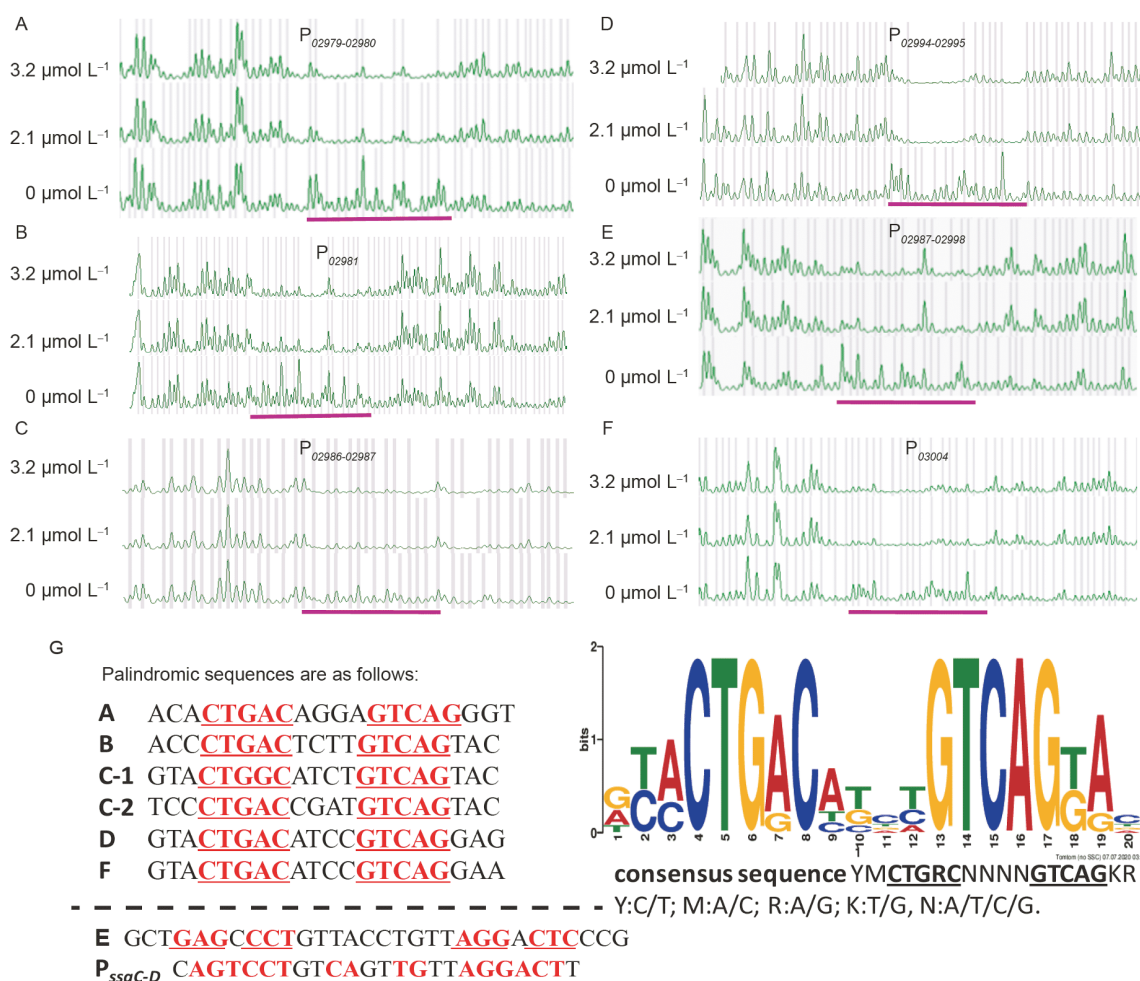

Figure 4 DNase I footprinting of SsaA binding to target genes. A–F, SsaA-protected sequence in P*02979–02980*, P*02981*, P*02986–02987*, P*02994–02995*, P*02997–02998* and P*03004*, which are marked by purple underlines. G, Conservative palindrome sequences of P*02979–02980*, P*02981*, P*02986–02987*, P*02994–02995*, P*02997–02998* and P*03004*. The consensus sequence logo was created using the MEME Suite FIMO program Version 5.1.1 (http://meme-suite.org/tools/fimo). The conservative palindrome sequence of P*02997–02998* is similar but not identical with P*ssaC-D* in *ssa*.

CTGGCATCTGTCAG-3′) and (5′-CTGACCGATGTCAG-3′) in P$_{02986-02987}$ were mutated to sequences (5′-TTGGCGGCGAATGG-3′) to generate mutants P$_{02986-02987M1}$ and P$_{02986-02987M2}$, respectively. EMSAs showed that mutated sequences (P$_{02986-02987M1}$ and P$_{02986-02987M2}$) also failed to form complexes with SsaA (Figure S4C–E in Supporting Information). The results showed that the above consensus sequences of both P$_{02981}$ and P$_{02986-02987}$ are essential for the binding of SsaA to its target genes.

To further verify the function of SsaA *in vivo*, three target genes (*SSGG-02981*, *SSGG-02987* and *SSGG-02994*) were selected for gene disruption. *SSGG-02981* is deduced to encode the condensation-thiolation (CT) domain of NRPS; *SSGG-02987* encodes the adenylation-thiolation and thioesterase (AT-TE) domain of NRPS; *SSGG-02994* is a putative NRPS adenylation (A) domain coding gene. The disruptions of *SSGG-02981*, *SSGG-02987* and *SSGG-02994* were carried out by replacing most of their DNA sequences with kanamycin resistance gene (*neo*) to generate disrupted strains Sr-ssaAOEΔ02981, Sr-ssaAOEΔ02987 and Sr-ssaAOEΔ02994, respectively. The HPLC and antimicrobial results showed that the above disrupted strains lost the ability to produce MRD and no inhibitory bioactivity against *P. aeruginosa* could be observed (Figure S5 in Supporting Information). It could be concluded that SsaA directly activates the transcription of structural genes in the cryptic *mrd*, which in turn positively controls the onset of MRD biosynthesis.

Effect of *SSGG-02995* partial sequence substitutions by *ssaA* corresponding sequences on MRD production

In the activation of the cryptic *mrd* of *S. roseosporus*, a very puzzling phenomenon was that the native regulatory gene—*SSGG-02995* was deduced to encode a protein with highly homologous to SsaA, however, its constitutive over-expression was unable to activate MRD production (Figure S1B and D in Supporting Information). In view of this, we speculated that some key regions of protein 02995 (gene product of *SSGG-02995*) might have unfavorable mutations during evolution, leading to its dysfunction. Detailed analysis showed that the different amino acids (AA) between SsaA and 02995 are mainly distributed in six regions of polypeptide chains, they are 1–28 AA (a), 36–76 AA (b), 115–150 AA (c), 186–218 AA (d), 245–253 AA (e) and 257–273 AA (f), respectively (Figure 5A). So we are curious to know which region is the main contributing factor for the functional difference between these two proteins. To this end, the corresponding DNA sequences of regions (a)–(f) (at nt positions 1–84 for (a), 109–228 for (b), 343–450 for (c), 556–654 for (d), 733–759 for (e) and 771–822 for (f)) in *SSGG-02995* were replaced by corresponding DNA sequences of *ssaA* to construct *SSGG-02995*-based modified DNA sequences (Data S1 in Supporting Information). A total

of 6 modified DNA sequences were obtained and the resulting recombinant plasmids under the control of P$_{hrdB}$ were expressed in Sr-WT to examine MRD production. After fermentation, corresponding derivative strains were analyzed by HPLC and antibacterial activity against *P. aeruginosa*. The HPLC spectra showed that only strain Sr-02995(115-150)OE could produce MRD (Figure 5B), but no MRD production was found in the other constructs. Sr-02995(115-150)OE is the modified *SSGG-02995(115-150)* (Data S1-2 in Supporting Information) overexpression strain, in which the coding sequence (343–450 nt) of pepetide region (c) in *SSGG-02995* is replaced by region c sequences of *ssaA*. The bioassay against *P. aeruginosa* also indicated that only Sr-02995(115-150)OE exhibited an inhibitory zone as that of Sr-ssaAOE (Figure 5C), which was consistent with HPLC analysis.

Molecular mechanism of MRD production mediated by SsaA

According to partial sequence substitutions, we deduced that peptide region **c** might be of significance for the bioactivity of 02995. To further understand the role of peptide region (c), the transcriptional analyses of Sr-ssaAOE, Sr-02995OE and Sr-02995(115-150)OE were conducted by RT-qPCR. The results showed that although transcriptional level of representative genes in Sr-02995(115-150)OE was still lower than that in Sr-ssaAOE, it was elevated several times compared with that of Sr-02995OE. For example, the transcriptional level increased 3.1 times for *SSGG-02979*, 2.3 times for *SSGG-02981* and 1.8 times for *SSGG-02997* on the 3rd-day; 3.6 times for *SSGG-02980*, 3.3 times for *SSGG-02986*, 2.1 times for *SSGG-02987*, 2.3 times for *SSGG-02994*, 2.5 times for *SSGG-02998* and 4.5 times for *SSGG-03004* on the 4th-day, respectively (Figure 2).

Another speculation was that the peptide region (c) (115–150 AA) might influence protein's function by altering the protein-DNA binding affinity. To validate the speculation, proteins 02995 and 02995st encoded by *SSGG-02995* and *SSGG-02995(115-150)* were over-expressed and purified for functional analyses (Figure S2 in Supporting Information). EMSAs were performed by using P$_{02979-02980}$, P$_{02981}$, P$_{02986-02987}$, P$_{02994-02995}$, P$_{02997-02998}$ and P$_{03004}$ as probes to test their binding affinity with His6-02995 and His6-02995st. Results showed that no His6-02995-DNA complex signals could be seen, implying that His6-02995 had no binding affinity with any DNA probes. On the contrary, the His6-02995-DNA complexes could be clearly observed, which exhibited that His6-02995 containing SsaA's peptide region (c) has binding ability with 6 DNA probes (Figure 3B). These results demonstrated the role of peptide region (c) and were in accordance with HPLC analysis and bioassay of Sr-02995OE and Sr-02995(115-150)OE.
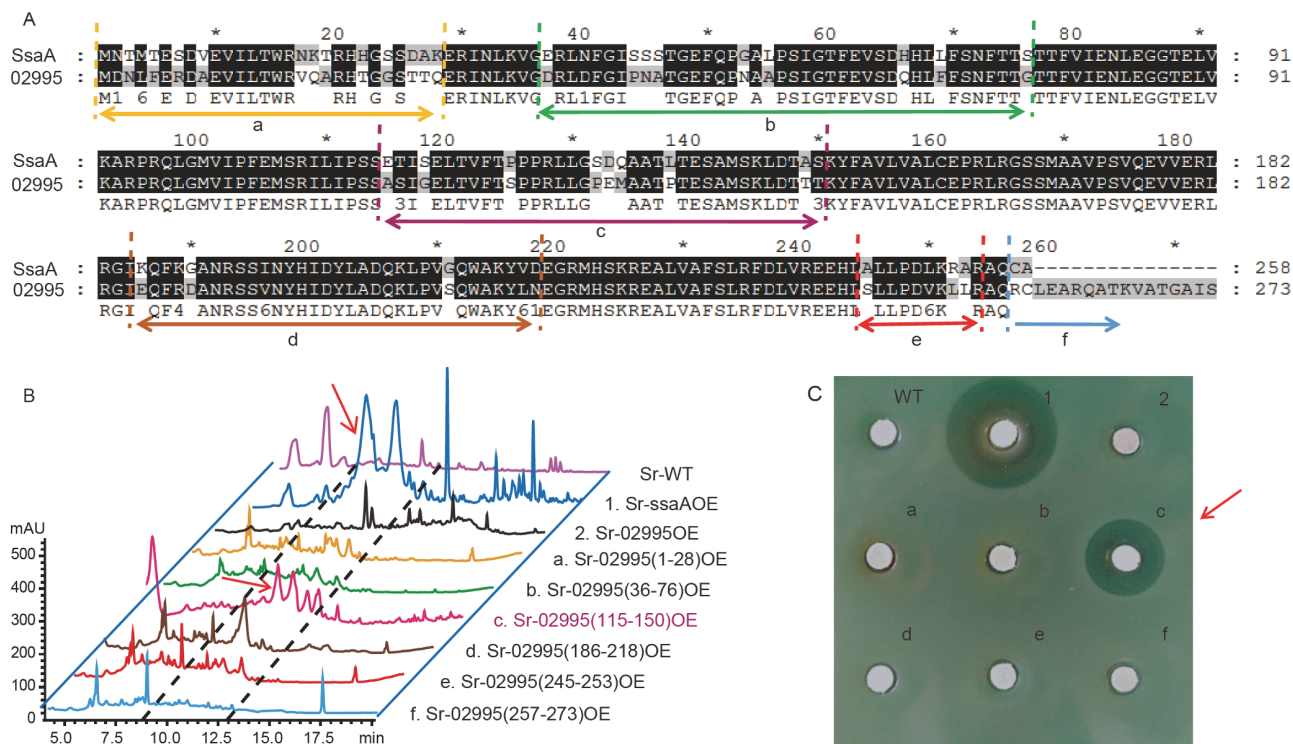
Figure 5   Effect of *SSGG-02995* partial sequence substitution by *ssaA* on MRD production. A, Six major different regions (a–f) between SsaA and 02995 are marked by dotted lines with different colors and also underlined arrows. Region a is 1–28 AA; region b is 36–76 AA; region c is 115–150 AA; region d is 186-218 AA; region e is 245-253 AA and region f is 257–273 AA. B, HPLC analyses of Sr-WT, Sr-ssaAOE, Sr-02995OE, Sr-02995(1-28)OE, Sr-02995(36-76)OE, Sr-02995(115-150)OE, Sr-02995(164-226)OE, Sr-02995(227-256)OE and Sr-02995(257-273)OE. Results showed that besides Sr-ssaAOE, Sr-02995(115-150)OE exhibited peaks at corresponding positions of MRD between two oblique dotted lines. C, Bioassays of fermentation extracts from Sr-WT, Sr-ssaAOE, Sr-02995OE and other sequence modified *SSGG-02995* overexpression strains against *P. aeruginosa*. Strain names in the picture are the same as Figure 5B. Interestingly, besides Sr-ssaAOE, only Sr-02995(115-150)OE exhibited inhibition zone in all sequence modified strains against *P. aeruginosa*, which was accordance with HPLC.

To better understand the molecular mechanism of action, we tried different conditions to obtain crystal structures of proteins SsaA, 02995 and 02995st. Unfortunately, no suitable crystals were observed. Accordingly, a molecular simulation method was performed on the three-dimensional structure of SsaA, 02995 and 02995st with Rosetta online server (http://robetta.bakerlab.org/) (Kim et al., 2004). The results showed that each of the three proteins was composed of two domains, namely the N-terminal fork head-associated (FHA) domain and the C-terminal DNA binding domain (DBD) with LuxR-type helix-turn-helix (HTH) motif, while the position of the peptide region (c) was exactly located at the junction region of the two domains (Figure 6A). In the simulated structures of SsaA and 02995st, the area of peptide region (c) forms a big opening pocket, that presumably favors the binding of protein to target DNA; whereas the pocket of peptide region (c) in 02995 is smaller compared with SsaA and 02995st, which may be unfavorable for its binding to target DNA (Figure 6B). The different simulated three-dimensional structure of peptide region (c) is consistent with the different functions of SsaA, 02995 and 02995st in activating the cryptic *mrd*, which may explain the

molecular mechanism why overexpression of *ssaA* can activate MRD production but overexpression of *SSGG-02995* fails to do so.

DISCUSSION

The epidemics of Ebola, coronavirus (COVID-19) and other highly pathogenic pathogens have brought severe challenges to the clinical drug application worldwide and call for more effective medicines (Khan et al., 2018; Jin et al., 2020). Genome sequencing and analyses reveal that streptomycetes, as the most important sources of antibiotics for clinical use, have been greatly underestimated for the potential of new drug leads owing to the presence of a large number of unexpressed BGCs for new natural products. Therefore, establishing new methods to activate these "cryptic" BGCs have become necessary (Nguyen et al., 2020). As we know, biosynthesis of secondary metabolites in *Streptomyces* is regulated stringently, and the regulatory genes, especially the CSRs, often function as the "molecular switches" for gene expression in the gene clusters (Yin et al., 2019). For ex-
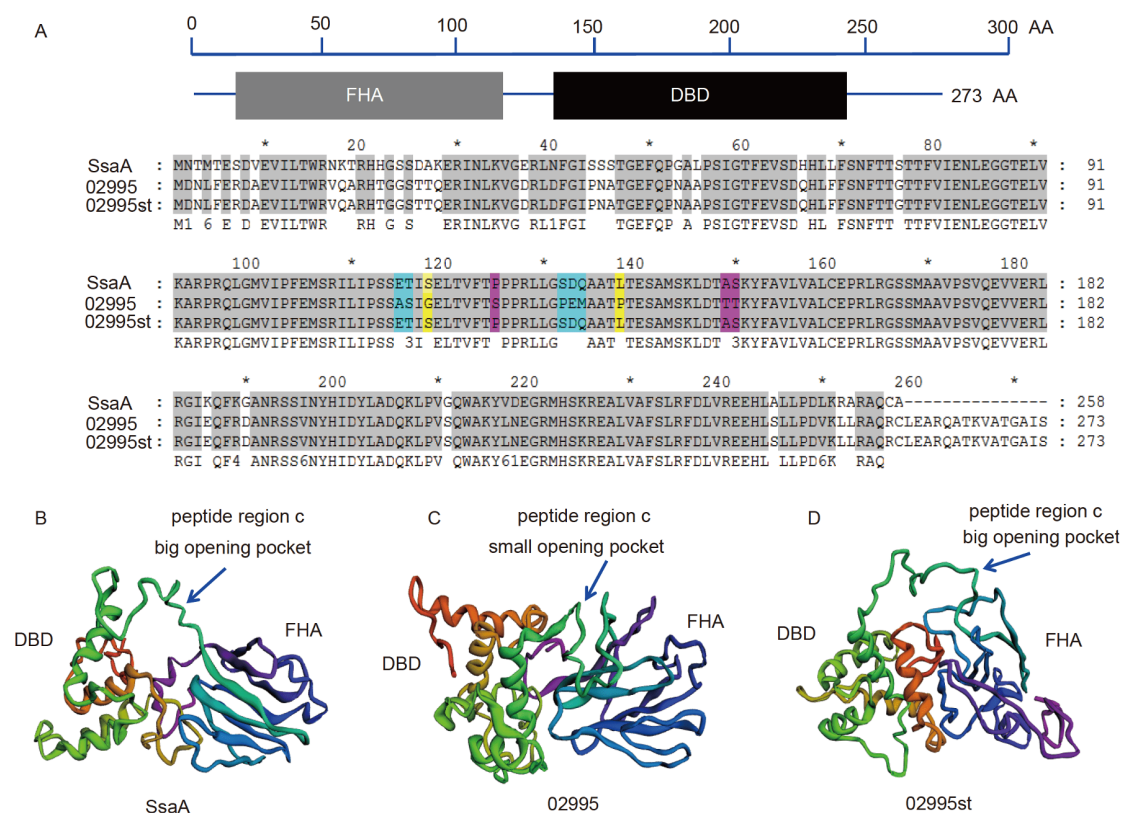
Figure 6 Sequence alignment of amino acids and putative structures of SsaA, 02995 and 02995st. A, Sequence alignment and predicted domains of SsaA, 02995 and 02995st. Differences within 115–150 AA are marked by different colors. B–D, Putative structure of SsaA, 02995 and 02995st, respectively. Proteins are proposed to have two major structure domains: N-terminal FHA domain and C-terminal LuxR-type HTH DBD.

ample, more than 20 LAL family CSRs in *Streptomyces* have been identified to be essential for antibiotic production, including AveR in avermectin biosynthesis of *S. avermitilis*, MilR in milbemycin biosynthesis of *S. bingchenggensis* and SlnR in salinomycin biosynthesis of *S. albus* (Liu et al., 2013; Guan et al., 2019; Li et al., 2019b). Thus, the genetic manipulation of CSRs has been an effective strategy to activate these "cryptic" BGCs. Our previous study noticed an interesting phenomenon that the endogenous regulatory gene *SSGG-02995* in the cryptic *mrd* of *S. roseosporus* could not activate the transcription of *mrd*, whereas overexpression of its homologous exogenous regulatory gene—*ssaA* from *Streptomyces* sp. strain SS could effectively activate the transcription of *mrd*. In this study, we demonstrate that SsaA encoded by *ssaA* acts as an activator for MRD biosynthesis and plays positive regulatory role by directly binding to the promoter regions of target genes in *mrd*. Moreover, the binding sites of SsaA in *mrd* are found to be less conservative compared with its originally reported binding sites in *ssa*, which may become the molecular basis of its interspecies regulation.

In searches of NCBI databases with BLASTP, SsaA orthologues are revealed to exist in multiple actinomycetes. Moreover, all the known orthologues are situated in biosynthetic gene clusters for uridyl peptide antibiotics, such as

PacA from the pacidamycin biosynthetic gene cluster pac, NpsM from the napsamycin biosynthetic gene cluster nap (Zhang et al., 2010; Kaysser et al., 2011), implying that this type of regulatory protein is likely to be common in uridyl peptide biosynthetic gene clusters of *Streptomyces* species. Therefore, it is possible to predict potential new uridyl peptides by simply retrieving the presence of SsaA orthologues from genome data of *Streptomyces* species. For example, *Streptomyces* sp. Tue6075, *Streptomyces* sp. BK038, *Streptomyces* sp. HCCB10043 and *Streptomyces* sp. WM6372 that have SsaA orthologues WP075265578.1, WP132755442.1, ESU49223.1 and KOU21872.1 also show potent BGCs of uridyl peptides according to genome scanning by antiSMASH. Interestingly, in the genome of *Nonomuraea* sp. WAC 01424, there are multiple NRPS-encoding genes and a pseudouracil synthase gene RSN09173 near the coding sequence of SsaA orthologue WP_125635936.1, suggesting that this species may have the potential to encode atypical uridyl peptides. However, no uridyl peptides or other nucleoside peptides produced by the above-mentioned actinomycetes have been reported yet, implying that these clusters are most likely to be silent under laboratory conditions. Analyses of the putative promoter regions in related gene clusters of these actinomycetes could also find similar 14-nt palindrome sequences (data not shown), which may be

potential SsaA binding sites, indicating that the over-expression of *ssaA* may be appropriate to activate related gene clusters to unleash the biosynthetic potential in these actinomycetes for valuable natural products.

High-level transcription of genes situated in BGC is an important prerequisite for the biosynthesis of secondary metabolites. Our results showed that genes in *mrd* did not transcribe at all in the fermentation cultures of Sr-WT, in-dicating the cryptic feature of this gene cluster. Whereas in Sr-ssaAOE, a *ssaA* over-expressed strain, the relevant genes in *mrd* were strongly upregulated on both the 3rd and 4th-days, and a large amount of MRD produced accordingly. This result indicated that SsaA could strongly activate the cryptic *mrd*, which was consistent with the results that SsaA showed binding affinity with target genes in *mrd in vitro*. Surprisingly, in Sr-02995OE, an *SSGG-02995* over-ex-pressed strain, the relevant genes in *mrd* did show weak transcription in fermentation cultures compared with that in Sr-ssaAOE, but the expected MRD could not be detected, implying that the transcriptional level of genes in *mrd* could not meet the demand for MRD biosynthesis. This was con-sistent with the phenomenon that protein 02995 could not effectively form complexes with promoter regions of *mrd in vitro*. Finally, in Sr-02995(115-150)OE, a modified gene *SSGG-02995(115-150)* over-expressed strain, the transcrip-tional levels of 10 representative genes in *mrd* increased to 1.8–4.5 times compared with those in Sr-02995OE on the 3rd or 4th-days' cultures. For example, the transcripts of *SSGG-02986* and *SSGG-02998* encoding NRPS increased to 3.3 times and 2.5 times on the 4th-day, respectively. Although MRD yield of Sr-02995(115-150)OE was lower than that of Sr-ssaAOE, MRD production could be detected definitely. The above mentioned results for Sr-02995(115-150)OE could also be supported by the binding affinity of protein 02995st to its target genes.

Many CSRs, such as TFRs and LALs, contain HTH do-mains for DNA-binding and ligand-binding domains (LBDs) for recognition of ligand molecules such as signaling mole-cules, antibiotics and oligopeptides, and the two domains are connected by random amino acids. Functions of HTH and LBD are reported, for example, SimR, a typical TFR, con-tains an N-terminal HTH according to its crystal structure. When SimR binds to target DNA sequence, its N-terminal extension binds in the minor groove, and SimR contacts some specific sequences in the major groove via its HTH motif (Le et al., 2011). However, less attention has been paid to the junction regions of different domains yet. Based on our above results, we speculate that the different three-dimen-sional structures of the junction regions (peptide region (c)) caused by the diversity in amino acid charges and polarities between SsaA and 02995 have an important contribution to their bioactivity, which result in that one could bind to their target and another can't do so. One possible explanation for

the mechanism is that a larger pocket between domains FHA and DBD of SsaA is beneficial for forming a "clamp" structure to contact target DNA, yet the corresponding pocket of 02995 is not big enough to do this. Another con-ceivable explanation is that many regulatory proteins such as CprB in *Streptomyces coelicolor* or TylP in *Streptomyces fradiae* need to form polymers to bind to their *cis*-elements (Bhukya et al., 2017; Ray et al., 2017). However, the evo-lutionary mutations in the peptide region (c) of 02995 may impede the polymerization process, so it lost its function. Our studies may provide an important new paradigm for understanding structure-function relationship to other reg-ulatory proteins.

MATERIALS AND METHODS

Strains, plasmids and growth conditions

Bacteria strains, plasmids and primers used in this study are listed in Tables 1, 2 and Table S1 in Supporting Information, respectively. *Streptomyces roseosporus* NRRL 15998 was used for activation and manipulation of the cryptic mur-eidomycin biosynthetic gene cluster. *Escherichia coli* JM109 was used as a general host for propagating plasmids. *E. coli* ET12567 (pUZ8002) was used for transferring plasmids into *Streptomyces* via intergeneric conjugation (Hopwood et al., 2000). *E. coli* C41 (DE3) and Rosetta (DE3) were used for recombinant protein expression. *Pseudomonas aeruginosa* O1 and 14 were used as indicator strains for bioassay of MRDs.

To harvest spores, *S. roseosporus* NRRL 15998 and its derivatives were grown on MM (0.5‰ L-Asn, 0.5‰ K$_2$HPO$_4$, 0.2‰ MgSO$_4$·7H$_2$O, 0.01‰ FeSO$_4$·7H$_2$O, 5% mannitol, 1% agar) at 28°C for 4 days. *S. roseosporus* NRRL 15998 was grown in AS-1 (0.1% yeast extract, 0.5% soluble starch, 0.25% NaCl, 1% Na$_2$SO$_4$, 0.02%, L-Ala, 0.02% L-Arg, 0.05% L-Asn, 2% agar) for intergeneric conjugation transfer experiment, and incubated in TSB (1.7% pancreatic digest of casein, 1.5% papaic digest of soybean, dextrose 0.25%, 0.5% NaCl) as seed culture; further fermentation for MRD production and RNA isolation was carried out on GYM (0.4% glucose, 0.4% yeast extract, 1% malt extract). *E. coli* strains were grown in LB (1% tryptone, 0.5% yeast extract, 1% NaCl, 1% agar) at 37°C for 16–20 h. When ne-cessary, antibiotics were added into MM or AS-1 medium with final concentration as follows: 100 µg mL^{-1} apramycin, 100 µg mL^{-1} kanamycin, 25 µg mL^{-1} chloramphenicol, 100 µg mL^{-1} hygromycin in LB medium for *E. coli*; 10 µg mL^{-1} apramycin, 25 µg mL^{-1} nalidixic acid in AS-1 medium for *S. roseosporus*.

Gene disruption

To disrupt *SSGG-02981*, the upstream and downstream

Table 1   Strains used in this study

Strains	Description/Genotype	Source/Ref.
Streptomyces roseosporus NRRL 15998	Wild type strain (Sr-WT)	Board Institute
Sr-ssaAOE	Sr-WT/pSET152::P$_{hrdB}$-*ssaA*	This work
Sr-02995OE	Sr-WT/pSET152::P$_{hrdB}$-*02995*	This work
Sr-02995(1-28)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(1-28)*	This work
Sr-02995(36-76)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(36-76)*	This work
Sr-02995(115-150)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(115-150)*	This work
Sr-02995(186-218)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(186-218)*	This work
Sr-02995(245-253)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(245-253)*	This work
Sr-02995(257-273)OE	Sr-WT/pSET152::P$_{hrdB}$-*02995(257-273)*	This work
Sr-ssaAOEΔ02981	*SSGG-02981* disruption mutant of Sr-ssaAOE	This work
Sr-ssaAOEΔ02987	*SSGG-02987* disruption mutant of Sr-ssaAOE	This work
Sr-ssaAOEΔ02994	*SSGG-02994* disruption mutant of Sr-ssaAOE	This work
Escherichia coli (*E. coli*)		
JM109	F', *proA$^+$B$^+$*, *lacI^q*, Δ*(lacZ)M15*/Δ*(lac-proAB)*, *gyrA96*, *recA1*, *relA1*, *endA1*, *hsdR17*	Invitrogen
C41(DE3)	F$^-$, *ompT*, *gal dcm hsd S$_B$* (r$_B^-$, m$_B^-$) (DE3)	Lucigen
Rosetta(DE3)	F$^-$, *ompT*, *gal dcm hsd S$_B$* (r$_B^-$, m$_B^-$) (DE3), pRARE(*arg*U, *arg*W, *ilex*, *gly*T, *leu*W, *pro*L) (Cam^R)	TSINGKE
SsaA-28a-C41	C41(DE3) containing pET28a::*ssaA*	This work
02995-28a-RS	Rosetta(DE3) containing pET28a::*02995*	This work
02995st-28a-RS	Rosetta(DE3) containing pET28a::*02995(115-150)*	This work
ET12567/pUZ8002	*dam$^-$ dcm$^-$ hsdM$^-$* pUZ8002	(Paget et al., 1999)
PAO1	Bioactivity indicator strain (*Pseudomonas aeruginosa*)	(Jiang et al., 2015)
PA14	Bioactivity indicator strain (*P. aeruginosa*)	(Jiang et al., 2015)

fragments of *SSGG-02981* were amplified by PCR with primer pairs 981upF/R and 981dnF/R, and a 1.0 kb *neo* fragment was amplified with primer pair kan-GF/GR (Pan et al., 2009). The three fragments were then ligated into the EcoRV-digested pKC1139 with Gibson assembly method to generate plasmid pKC1139-02981D. To disrupt *SSGG-02987* and *SSGG-02994,* similar strategies were used to generate plasmids pKC1139-02987D and pKC1139-02994D, except that different primer pairs were used (Table S1 in Supporting Information). All recombinant plasmids were verified by PCR using their corresponding primer pairs 981crsF/R, 987crsF/R and 994crsF/R, respectively. The resulting plasmids were first introduced into the *E. coli* ET12567 (pUZ8002) and subsequently transferred into Sr-WT via *E. coli-Streptomyces* intergeneric conjugation. *S. roseosporus* exconjugants were incubated on AS-1 petri dishes with kanamycin at 40°C for 4 days and at 28°C for additional 4 days. The kanamycin-resistant and apramycin-sensitive transformants were selected as the double-cross-over disrupted strains and named as Sr-Δ02981, Sr-Δ02987 and Sr-Δ02994. All resulting recombinant strains were confirmed by PCR. Finally, pSET152::P$_{hrdB}$-*ssaA* was introduced into Sr-Δ02981, Sr-Δ02987 and Sr-Δ02994 via

conjugal transfer to obtain Sr-ssaAOEΔ02981, Sr-ssaAOEΔ02987 and Sr-ssaAOEΔ02994, respectively.

Expression of *ssaA, SSGG-02995, SSGG-02995(115-150)* and other modified genes

To over-express *ssaA* (or *SSGG-02995*) in Sr-WT, the ORF of *ssaA* (or *SSGG-02995*) was amplified using primer pair ssaGF/R (or 995GF/R), and the constitutive promoter P$_{hrdB}$ from *S. coelicolor* was amplified using primer pair hrdGF/R. The two fragments were then inserted into the EcoRV-digested pSET152 with Gibson assembly method to generate pSET152::P$_{hrdB}$-*ssaA* (or pSET152::P$_{hrdB}$-*02995*).

SSGG-02995(115-150) was a modified gene of *SSGG-02995*, its 343 to 450 nt encoding 115–150 AA was substituted by the 343–450 nt of *ssaA*. To obtain this modified gene, plasmid pIJ778::P$_{hrdB}$-*02995* was firstly constructed as follows: P$_{hrdB}$-*02995* (1.25 kb) was amplified by PCR using pSET152::P$_{hrdB}$-*02995* as template with primer pair hrdGF778/995GR778, meanwhile pIJ778 (2.5 kb) was amplified by primer pair 778F/R. The two fragments as mentioned above were ligated together with Gibson assembly method to obtain a 3.75 kb plasmid pIJ778::P$_{hrdB}$-*02995*.

Table 2 Plasmids used in this study

Plasmids	Description	Source/Ref.
pSET152	$aac(3)IV$, $lacZ$, rep^{pMB1*} $att\phi C31$, $oriT$	(Hopwood et al., 2000)
pKC1139	$aac(3)IV$, *E. coli-Streptomyces* shuttle plasmid contains a *Streptomyces* temperature-sensitive origin of replication pSG5	(Hopwood et al., 2000)
pIJ778	Spec^R, Strep^R; $aadA$ from Ω-fragment, $oriT$	(Gust et al., 2003)
pET28a(+)	Kan^R, expression vector	Novagen
pSET152::P$_{hrdB}$-ssaA	pSET152 containing P$_{hrdB}$ of *S. coelicolor* A3(2) and *ssaA*, used for overexpression of *ssaA*	(Jiang et al., 2015)
pSET152::P$_{hrdB}$-02995	pSET152 containing P$_{hrdB}$ and *SSGG-02995*, used for over-expression of *SSGG-02995*	(Jiang et al., 2015)
pIJ778::P$_{hrdB}$-ssaA	pIJ778 containing sequence P$_{hrdB}$-ssaA	This work
pIJ778::P$_{hrdB}$-02995	pIJ778 containing sequence P$_{hrdB}$-02995	This work
pSET152::P$_{hrdB}$-02995(1-28)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(1-28)*	This work
pSET152::P$_{hrdB}$-02995(36-76)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(36-76)*	This work
pSET152::P$_{hrdB}$-02995(115-150)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(115-150)*	This work
pSET152::P$_{hrdB}$-02995(186-218)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(186-218)*	This work
pSET152::P$_{hrdB}$-02995(245-253)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(245-253)*	This work
pSET152::P$_{hrdB}$-02995(257-273)	pSET152 containing P$_{hrdB}$ and sequence of *SSGG-02995(257-273)*	This work
pKC1139-02981D	pKC1139 derivative, used for disruption of *SSGG-02981* with replacement of kanamycin resistance gene *neo*	This work
pKC1139-02987D	pKC1139 derivative, used for disruption of *SSGG-02987* with replacement of *neo*	This work
pKC1139-02994D	pKC1139 derivative, used for disruption of *SSGG-02994* with replacement of *neo*	This work
pET28a::ssaA	pET28a containing *ssaA* coding sequence	This work
pET28a::02995	pET28a containing *SSGG-02995* coding sequence	This work
pET28a::02995(115-150)	pET28a containing *SSGG-02995(115-150)* coding sequence	This work
pSET152::P$_{79-80}$	pSET152 containing full sequence of $P_{02979-02980}$	This work
pSET152::P$_{81}$	pSET152 containing full sequence of P_{02981}	This work
pSET152::P$_{86-87}$	pSET152 containing full sequence of $P_{02986-02987}$	This work
pSET152::P$_{94}$	pSET152 containing full sequence of P_{02994}	This work
pSET152::P$_{97-98}$	pSET152 containing full sequence of $P_{02997-02998}$	This work
pSET152::P$_{04}$	pSET152 containing full sequence of P_{03004}	This work

Subsequently, a 3.65 kb fragment, in which the 343–450 nt of *SSGG-02995* was deleted from pIJ778::P$_{hrdB}$-02995, was amplified by PCR using pIJ778::P$_{hrdB}$-02995 as template with primer pair 995F3/R3, and a 108 bp fragment for 343–450 nt of *ssaA* was amplified by PCR with primer pair ssaF3/R3. Then, the above two fragments of 3.65 kb and 108 bp were ligated together with Gibson assembly method to obtain pIJ778::P$_{hrdB}$-02995(115-150). Finally, P$_{hrdB}$-02995 (115-150) was amplified by PCR using primer pair hrdGF/ 995GR and subcloned into the EcoRV-digested pSET152 with Gibson assembly method to generate pSET152::P$_{hrdB}$-02995(115-150). Similar strategy was used to construct other plasmids in which some nucleotide sequences of *SSGG-02995* at positions 1–84 nt (1–28 AA), 109–228 nt (36–76 AA), 556–654 nt (186–218 AA), 733–759 nt (245–253 AA) and 771–822 nt (257–273 AA) were substituted by corresponding sequences of *ssaA*, and the resulting plasmids designated as pSET152::P$_{hrdB}$-02995(1-28), pSET152::P$_{hrdB}$- 02995(36-76), pSET152::P$_{hrdB}$-02995(186-218), pSET152:: P$_{hrdB}$-02995(245-253) and pSET152::P$_{hrdB}$-02995(257-273), respectively.

After construction, all the pSET152-based recombinant plasmids were introduced into strain Sr-WT via the conjugal transfer to obtain corresponding mutant strains (Table 1). Mutant strains containing P$_{hrdB}$-ssaA, P$_{hrdB}$-02995, P$_{hrdB}$- 02995(115-150) and other modified genes were further confirmed by PCR, and the primer pairs used were listed in Table S1 in Supporting Information.

Expression and purification of SsaA, 02995 and 02995st

SsaA, 02995 and 02995st coding regions were amplified by PCR with prime pairs ssaAPF/R, 995PF/R and 995stPF/R (Table S1 in Supporting Information) using fragments P$_{hrdB}$- ssaA, P$_{hrdB}$-02995 and P$_{hrdB}$-02995(115-150) as templates, respectively. After digestion of PCR products with EcoRI

and HindIII, these fragments were inserted into the same sites of pET-28a(+) to give plasmid pET28a::*ssaA*, pET28a::*02995* and pET28a::*02995(115-150)*, and then they were transformed into *E. coli* C41(DE3) and/or Rosetta (DE3) to overexpress His$_6$-SsaA, His$_6$-02995 and His$_6$-02995st, respectively. The optimum expression temperature is 24°C for SsaA and 16°C for the other two proteins. Protein purification was performed with a nickel-nitrilotriacetic acid-agarose column as described previously (Zou et al., 2014).

EMSAs and DNase I footprinting

EMSAs were performed as described previously (Li et al., 2016; Wang et al., 2018). Six promoters used as probes for EMSAs were amplified by PCR using genomic DNA of Sr-WT as template with primer pairs listed in Table S1 in Supporting Information. Three sequence modified DNA probes P$_{02981M}$, P$_{02986–02987M1}$ and P$_{02986–02987M2}$ for EMSA were synthesized in Beijing TSINGKE Biological Technology Company. A gradient concentration of SsaAs was incubated with 30 ng probes in a 20 µL reaction mixture at 25°C for 30 min, and then the samples were loaded on 4% (w/v) native polyacrylamide gels. After electrophoresis, DNA in the gel was stained with SYBR Gold nucleic acid gel stain for 30 min and photographed under UV transillumination using Quantity One.

DNase I footprinting assays were performed as described previously (Xu et al., 2017; Wang et al., 2018). Six putative promoter regions were amplified by PCR using primer pairs listed in Table S1 in Supporting Information to obtain fluorescently unlabeled DNA probe, then these probes were individually inserted into the EcoRV site of pSET152. The resulting recombinant plasmids were used as templates for fluorescent probe amplification using primer pair fotF/R (Table S1 in Supporting Information). Primer fotF was 5′-labled by fluorescent group 6-carboxy-fluorescein (FAM) and primer fotR was 5′-labled by fluorescent group hexachlorofluorescein (HEX). The fluorescent probes (~200 ng) and different amount of SsaA were added to a 50 µL reaction volume and incubated at 25°C for 30 min. Transient DNase I (Promega, USA) digestions were carried out for 35–40 s at 25°C and stopped by adding 5.5 µL of EGTA immediately, then incubated at 65°C for 10 min. The digested DNA complexes were recovered and purified by Wizard SV Gel and PCR Clean-Up System kit (Promega), then analyzed by a 3730XL DNA analyzer with software GeneMarker version 2.2.

RNA isolation and RT-qPCR analyses

Total RNAs were isolated from mycelia of *S. roseosporus* and its derivative strains at three time points (48, 72 and 96 h). RNA isolation, genomic DNA removal, reverse transcription and real-time quantitative PCR (RT-qPCR) were performed as described previously (Zou et al., 2014). All primers used for co-transcriptional analyses and RT-qPCR were listed in Table S1 in Supporting Information. Ribosome S12 rRNA of *S. roseosporus* was used as internal control (Huang et al., 2017). Data shown were the averages of three replicates. Error bars indicated means±standard deviations.

Bioassay, HPLC and HRMS analyses of MRD production

S. roseosporus NRRL 15998 and its derivative strains were inoculated into tubes containing 5 mL liquid TSB and incubated at 28°C, 220 rpm for 2 days as seed culture; then 1.0 mL of seed culture was transferred into a 500-mL shaking flask containing 100 mL of GYM and incubated at 28°C, 220 rpm for 7 days. After fermentation, 2.0 mL supernatant was concentrated to dryness in a vacuum and redissolved in 200 µL ddH$_2$O for bioassay. The rest 98 mL supernatant was adsorbed by 20 mL HP-20 macroporous resin in a 250 mL shaking flask at 160 rpm for 30 min. Subsequently the HP-20 resin was eluted with 80 mL ethyl alcohol in another 250-mL shaking flask at 160 rpm for 30 min. After the eluted alcohol liquor was evaporated to dryness in a vacuum, the dried residue was redissolved in 2 mL of methanol and then used for HPLC and HRMS analyses.

To detect the antibacterial bioactivity of MRD, oxford cups were placed on LB agar containing 1% (v/v) *P. aeruginosa* culture in petri-dish and then 100 µL of the concentrated supernatant was added into each cup, and kept at room temperature for 40–60 min. After further incubation at 37°C for 10–12 h, the antibacterial activity was estimated by measuring the diameter of the inhibition zones.

Twenty microliters of each sample dissolved in methanol was analyzed on Shimadzu SPD-M20A HPLC (Xbridge ODS, 4.6- by 150-mm column, 5 µm; flow rate, 1.0 mL min^{-1}; photodiode array detector, 190 to 800 nm). The linear gradient elution was as follows: 0.1 min, 20% B; 15.0 min, 100 % B; 20.5 min, 100% B; 20.51 min, 20% B; 26.5 min, 20% B. Mobile phases consisted of A—water containing 0.1% formic acid and B—methanol containing 0.1% formic acid.

HRMS experiments were performed on Waters ACQUITY UPLC® and Xevo G2 Qtof MS systems equipped with an electrospray ion source and hybrid quadrupole-time-of-flight (Q-TOF) mass spectrometer. Data were collected by MassLynx V4.1 in positive mode.

References

Alford, M.A., Baghela, A., Yeung, A.T.Y., Pletzer, D., and Hancock, R.E.W. (2020). NtrBC regulates invasiveness and virulence of *Pseudomonas aeruginosa* during high-density Infection. Front Microbiol 11, 773.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic Acids Res 43, W39–W49.

Bhukya, H., Jana, A.K., Sengupta, N., and Anand, R. (2017). Structural and dynamics studies of the TetR family protein, CprB from *Streptomyces coelicolor* in complex with its biological operator sequence. J Struct Biol 198, 134–146.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H., and Weber, T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47, W81–W87.

Breidenstein, E.B.M., de la Fuente-Núñez, C., and Hancock, R.E.W. (2011). *Pseudomonas aeruginosa*: all roads lead to resistance. Trends Microbiol 19, 419–426.

Bugg, T.D.H., Lloyd, A.J., and Roper, D.I. (2006). Phospho-MurNAc-pentapeptide translocase (MraY) as a target for antibacterial agents and antibacterial proteins. Infect Disord Drug Targets 6, 85–106.

Chatterjee, S., Nadkarni, S.R., Vijayakumar, E.K.S., Patel, M.V., Gangul, B.N., Fehlhaber, H.W., and Vertesy, L. (1994). Napsamycins, new *Pseudomonas* active antibiotics of the mureidomycin family from *Streptomyces* sp. HIL Y-82, 11372. J Antibiot 47, 595–598.

Chen, R.H., Buko, A.M., Whittern, D.N., and McAlpine, J.B. (1989). Pacidamycins, a novel series of antibiotics with anti-*Pseudomonas aeruginosa* activity. II. Isolation and structural elucidation. J Antibiot 42, 512–520.

Du, D., Wang, L., Tian, Y., Liu, H., Tan, H., and Niu, G. (2015). Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in *Streptomyces*. Sci Rep 5, 8740.

Guan, H., Li, Y., Zheng, J., Liu, N., Zhang, J., and Tan, H. (2019). Important role of a LAL regulator StaR in the staurosporine biosynthesis and high-production of *Streptomyces fradiae* CGMCC 4.576. Sci China Life Sci 62, 1638–1654.

Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci USA 100, 1541–1546.

Hopwood, D.A., Kieser, T., Bibb, M., Buttner, M., and Chater, K. (2000). Practical *Streptomyces* Genetics. Norwich: John Innes Foundation.

Huang, X., Ma, T., Tian, J., Shen, L., Zuo, H., Hu, C., and Liao, G. (2017). *wblA*, a pleiotropic regulatory gene modulating morphogenesis and daptomycin production in *Streptomyces roseosporus*. J Appl Microbiol 123, 669–677.

Inukai, M., Isono, F., Takahashi, S., Enokita, R., Sakaida, Y., and Haneishi, T. (1989). Mureidomycins A-D, novel peptidylnucleoside antibiotics with spheroplast forming activity. I. Taxonomy, fermentation, isolation and physico-chemical properties. J Antibiot 42, 662–666.

Isono, F., and Inukai, M. (1991). Mureidomycin A, a new inhibitor of bacterial peptidoglycan synthesis. Antimicrob Agents Chemother 35, 234–236.

Isono, F., Inukai, M., Takahashi, S., Haneishi, T., Kinoshita, T., and Kuwano, H. (1989a). Mureidomycins A-D, novel peptidylnucleoside antibiotics with spheroplast forming activity. II. Structural elucidation. J Antibiot 42, 667–673.

Isono, F., Katayama, T., Inukai, M., and Haneishi, T. (1989b). Mureidomycins A-D, novel peptidylnucleoside antibiotics with spheroplast forming activity. III. Biological properties. J Antibiot 42, 674–679.

Jiang, L., Wang, L., Zhang, J., Liu, H., Hong, B., Tan, H., and Niu, G. (2015). Identification of novel mureidomycin analogues via rational activation of a cryptic gene cluster in *Streptomyces roseosporus* NRRL 15998. Sci Rep 5, 14111.

Jin, Y.H., Zhan, Q.Y., Peng, Z.Y., Ren, X.Q., Yin, X.T., Cai, L., Yuan, Y.F., Yue, J.R., Zhang, X.C., Yang, Q.W., et al. (2020). Chemoprophylaxis, diagnosis, treatments, and discharge management of COVID-19: An evidence-based clinical practice guideline (updated version). Mil Med Res 7, 41.

Katz, L., and Baltz, R.H. (2016). Natural product discovery: past, present, and future. J Ind Microbiol Biotechnol 43, 155–176.

Kaysser, L., Tang, X., Wemakor, E., Sedding, K., Hennig, S., Siebenberg, S., and Gust, B. (2011). Identification of a napsamycin biosynthesis gene cluster by genome mining. Chembiochem 12, 477–487.

Khan, S., Muhammad, S., Rauf, A., Khan, A., Rizwan, M., Patel, S., Khan, H., Mahasneh, A.M., and Mubarak, M.S. (2018). Comprehensive review on Ebola (EBOV) virus: Future prospects. Infect Disord Drug Targets 18, 96–104.

Kim, D.E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32, W526–W531.

Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G.L., and Aigle, B. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. Proc Natl Acad Sci USA 108, 6258–6263.

Le, T.B.K., Schumacher, M.A., Lawson, D.M., Brennan, R.G., and Buttner, M.J. (2011). The crystal structure of the TetR family transcriptional repressor SimR bound to DNA and the role of a flexible N-terminal extension in minor groove binding. Nucleic Acids Res 39, 9433–9447.

Li, C., He, H., Wang, J., Liu, H., Wang, H., Zhu, Y., Wang, X., Zhang, Y., and Xiang, W. (2019a). Characterization of a LAL-type regulator NemR in nemadectin biosynthesis and its application for increasing nemadectin production in *Streptomyces cyaneogriseus*. Sci China Life Sci 62, 394–405.

Li, D., Zhang, J., Tian, Y., and Tan, H. (2019b). Enhancement of salinomycin production by ribosome engineering in *Streptomyces albus*. Sci China Life Sci 62, 276–279.

Li, Q., Wang, L., Xie, Y., Wang, S., Chen, R., and Hong, B. (2013). SsaA, a member of a novel class of transcriptional regulators, controls sansanmycin production in *Streptomyces* sp. strain SS through a feedback mechanism. J Bacteriol 195, 2232–2243.

Li, Y., Li, J., Tian, Z., Xu, Y., Zhang, J., Liu, W., and Tan, H. (2016). Coordinative modulation of chlorothricin biosynthesis by binding of the glycosylated intermediates and end product to a responsive regulator ChlF1. J Biol Chem 291, 5406–5417.

Liu, G., Chater, K.F., Chandra, G., Niu, G., and Tan, H. (2013). Molecular regulation of antibiotic biosynthesis in *Streptomyces*. Microbiol Mol Biol Rev 77, 112–143.

Liu, X., Zheng, G., Wang, G., Jiang, W., Li, L., and Lu, Y. (2019). Overexpression of the diguanylate cyclase CdgD blocks developmental transitions and antibiotic biosynthesis in *Streptomyces coelicolor*. Sci China Life Sci 62, 1492–1505.

Lu, F., Hou, Y., Zhang, H., Chu, Y., Xia, H., and Tian, Y. (2017). Regulatory genes and their roles for improvement of antibiotic biosynthesis in *Streptomyces*. 3 Biotech 7, 250.

Mashalidis, E.H., Kaeser, B., Terasawa, Y., Katsuyama, A., Kwon, D.Y., Lee, K., Hong, J., Ichikawa, S., and Lee, S.Y. (2019). Chemical logic of MraY inhibition by antibacterial nucleoside natural products. Nat Commun 10, 2917.

McLean, T.C., Wilkinson, B., Hutchings, M.I., and Devine, R. (2019). Dissolution of the disparate: co-ordinate regulation in antibiotic biosynthesis. Antibiotics 8, 83.

Nguyen, C.T., Dhakal, D., Pham, V.T.T., Nguyen, H.T., and Sohng, J.K.

(2020). Recent advances in strategies for activation and discovery/characterization of cryptic biosynthetic gene clusters in *Streptomyces*. Microorganisms 8, 616.

Niu, G., Li, Z., Huang, P., and Tan, H. (2019). Engineering nucleoside antibiotics toward the development of novel antimicrobial agents. J Antibiot 72, 906–912.

Niu, G., and Tan, H. (2015). Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. Trends Microbiol 23, 110–119.

Paget, M.S.B., Chamberlin, L., Atrih, A., Foster, S.J., and Buttner, M.J. (1999). Evidence that the extracytoplasmic function sigma factor çE is required for normal cell wall structure in *Streptomyces coelicolor* A3 (2). J Bacteriol 181, 204–211.

Pan, Y., Liu, G., Yang, H., Tian, Y., and Tan, H. (2009). The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. Mol Microbiol 72, 710–723.

Ray, S., Maitra, A., Biswas, A., Panjikar, S., Mondal, J., and Anand, R. (2017). Functional insights into the mode of DNA and ligand binding of the TetR family regulator TylP from *Streptomyces fradiae*. J Biol Chem 292, 15301–15311.

Wang, W., Zhang, J., Liu, X., Li, D., Li, Y., Tian, Y., and Tan, H. (2018). Identification of a butenolide signaling system that regulates

nikkomycin biosynthesis in *Streptomyces*. J Biol Chem 293, 20029–20040.

Xie, Y., Chen, R., Si, S., Sun, C., and Xu, H. (2007). A new nucleosidyl-peptide antibiotic, sansanmycin. J Antibiot 60, 158–161.

Xu, J., Zhang, J., Zhuo, J., Li, Y., Tian, Y., and Tan, H. (2017). Activation and mechanism of a cryptic oviedomycin gene cluster via the disruption of a global regulatory gene, *adpA*, in *Streptomyces ansochromogenes*. J Biol Chem 292, 19708–19720.

Yin, H., Wang, W., Fan, K., and Li, Z. (2019). Regulatory perspective of antibiotic biosynthesis in *Streptomyces*. Sci China Life Sci 62, 698–700.

Zhang, W., Ostash, B., and Walsh, C.T. (2010). Identification of the biosynthetic gene cluster for the pacidamycin group of peptidyl nucleoside antibiotics. Proc Natl Acad Sci USA 107, 16828–16833.

Zhang, X., Hindra, X., and Elliot, M.A. (2019). Unlocking the trove of metabolic treasures: activating silent biosynthetic gene clusters in bacteria and fungi. Curr Opin Microbiol 51, 9–15.

Zhang, Y.Y., Zou, Z.Z., Niu, G.Q., and Tan, H.R. (2013). *jadR** and *jadR2* act synergistically to repress jadomycin biosynthesis. Sci China Life Sci 56, 584–590.

Zou, Z., Du, D., Zhang, Y., Zhang, J., Niu, G., and Tan, H. (2014). A γ-butyrolactone-sensing activator/repressor, JadR3, controls a regulatory mini-network for jadomycin biosynthesis. Mol Microbiol 94, 490–505.

SUPPORTING INFORMATION

The supporting information is available online at 10.1007/s11427-020-1892-3. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.