NOTES

Mutational Analysis of the Regulatory Region of the srfA Operon in Bacillus subtilis

MICHIKO M. NAKANO* AND PETER ZUBER
Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, Shreveport, Louisiana 71130

Received 19 October 1992/Accepted 10 December 1992

Transcription of the Bacillus subtilis srfA operon is dependent on the transcriptional activator ComA. Mutational analysis of the srfA regulatory region suggests that two regions of dyad symmetry upstream of the srfA promoter may function in transcriptional activation by facilitating a cooperative interaction between ComA dimers.

The srfA operon, which encodes the multifunctional enzymes (28, 46) that catalyze the biosynthesis of the lipopeptide surfactin in Bacillus subtilis (24, 25), is also required for competence development (24, 41) and efficient sporulation (24). Strains that do not produce surfactin possess srfA but lack an intact sfp gene which functions in the production of surfactin through an unknown mechanism (23, 29). srfA was previously shown to occupy an intermediate position in the regulatory cascade that controls late competence gene expression (6–8, 12, 24, 30, 41). The function of the early competence genes comp (44), comA (42), comQ (43), and spo0K (32, 36) in competence establishment is primarily to activate srfA operon expression (12, 30). Expression of srfA is controlled at the transcriptional level by the two-component regulatory pair (1, 2, 40) ComP and ComA (12, 27, 41). ComP is presumably a histidine kinase sensor protein with a predicted membrane-spanning domain (44). It is believed to modulate the activity of its cognate response regulator, ComA (42), through phosphorylation. Phosphorylated ComA has been previously thought to act as a transcriptional activator of srfA as well as degQ (20, 45), a pleiotropic regulatory gene which affects the expression of genes encoding degradative enzymes, and gsaA (21, 22), a glucose starvation-induced operon that acts as a negative regulator of sporulation initiation.

Our previous work has shown that a region of dyad symmetry (TTGCGG-N4-CGGCAA: ComA box 1) between positions −118 and −103 of the srfA promoter plays a critical role in the positive control of srfA transcription by the ComP-ComA signal transduction system (26). Another comA-regulated gene, degQ, has an imperfect dyad symmetry (TTGCGG-N4-CGGCAG) from positions −71 to −56, which is nearly identical to that of the ComA box 1 found in the srfA regulatory region (20, 26, 45). A recent report also identified an imperfect inverted repeat upstream of the gsaA promoter region between positions −75 and −60 (TTGCGG-N4-CGGAAA) which shows homology to the dyad symmetry in the srfA and degQ promoter regions (21). These findings, together with the fact that ComA contains a possible helix-turn-helix motif characteristic of DNA binding proteins (42), suggest that phosphorylated ComA binds to the ComA box to induce srfA transcription at the onset of the stationary phase of the growth curve. A further comparison of the promoter sequences of degQ, gsaA, and srfA identified another weak dyad symmetry (TTTGGG-N4-CGGCAAG) in the region from positions −74 to −59 of PsrfA. To examine whether this putative ComA box (ComA box 2) is functional in srfA transcription initiation, mutant derivatives of the srfA promoter were constructed.

Two M13mp9 derivatives containing copies of the srfA promoter regions were used for oligonucleotide-directed mutagenesis. One of them, msf29, carries 300-bp PsrfA (the srfA promoter region) between the 5′-end DraI and HaeIII sites. This fragment contains the two ComA boxes. The PsrfA fragment from msf29 was subcloned in front of the promoterless lacZ gene of pTKlac (15) (pMMN92 in reference 26 and Fig. 1). The other M13 derivative, msf33, contains the HinII and PvuII fragments of PsrfA harboring only ComA box 2. pMMN88 (26) (Fig. 1) was constructed by introducing the PsrfA fragment from msf33 into pTKlac. The srfA promoter regions in msf29 and msf33 were mutagenized with synthetic oligonucleotides by the procedure described previously (26). The clones with the desired mutations were identified by DNA sequencing and examined for the absence of other spontaneous mutations before further studies were performed. The mutated PsrfA fragments were inserted into pTKlac in order to construct plasmids designed for assaying promoter activity. pMMN102 was derived from msf33, and pMMN97, −101, and −103 were constructed from msf29. pMMN104 with a 10-bp insertion between the two ComA boxes was constructed from the M13 derivative which contained PsrfA of pMMN103 (5-bp insertion between the ComA boxes). The oligonucleotides used for the mutagenesis are as follows (the bases changed are underlined, and the bases inserted are overlined): pMMN97, shown in reference 26; pMMN101, AATCTTGGCCGATCCCGACAGA ACT; pMMN102, AATCTTGGCCGATCCCGACAGA CT; pMMN103, GCTGTAAGGCAAACTGATACGAAACTGGCAT; and pMMN104, GTAAATGACTGGATCCTTGAGAATTCGAC. The oligonucleotide used for the insertion of 5 bp was designed so that a new EcoRV site was created. Insertion of 10 bp resulted in the loss of the EcoRV site and the generation of a new BamHI site. This

* Corresponding author.
FIG. 1. Mutational analysis of ComA boxes. The promoter region of the srfA operon is shown. l-met and P with a large arrow show the srfA translational and transcriptional start sites, respectively. The region containing two dyad symmetries (indicated by small arrows) is expanded below the map, and the sequence of the ComA boxes is shown. The detailed surrounding sequence was reported in our previous work (26). The nucleotides written in capital letters represent complementary bases of the dyad symmetry region, and those in lowercase letters are mismatched. The nucleotides marked by dots are those generated by site-directed mutagenesis. Arrowheads at 5 and 10 bp are the nucleotide insertions between ComA boxes. The plasmids with the mutated promoters fused to lacZ are shown as pMMN numbers. β-Galactosidase specific activities directed by the intact and mutated promoters were measured during growth and sporulation in the wild-type (W. T.) (OKB105) and ΔcomP comA) (OKB167) strains. The maximal activities are listed in the column on the right.

difference in restriction enzyme cleavage patterns was used to identify the mutations. The pTKlac derivatives carrying the wild-type and the mutant srfA promoters were used to transform B. subtilis ZB307A, a lysogen of SPβc2del2::Tn917::psK10Δ6 (47) with selection for chloramphenicol resistance (Cm'). SPβ lyases carrying the srfA-lacZ fusions from ZB307A were prepared as described in reference 27, and the lysates were used to infect and lysogenize B. subtilis OKB105 and OKB167 [ΔcomP comA) cells. The activity of β-galactosidase was monitored during exponential growth and sporulation, and the maximal activities are summarized in Fig. 1. As shown in our previous work (26), deletion of ComA box 1 (pMMN88) and a mutation which weakened the dyad symmetry of ComA box 1 (pMMN97) led to severe reduction of srfA transcription compared with that by the wild-type promoter (pMMN92). The complete elimination of dyad symmetry in ComA box 2 also resulted in the loss of β-galactosidase activity (pMMN101), in spite of the presence of the intact ComA box 1. This result shows that ComA box 2, as well as ComA box 1, is essential for srfA transcription.

Figure 1 also shows that pMMN102, which has perfect dyad symmetry in ComA box 2 and a deletion of ComA box 1, resulted in wild-type levels of the promoter activity when introduced into B. subtilis cells. This strongly argues that ComA box 2 is directly involved in activation of srfA transcription and that ComA box 1 may facilitate binding of ComA to the weak ComA box 2. To examine the nature of the interaction between the two putative ComA-binding sites, a 5-bp insertion was created between ComA box 1 and ComA box 2, which resulted in the positioning of the two ComA boxes on opposite faces of the DNA helix (pMMN103 in Fig. 1) and a significant reduction in promoter activity. A 10-bp insertion between the two ComA boxes which placed the centers on the same side of the DNA helix resulted in levels of promoter activity that were twofold higher than those by wild-type srfA-lacZ.

The data presented here strongly suggest that the ComA protein cooperatively binds to the regulatory region of srfA at two sequences of dyad symmetry (the ComA boxes). It is proposed that ComA dimers bind to the two regions of dyad symmetry upstream of the srfA promoter, since they are required for srfA transcription and since such sequences are often the sites of interaction with transcriptional regulatory proteins that possess an α-helix-turn-α-helix motif. An amino acid sequence found in the ComA protein. On the basis of this and the mutational analysis described herein, the most likely model for interaction of ComA proteins with the srfA promoter is shown in Fig. 2. The upstream ComA box 1, with the higher affinity for ComA, is first occupied by a ComA dimer. The model includes the formation of a ComA tetramer which facilitates binding to the low-affinity ComA box 2. This cooperative binding results in the formation of a loop consisting of the intervening DNA. Direct interaction between RNA polymerase and the ComA dimer bound to the ComA box 2 may be essential for srfA transcription. This model is supported by the DNase I footprinting analysis by Ruggiani and Dubnau (35).

That an interaction between the two ComA-DNA complexes occurs is supported by the finding that the function of the regulatory sequences is helix face dependent. The centers of the two ComA boxes are separated by four turns of the helix (44 bp), and the insertion of 5 bp (one-half a turn) resulted in a substantial loss of activity. Insertion of 10 bp between the ComA boxes creates five helical turns and results in a twofold-higher level of srfA-lacZ activity than what is observed with wild-type srfA-lacZ. The higher

FIG. 2. Schematic model for srfA transcriptional initiation. ComA box 1, with a higher affinity for a ComA dimer, is shown by large arrows, and ComA box 2, with a lower affinity, is shown by small arrows. The tetramer formation of ComA proteins causes the intervening DNA loop formation. The interaction of RNA polymerase (RNA Pol) with ComA allows srfA transcription.
β-galactosidase activity may be explained by the larger loop size allowing for easier loop formation, although, in the case of araBAD, no lower limit, in terms of function, was found for the loop size (16). Another explanation for the higher activity is that 54 bp in pMMN104 may be closer to an integral turn of the helix than is 44 bp in pMMN92. The helical repeat of linear DNA in vitro is reported to be about 10.5 bp per turn, and the in vivo helical repeat of the DNA between araI and arao2 is shown to be 11.1 bp per turn (16). Oscillations in activation as a function of spacing between the two ComA boxes may similarly provide information about the helical repeat distance of the PsrFA DNA in vivo. Similar DNA looping in prokaryotic systems has been known to be involved in transcriptional regulation (for reviews, see references 33, 38, and 39) of the artificially constructed λ operators (13) and the repression of the gal (11, 18), lac (3, 10, 31), and deo (19) operons in Escherichia coli. In more complex cases, such as the araBAD operon, repression is mediated by DNA looping (9) and by arabinose-mediated loop breaking when the operon is induced (17). DNA loop formation which functions in transcriptional activation includes the glnALG operon (34), the glnHPQ operon (4) in E. coli, and the rif operon (14) in Klebsiella pneumoniae, the activation of which requires α^44 RNA polymerase and the phosphorylated form of NRk (NtrC) or NiFA. Involvement of integration host factor-induced DNA bend is also required in these last two cases to facilitate loop formation. Little is known about the requirements for DNA loop formation in the regulation of gene expression in B. subtilis. Czaplewski et al. reported that the AhrC hexamer binds at two distinct sites within argC, forming a repression loop (5). DNA bending was suggested to facilitate binding of Spo0A to two sites in the spoIIG promoter (37).

We thank D. Dubnau for valuable discussions and critical reading of the manuscript.

This work was supported by grants GM39479 and GM45898 (to P.Z.) from the National Institutes of Health.

REFERENCES