JOURNAL OF BACTERIOLOGY, Mar. 2001, p. 1921–1927 Vol. 183, No. 6
0021-9193/01/$04.00+0 DOI: 10.1128/JB.183.6.1921–1927.2001

The −10 Region Is a Key Promoter Specificity Determinant for the *Bacillus subtilis* Extracytoplasmic-Function σ Factors σ^X and σ^W

JIAN QIU AND JOHN D. HELMANN*

Department of Microbiology, Cornell University, Ithaca, New York 14853-8101

Received 22 September 2000/Accepted 20 December 2000

Transcriptional selectivity derives, in large part, from the sequence-specific DNA-binding properties of the σ subunit of RNA polymerase. There are 17 σ factors in *Bacillus subtilis* which, in general, recognize distinct sets of promoters. However, some σ factors have overlapping promoter selectivity. We hypothesize that the overlap between the regulons activated by the σ^X and σ^W factors can be explained by overlapping specificity for the −10 region: σ^X recognizes −10 elements with the sequence CGAC and σ^W recognizes CGTA, while both can potentially recognize CGTC. To test this model, we mutated the σ^X-specific autoregulatory site (P_X), containing the −10 element CGAC, to either CGTC or GCTA. Conversely, the σ^W autoregulatory site (P_W) was altered from CGTA to CGTC or CGAC. Transcriptional analyses, both in vitro and in vivo, indicate that changes to the −10 element are sufficient to switch a promoter from the σ^X to the σ^W regulon or, conversely, from the σ^W to the σ^X regulon, but context effects clearly play an important role in determining promoter strength. It seems likely that these subtle differences in promoter selectivity derive from amino acid differences in conserved region 2 of σ, which contacts the −10 element. However, we were unable to alter promoter selectivity by replacements of two candidate recognition residues in σ^W.

While the sequencing of bacterial genomes is proceeding at a rapid pace, functional annotation remains a formidable challenge. Typically, half or more of all predicted open reading frames encode proteins which have no functionally characterized homologs or which are related only to large classes of proteins with a wide range of functions (e.g., transporters or oxidoreductases). Additional clues to gene function can often be gleaned from careful analysis of operon organization (13, 23) or by identifying groups of genes (stimulons) that are coordinately activated under specific conditions (12, 19). Interpreting global transcriptional profiles requires that genes be grouped into regulons that share a common regulatory factor. Regulons activated by secondary σ factors are often a significant component of the stimulons activated in bacteria under specific stress conditions or in response to changing environmental conditions (3–5).

Sequencing of the *Bacillus subtilis* genome revealed genes for seven previously unidentified σ factors, all belonging to the extracytoplasmic-function (ECF) subfamily (15). Mutants with alterations in these genes are viable and do not have obvious phenotypes, although the mutant strains are often somewhat more sensitive to selected stress conditions (7, 8). Therefore, to define the roles of the ECF σ factors in *B. subtilis,* we have sought to identify target genes dependent on ECF σ factors for their expression (10, 11).

ECF σ factors often positively regulate their own synthesis (16, 21). The identification of the corresponding autoregulatory promoters provides useful clues to promoter selectivity which can then be used to search the genome for additional target sites. To develop this strategy, we analyzed a large collection of point mutations in the σ^X-dependent autoregulatory promoter (P_X) to define bases critical for activity (11). As expected for a σ^{70} class holoenzyme, the critical bases are clustered near −35 and −10 relative to the transcription start point (tGtAACN$_{17}$CGaC; bases with no allowable substitutions are in uppercase). Using this information, we were able to identify a number of promoters that are recognized by σ^X both in vivo and in vitro. However, some of the target genes we identified were still transcribed, from the same start point, even in a *sigX* null mutant (11). This suggested that at least one other holoenzyme has an overlapping specificity with σ^X.

In parallel with these studies of the σ^X regulon, we also initiated an analysis of the σ^W regulon. Like *sigX, sigW* is transcribed from an autoregulatory promoter element, P_W (TGAAACN$_{16}$CGTA) (9). Analysis of the genome revealed 15 additional operons with candidate promoters identical to P_W (in the −35 and −10 elements and spacer length), and all 15 of these sites are σ^W dependent both in vivo and in vitro (10). Thus, promoter sequence comparisons have proven to be a valuable approach to defining σ factor regulons.

Despite considerable sequence similarity between P_W and P_X, these promoters are exclusively recognized by the cognate σ in vivo and in vitro (9–11). Sequence comparisons, in conjunction with the mutational analysis of P_X, suggested that this selectivity might derive from the −10 region sequences. P_X contains the −10 element sequence CGAC, while P_W contains CGTA (9). Characterization of the σ^X and σ^W regulons also identified several promoters with the −10 region sequence CGTC. In vitro, these promoters seem to be recognized by both holoenzyme forms (9). In vivo, these promoters seem to depend primarily upon σ^X, but often both σ^X and σ^W contribute to expression.

It is difficult to accurately predict promoters based solely on −35 and −10 consensus sequences. Context effects may play a large role in promoter selectivity, and important discriminatory elements may residue outside the classically defined −35 and

* Corresponding author. Mailing address: Department of Microbiology, Cornell University, Ithaca, NY 14853-8101. Phone: (607) 255-6570. Fax: (607) 255-3904. E-mail: jdh9@cornell.edu.

−10 elements. As a test of our model for promoter recognition by σ^X and σ^W, we have engineered mutations within the −10 element to convert P_X into a σ^W-dependent promoter and, conversely, to convert P_W into a σ^X-dependent promoter. Our results indicate that changes to the −10 region are sufficient for altering holoenzyme selectivity both in vivo and in vitro. These observations support the notion that the −10 region is a critical selectivity determinant for these two ECF σ factors.

MATERIALS AND METHODS

Bacterial strains, growth media, and antibiotics. All *B. subtilis* and *Escherichia coli* strains used in this work are listed in Table 1. Strains were grown at 37°C with vigorous shaking in Luria broth (LB) medium unless otherwise indicated. In *E. coli*, ampicillin resistance was selected by using 100 μg of ampicillin/ml. In *B. subtilis*, antibiotics used for selection were neomycin at 8 μg/ml, spectinomycin at 100 μg/ml, kanamycin at 20 μg/ml, and macrolides-lincomycin-streptogramin B at 25 μg/ml (lincomycin) and 1 μg/ml (erythromycin).

Mutagenesis of P_X and P_W. To introduce mutations into P_X, 1 nmol of oligonucleotide 313 (−10 region, CGTAaa [−10 consensus bases are in uppercase]) or 314 (−10 region, CGTAta) were mixed with 1 nmol of the reverse oligonucleotide XH135 in 50 μl of TMED buffer (10 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol [DTT], 50 mM NaCl), heated at 95°C, and cooled slowly to allow annealing. Deoxynucleoside triphosphates (10 μl) at 25 mM and 2 μl of Sequenase, version 2, were added, and the oligonucleotides were extended at 37°C for 1 h. The duplex product was purified using a Qiagen purification kit, digested with *Hin*dIII and *Bam*HI, and ligated into pJPM122 to construct pJQ1 and pJQ2, respectively. After transformation into *E. coli* DH5α with selection for ampicillin resistance, plasmids were recovered and the sequence of the promoter region was verified by DNA sequencing. A plasmid containing a −10 element of CGTCaa was obtained from the saturation mutagenesis studies by Huang and Helmann (11) and renamed pJQ3. The plasmids pJQ1, pJQ2, and pJQ3 contain the −44 to +11 regions of P_X variants.

To introduce mutations into P_W, P_W and its variants were amplified using the forward primer XH180 (with a *Hin*dIII site) and one of three reverse primers (416 to 418) (Table 2). The reverse primers carry a *Bam*HI site and the indicated −10 element (Table 2). The resulting PCR fragments were then cloned into pJPM122 (26) to construct pJQ8, pJQ9, and pJQ10, following the procedures described above. The plasmids pJQ8, pJQ9, and pJQ10 contain the −79 to +6 regions of P_W and its variants.

Construction and analysis of SPβ reporter phage. To recombine the promoter-*cat-lacZ* fusions into the SPβ prophage, each pJPM122 derivative was linearized by digestion with *Sca*I and transformed into ZB307A [W168 SPβc2Δ2::Tn*917*::pSK10Δ6 (MLS^r)] (29) with selection for neomycin resistance. To transduce each reporter fusion into various genetic backgrounds, SPβ lysates were prepared by heat induction at 50°C for 10 min followed by continued incubation at 37°C for 90 min (1). The resulting lysates were used to transduce recipient strains using standard techniques (Table 1).

Expression levels for each reporter fusion were determined by β-galactosidase assays of cultures grown in LB media. Each strain was grown overnight in LB medium containing appropriate antibiotics and diluted 100-fold into LB medium without antibiotics. Samples of cells were taken after growth at 37°C for 3 h, harvested, and frozen at −80°C. β-Galactosidase activity was assayed as described by Miller (20).

Overproduction and purification of σ^W mutants. Mutagenesis of *sigW* was achieved by two rounds of PCR with Vent DNA polymerase (New England Biolabs). Primer 452 is located at the 5′ end of *sigW* and contains a ribosome binding site and an *Xba*I restriction site, and primer 453 is located at the 3′ end of *sigW* and contains a *Xho*I restriction site. Primers 454 and 455 introduce mutations in the codon of Arg75 on the sense and antisense strands, respectively. Primer 456 and 457 introduce mutations in the codons for both Arg75 and Asn79. In the first round of PCR, primer 452 and primer 455 or 457 were used to amplify the mutated 5′ fragments of the *sigW* gene from *B. subtilis* chromosomal DNA, and primer 453 and primer 454 or 456 were used to amplify the mutated 3′ fragments of *sigW* gene. In the second round of PCR, the products from the first round were used as templates and oligonucleotides 452 and 453 were used as primers. The amplified fragments from the second round were digested with *Xba*I and *Xho*I and cloned into pET17b to construct pJQ27 and pJQ28. The sequences of both cloned mutant *sigW* genes were confirmed by DNA sequencing.

For purification of σ^W mutants, pJQ27 and pJQ28 were transformed into *E. coli* BL21/DE3 to generate strains HE4524 and HE4525. Both strains were grown to mid-logarithmic phase at 37°C in 500 ml of LB medium with 100 μg of ampicillin/ml. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to 0.4 mM, and cells were harvested after further incubation for 3 h. After centrifugation, the cell pellets were suspended in 20 ml of disruption buffer (50 mM Tris-HCl [pH 8.0], 2 mM EDTA, 0.1 mM DTT, 1 mM β-mercaptoethanol, 233 mM NaCl, 10% glycerol) and lysed by sonication, and the inclusion bodies were

TABLE 1. Strains and plasmids used in this work

Strain or plasmid	Description	Source or derivation
E. coli		
DH5α	*supE44*Δ *lacU169* (ϕ80*lacZ*Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Lab stock
BL21	F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm*	Novagen
BL21/DE3	BL21 with λ DE3 (T7 RNAP under *lac* control)	Novagen
HE4524	BL21/DE3 containing pJQ27	This work
HE4525	BL21/DE3 containing pJQ28	This work
B. subtilis		
CU1065	W168 *trpC2 att*SPβ	Lab stock
ZB307A	W168 SPβc2Δ2::Tn*971*::pSK10Δ6 (MLS^r)	29
HB7007	CU1065 *sigX*::*spc* (Spc^r) $SigX^-$	8
HB7013	CU1065::pVA29 (MLS^r) $RsiX^-$	8
HB7022	CU1065 SPβ7019 (MLS^r Neo^r) P_X-*cat-lacZ*	8
HB0010	CU1065 *sigW*::*kan* (Km^r) $RsiW^-$	
HB0020	CU1065 *sigW*::*erm* (MLS^r) $SigW^-$	
HB0030	HB0020 *sigX*::*spc* (MLS^r Spc^r) $SigX^-$ $SigW^-$	
HB7019	ZB307A SPβ7019 (MLS^r Neo^r) P_X-*cat-lacZ* (wild type; CGACta)	8
HB4506	ZB307A × (pJQ1 × *Sca*I) P_X-*cat-lacZ* with −10 CGTAaa (X3)	This work
HB4507	ZB307A × (pJQ2 × *Sca*I) P_X-*cat-lacZ* with −10 CGTAta (X2)	This work
HB4508	ZB307A × (pJQ3 × *Sca*I) P_X-*cat-lacZ* with −10 CGTCta (X1)	This work
HB4521	ZB307A × (pJQ8 × *Sca*I) P_X-*cat-lacZ* (wild type; CGTA)	This work
HB4522	ZB307A × (pJQ9 × *Sca*I) P_X-*cat-lacZ* with −10 CGTC (W1)	This work
HB4523	ZB307A × (pJQ10 × *Sca*I) P_X-*cat-lacZ* with −10 CGAC (W2)	This work
Plasmids		
pJPM122	Integrational plasmid for *cat-lacZ* fusion construction	26
pET17b	Expression vector under the control of T7 promoter	Novagen
pJQ1	P_X variant with a CGTAaa −10 element in pJPM122	This work
pJQ2	P_X variant with a CGTAta −10 element in pJPM122	This work
pJQ3	P_X variant with a CGTCaa −10 element in pJPM122	11
pJQ8	P_W with the consensus CGTA −10 element in pJPM122	This work
pJQ9	P_W variant with a CGTC −10 element in pJPM122	This work
pJQ10	P_W variant with a CGAC −10 element in pJPM122	This work
pJQ27	The *sigW* gene with R75S mutation cloned into pET17b	This work
pJQ28	The *sigW* gene with R75S N79H mutations in pET17b	This work

TABLE 2. Oligonucleotides used in this study

Oligonucleotide	Sequence (5′-3′)[a]	Function
313	GCATAG*AAGCTT*AGTTGTAATGTAACTTTTCAAGCTATTCATACGTAAAAAAAGTGAACGGAG	P_X mutagenic primer (X2)
314	GCATAG*AGAGCTT*AGTTGTAATGTAACTTTTCAAGCTATTCATACGTATAAAAAGTGAACGGAG	P_X mutagenic primer (X3)
XH135	CGC*GGATCC*CCTCCGTTCACTTTT	P_X reverse primer
XH180	ACGAAT*AAGCTT*CTACACCCTGCCAAA	P_W forward primer
416	CGC*GGATCC*GGTCTGTATGTATACGAGCTTCGTTTCAAAAG	P_W primer (WT)
417	CGC*GGATCC*GGTCTGTATGTAGACGAGCTTCGTTTCAAAAG	P_W mutagenic primer (W1)
418	CGC*GGATCC*GGTCTGTATGTAGTCGAGCTTCGTTTCAAAAG	P_W mutagenic primer (W2)
452	ACA*TCTAGA*CCGGTGAAGGCAGAGG	*sigW* forward (*Xba*I)
453	TGTTCA*CTCGAG*CTCATTTCATCACCCCAC	3′ end of *sigW* (*Xho*)
454	GGCTTTATAGTATCGCGACCAATTTG	R75 mutagenic (sense)
455	CAAATTGGTCGCGATACTATAAAGCCAAG	R75 mutagenic (antisense)
456	GGCTTTATAGTATCGCGACCCATTTGACCATTG	R75 and N79 mutagenic (sense)
457	GGTCAAATGGGTCGCGATACTATAAAGCCAAG	R75 and N79 mutagenic (antisense)
435	CATAG*AAGCTT*AGTTGTAATGTAACTTTTCAAG	P_X primer for pJPM122 derivatives
366	ACTCTCCGTCGCTATTGTAACCAG	*cat* gene; for promoter amplification for runoffs

[a] Introduced restriction sites used for cloning are in italics, −10 region sequences (and relevant flanking bases for P_X) are underlined, and codons altered by mutagenesis are double underlined.

recovered by centrifugation. The inclusion bodies were washed twice with 100 ml of TEDG buffer (50 mM Tris-HCl [pH 8.0], 0.1 mM EDTA, 1 mM DTT, 10% glycerol) containing 0.5% (vol/vol) Triton X-100 and 10 mM EDTA and then dissolved in 10 ml of TEDG–6 M guanidine hydrochloride. A portion (2.5 ml) was gradually diluted to 250 ml with TEDGX (TEDG containing 0.01% Triton X-100) to allow renaturation of σ^W mutants and then loaded onto a 10-ml heparin-Sepharose CL-6B column equilibrated with TEDGX. After being washed with 80 ml TEDGX–0.2 M NaCl, the σ^W mutant proteins were eluted with TEDGX–0.5 M NaCl. The peak fractions were concentrated with Centricon 10 from Amicon and stored at −80°C.

In vitro transcription assays. Runoff transcription assays were performed with DNA products from PCR as the templates. P_X and its variants were amplified from *B. subtilis* HB7022 [CU1065 SPβ7019 P_X-*cat-lacZ* (MLS^r Neo^r)] chromosomal DNA or plasmid pJQ1, pJQ2, or pJQ3 with primers 435 and 366. P_W and its variants were amplified from plasmid pJQ8, pJQ9, or pJQ10 with primers XH180 and 366. Primer 366 is located within the *cat* gene, and the PCR-amplified products contain the promoter regions and a 263-bp 5′ fragment of the *cat* gene.

B. subtilis core RNA polymerase (RNAP) (14), σ^X (8), σ^W (9), and δ (17) preparations have been described previously. Typical transcription reaction mixtures (20 μl) contained 0.36 pmol of core RNAP, 4.5 pmol of σ, 4.2 pmol of δ, and 0.04 pmol of template DNA in transcription buffer (20 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 50 mM KCl, 0.5 mM DTT, 0.1 mg of bovine serum albumin/ml, 5% [vol/vol] glycerol, and the RNase inhibitor RNasin from Promega at 0.8 U/reaction), to which were added nucleoside triphosphate mixtures containing 10 nmol of ATP, GTP, and CTP, 1 nmol of UTP, and 0.6 pmol of [α-^{32}P]UTP (3,000 Ci/mmol). Core RNAP, σ, and δ were mixed on ice for 15 min to form RNAP holoenzyme before the addition of template DNA and incubation at 37°C for 10 min to allow promoter binding. Nucleoside triphosphates were added, and transcription was allowed to proceed for 7 min at 37°C prior to addition of 6 μg of heparin/reaction and an additional 11 min of incubation. Reactions were terminated by the addition of 80 μl of stop solution (2.5 M NH_4 acetate, 10 mM EDTA, and 0.1 mg of glycogen/ml), extracted with phenol-chloroform, and precipitated with ethanol. The pellets were dissolved in 8 μl of loading buffer (20 μg of xylene cyanol FF/ml, 20 μg of bromophenol blue/ml, and 60 mg of urea/ml in 1× Tris-borate-EDTA buffer) and subjected to 8 M urea–6% polyacrylamide gel electrophoresis. Reaction products were visualized by using a Molecular Dynamics PhosphorImager system and ImageQuant software.

Experiments to test the selectivity of σ^W mutant proteins were performed as described above or using a template competition assay containing both P_W and the CGTC variant. To distinguish the RNA products from the two fragments, the PCR product carrying the CGTC P_W variant was digested with *Dde*I. This reduces the size of the PCR fragment from 353 to 281 bp and leads to a concomitant decrease in the length of the runoff transcript. Transcripts were quantified using a Molecular Dynamics PhosphorImager system and ImageQuant software, and the molar ratios were calculated after taking into account the difference in UMP content of each runoff RNA.

RESULTS

Activity of P_X and P_W variants in vitro. P_X (the σ^X-dependent promoter preceding *sigX*) and P_W (the σ^W-dependent promoter preceding *sigW*) have very similar sequences in the −10 and −35 elements, yet $E\sigma^X$ cannot recognize P_W nor can $E\sigma^W$ recognize P_X. Comparison of P_X with P_W, in the vicinity of the −35 and −10 elements, reveals seven sequence differences that might account for the mutually exclusive recognition of these two promoters (Fig. 1). Previous mutational analysis of P_X indicates that changing the sequences at five of these positions does not eliminate promoter activity (11), and alignment of known σ^X-dependent promoters supports the idea that these are unlikely to be key selectivity determinants (11). Nor is the difference in spacer length between P_X and P_W likely to be an important factor: both σ^X and σ^W can recognize promoters with either 16- or 17-base spacer regions (10, 11). These observations led us to focus our attention on the −10 element.

```
                    A
               A    C (Δ)          N  T T        +1
P_X    aaTGTAACttttcaagctattcataCGACaaaaaagtGaacg
         1    2        3       4  567
P_W    atTGAAACcttttgaaa-cgaagctCGTAtacatacaGAccg
```

FIG. 1. Sequence comparison of P_X and P_W. The sequence of the σ^X-dependent autoregulatory site, P_X, is shown, with critical bases, as judged by mutational analysis (11), in bold. Alignment of P_X with P_W reveals a high degree of sequence similarity with related −35 and −10 sequences, also in bold. The numbers represent seven positions that might account for the mutually exclusive recognition of these two promoters. Positions 1, 2, 4, and 7 are viewed as unlikely discriminatory features, since the corresponding mutations in P_X retain at least 25% of wild-type activity in vivo (11). A change of A to T at position 5 also retains activity in vivo and this activity is eliminated in a *sigX* mutant. A variant of P_X with a single base deletion (Δ), to generate a 16-bp spacer as in P_W, retains activity, although at a considerably lower level than P_X. This analysis, together with alignment of promoters shown previously to depend on σ^X, σ^W, or both, suggests a critical role for position 6 (located at −9 relative to the most common transcriptional start point) in the discrimination of P_X from P_W (9).

TABLE 3. Sequences of P_X, P_W and variants used in this study

Promoter	Sequence[a]	IVT[b]	β-Gal[c]
P_X	aaTGTAACttttcaagctattcataCGACaaaaaagtgaag	X	X
X1	aaTGTAACttttcaagctattcataCG**T**Caaaaaagtgaacg	X, w	X
X2	aaTGTAACttttcaagctattcataCG**TA**aaaaaagtgaacg	W	NA
X3	aaTGTAACttttcaagctattcataCG**TAt**aaaaagtgaacg	W	W
P_W	atTGAAACcttttgaaa-cgaagctCGTAtacatacagaccg	W	W
W1	atTGAAACcttttgaaa-cgaagctCGT**C**tacatacagaccg	X, W	X, w
W2	atTGAAACcttttgaaa-cgaagctCG**AC**tacatacagaccg	X	X, w

[a] Conserved −35 and −10 elements are in uppercase. Bases altered by mutagenesis are in bold. The experimentally mapped transcription start sites are underlined.
[b] The activity of each promoter in in vitro transcription (IVT) assays is indicated. A lowercase letter indicates weak activity.
[c] The activity of each promoter in vivo is summarized. NA, no activity was detected.

To determine whether the −10 element is the key determinant that distinguishes between $E\sigma^X$ and $E\sigma^W$, we mutated the −10 elements so that P_X acquired the −10 element CGTC or CGTA and P_W acquired the −10 element CGTC or CGAC (Table 3). Reconstituted RNAP holoenzyme was then used to determine promoter activity in runoff transcription assays (Fig. 2). Under our reaction conditions, core RNAP alone ($\beta\beta'\alpha_2\delta$) could not recognize or transcribe from either P_X or P_W, and enzyme reconstituted with σ^A recognized a σ^A-dependent promoter with high activity but did not recognize either P_X or P_W (data not shown). Consistent with previous reports (9), reconstituted $E\sigma^X$ directs transcription from P_X, but not P_W, while $E\sigma^W$ directs transcription from P_W, but not P_X.

Alteration of the P_X −10 element from CGAC to CGTC, a single base change, allows in vitro recognition by the $E\sigma^W$ holoenzyme, but the level of transcription is still less than that achieved with $E\sigma^X$. One additional base change, to generate a −10 region of sequence CGTA results in a promoter that can only be recognized by $E\sigma^W$. Introduction of an adjacent T residue (corresponding to position 7) (Fig. 1) results in a much stronger promoter while retaining the strong preference for the $E\sigma^W$ holoenzyme. Thus, two or three base changes in the P_X −10 element are sufficient to switch this sequence from one exclusively recognized by σ^X to one preferentially recognized by σ^W.

Similar results were found with the P_W variants. When the −10 element of P_W was changed from CGTA to CGTC, the resulting promoter could be recognized by both $E\sigma^X$ and $E\sigma^W$ with nearly equal activity. One additional base change, to generate a −10 sequence CGAC, results in a P_W variant that is preferentially recognized by $E\sigma^X$. Thus, the −10 elements of P_X and P_W determine whether the promoter is σ^X and/or σ^W dependent in vitro.

Activity of P_X variants in vivo. To test the promoter activities of the P_X and P_W variants in vivo, we cloned the promoters into an SPβ-derived prophage to generate promoter *lacZ* operon fusions. The resulting reporter fusions were transduced into the wild-type strain (CU1065), mutant strains altered in *sigX* (HB7007), *rsiX* (HB7013), *sigW* (HB0020), or *rsiW* (HB0010), or the double *sigX sigW* mutant (HB0030).

Consistent with previous work (8), P_X is active in CU1065 but not in the *sigX* mutant (Fig. 3). As expected for a σ^X-dependent promoter, activity increases in the *rsiX* anti-σ factor mutant. Unexpectedly, activity of P_X is also reduced several-fold in a *sigW* mutant. The origin of this effect is unclear, since previous analyses failed to reveal a comparable effect. Indeed, activity of most σ^X-dependent promoters is slightly elevated in *sigW* mutant strains (M. Cao and J. D. Helmann, unpublished data). The basis for this discrepancy is not clear.

The CGTC P_X variant has an in vivo profile virtually indistinguishable from P_X: activity is reduced to background levels in the *sigX* mutant strain (Fig. 3). This suggests that in vivo σ^W does not contribute significantly to expression, despite the fact that $E\sigma^W$ can recognize, albeit weakly, this promoter in vitro.

The CGTAaa P_X variant (bases flanking the four −10 consensus positions are in lowercase) did not show any measurable activity in any recipient strain. Notice that the CGTAaa P_X variant has a much weaker activity than the CGTAta P_X variant in vitro, although it is also σ^W dependent (Fig. 2). Consistent with the in vitro transcription results, the CGTAta P_X variant is no longer dependent on σ^X for expression, and instead, activity is reduced to background levels (<1 Miller unit) in the *sigW* mutant. As expected for a σ^W-dependent promoter, activity is elevated slightly in an *rsiW* mutant. In addition, activity is elevated in the *sigX* mutant and decreased slightly in an *rsiX* background. The latter results are consistent with previous observations that activities of σ^W and σ^X promoters are often mutually antagonistic: increased activity of one leads to decreased activity of the other (9, 10). The basis for this effect is not yet understood. Thus, despite the fact that the in vivo activity of the P_X CGTA variant is low (~5 Miller units), it has all the hallmarks of a σ^W-dependent promoter sequence.

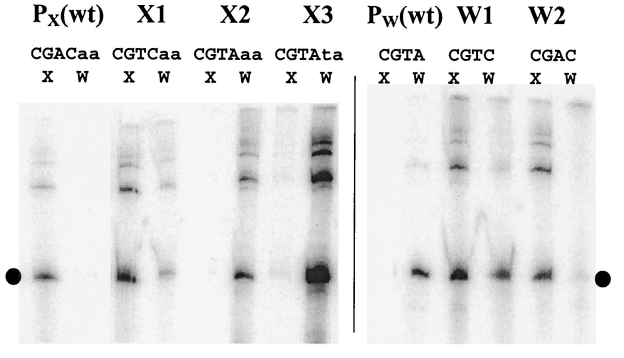

FIG. 2. In vitro transcription of P_X, P_W, and their variants by $E\sigma^X$ and $E\sigma^W$ holoenzymes. The −10 elements of the promoters are indicated. Downstream flanking bases are shown for the σ^X −10 element, since one base was altered between the X2 and X3 variants. The dot indicates the expected runoff products. wt, wild type.

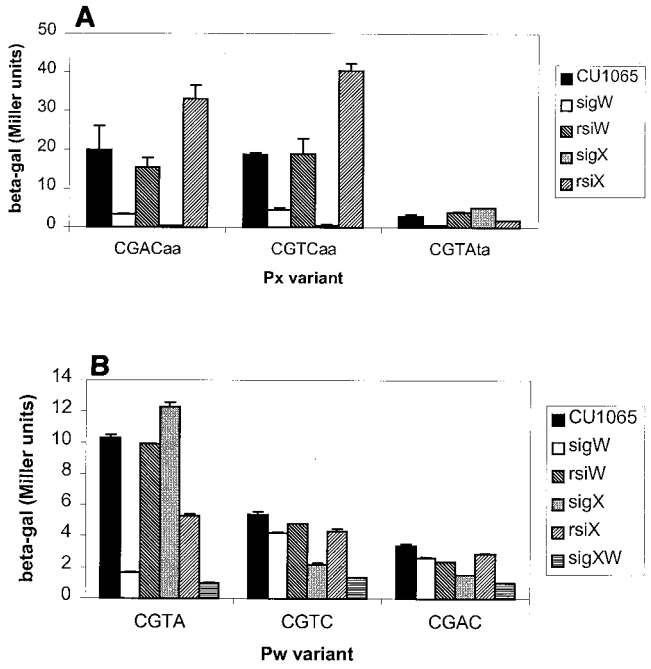

FIG. 3. In vivo expression driven by P_X, P_W, and their variants. β-Galactosidase activities (beta-gal) were measured for strains carrying the indicated promoter variant. (A) Results for wild-type P_X (CGACaa) and its −10 region variants. Note that the CGTAta variant used (X3) was altered at three positions (underlined), since the X2 variant (Table 3) did not have detectable activity in vivo. (B) Results for wild-type P_W (CGTA) and its −10 region variants. Error bars represent the standard deviations from at least two assays.

Activity of P_W variants in vivo. As expected, P_W directs the synthesis of β-galactosidase in the wild type (CU1065), but activity is completely lost in the *sigW* mutant and actually increases slightly in the *sigX* mutant strain. With the P_W variant with a CGTC sequence in the −10 region, mutation of *sigW* has only a small effect on activity while a *sigX* mutation leads to a significant decrease in activity. In the double *sigX sigW* mutant, activity is reduced further still, to near background. Thus, this single base change has converted a σ^W-dependent promoter to a one primarily dependent on σ^X in vivo (Fig. 3), consistent with in vitro transcription results (Fig. 2). Similar results are seen with the P_W variant with a CGAC −10 region.

Taken together, the in vitro and in vivo expression studies demonstrate that sequence changes localized to the −10 region can convert a σ^X-dependent to a σ^W-dependent promoter and vice versa. However, in each case the total promoter activity is significantly reduced. Thus, there are likely to be other sequence features within the promoter region that, while not essential for recognition, nevertheless help optimize the promoter for activity with the cognate holoenzyme. In addition, the in vivo activity of these promoters may be affected by regulatory proteins not present in the purified in vitro system.

Promoter specificity of σ^W region 2 mutants. Numerous previous studies indicate that region 2 of σ factors interacts with the −10 element of promoter DNA (see references 3, 5, and 22 for reviews). Sequence comparisons of σ^X, σ^W, and other σ factors identify two amino acids that are candidates for contacting the −10 region consensus element. Wilson modeled this region of ECF σ factors as an alpha helix, based on the three-dimensional structure of this domain in *E. coli* σ^{70} (28), and noted that there is a correlation between the identity of surface-exposed amino acids and the −10 element sequences recognized by the corresponding holoenzymes (Fig. 4) (28). Specifically, she hypothesized that His64 and Ser60 of σ^X recognize the critical AC in the −10 element of P_X, while Asn79 and Arg75 of σ^W recognize the TA in the −10 element of P_W. To test this model, we mutated the *sigW* gene to allow expression of σ^W variants having either one or both of these candidate selectivity determinants replaced with the corresponding residues from σ^X. The R75S and R75S N79H mutants were expressed in *E. coli* and purified using procedures similar to that used for wild-type σ^W (9).

We used in vitro transcription assays to determine if the σ^W mutations affected promoter recognition. Like σ^W, both the R75S and the R75S N79H σ^W mutant proteins recognize P_W, but not P_X, in vitro. Neither mutant holoenzyme is able to

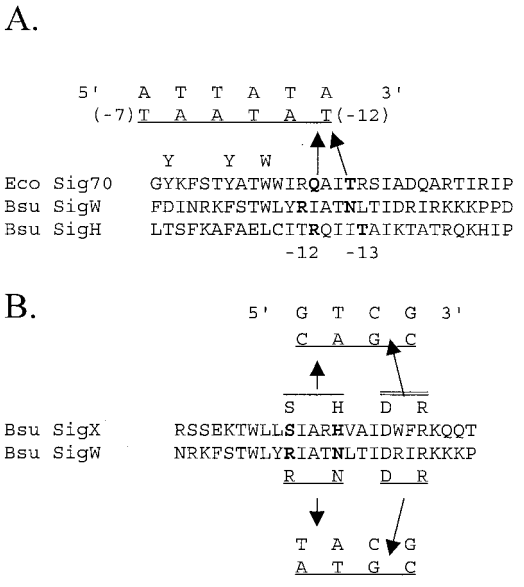

FIG. 4. Alignment of the −10 recognition domain (region 2) of σ factors and a model for promoter recognition. (A) The portion of conserved σ region 2 corresponding to the DNA −10 region recognition helix (helix 14 in the crystal structure of a fragment of σ^{70}) (18) is shown. In *E. coli* σ^{70}, residues Q437 and T440 have been implicated in recognition of the T residue at the −12 position of the −10 region consensus (25, 27). Note that the −10 region is written in inverted orientation (the transcriptional start site would be to the left and the −35 region to the right). Genetic experiments with *B. subtilis* σ^H also support a role for this region in −10 region recognition and contribute to a model orienting the recognition helix with its amino-terminal end at the downstream end of the −10 region (2). The amino-terminal end of this helix contains multiple aromatic amino acids (two Y and two W) that form a single-stranded DNA-binding surface that makes contributions both to promoter melting and to −10 region recognition (reviewed in reference 6). The corresponding region of σ^W is shown based on a multiple-sequence alignment of ECF σ family members against other σ^{70} family members (16). N79 of σ^W is a candidate recognition residue that aligns with σ^{70} Q437, while R75 aligns with the conserved W residues in σ^{70}. (B) Modeling region 2 of σ^W and σ^X as an α helix reveals four amino acids that might contribute to −10 region recognition as shown (28). In each case, the conserved R and D residues would recognize the upstream CG dinucleotide common to both σ^X and σ^W promoters while the two amino-terminal residues would distinguish between the downstream portion of the −10 element.

recognize the CGAC P_W variant. Thus, substitution of either or both of these amino acids failed to confer a σ^X-like selectivity on the resulting holoenzyme. Both mutants, like wild-type σ^W, can recognize the CGTC P_W variant. Reasoning that perhaps transcriptional selectivity had been altered only marginally by these amino acid changes, we set up a mixed-template transcription system containing both the wild-type P_W and the CGTC variant (see Materials and Methods). The molar ratios of transcripts from P_W to those from the CGTC P_W variant were 7.2, 9.9, and 9.8 for the wild-type σ^W, the R75S mutant, and the R75S N79H mutant, respectively. Thus, these mutations did not allow σ^W to recognize the CGTC −10 region better than the wild type, as might have been expected if the mutant proteins had a selectivity more similar to that of σ^X.

DISCUSSION

We have used in vitro transcription and β-galactosidase assays to investigate the role of the P_X and P_W −10 elements in recognition by $E\sigma^X$ and $E\sigma^W$. Altering the −10 element of P_X from CGAC to CGTA makes the promoter σ^W dependent, while changing the −10 element of P_W from CGTA to either CGAC or CGTC makes the promoter recognizable by $E\sigma^X$. This indicates that the −10 region is the major sequence element that distinguishes a σ^X-dependent from a σ^W-dependent promoter. In contrast, the −35 elements of promoters under ECF σ control often have similar sequence features, including a conserved AAC trinucleotide motif (21). Comparisons of known σ^W- and σ^X-dependent promoters are consistent with this model for the −10 region as a key selectivity determinant and have failed to reveal any plausible discriminatory sequences in the −35 region.

Several lines of evidence suggest that other sequence elements also function in promoter recognition and also play a major role in determining promoter strength. For example, the *B. subtilis* genome contains 27 perfect matches to $TGAAACN_{16}CGTA$, including at least 16 active, σ^W-dependent promoters. However, the remaining 11 sites are not positioned upstream of genes and are thus unlikely to be active promoters (10). Additional sequence elements are postulated to help distinguish the 16 active σ^W-dependent promoters from the other 11 sites with identical −35 and −10 elements. Previously, we suggested that these additional sequence features might include upstream promoter elements between −40 and −70, a pyrimidine-rich region just downstream of the −35 element (Fig. 1), and additional sequence determinants near the −10 element (10). The complexity of promoter recognition is also apparent from the observation that mutating the −10 element of P_W from CGTA to CGAC results in a σ^X-dependent promoter, but one much less active than P_X (at least in vivo). Conversely, mutating the −10 element of P_X from CGAC to CGTA results in a weak σ^W-dependent promoter. Thus, while the −10 element plays a dominant role in promoter selectivity, overall activity of the promoter is strongly influenced by context. The origins of these context effects on promoter strength are not yet clear. It is possible that regulatory proteins present in vivo, but lacking in our in vitro system, also play an important role in governing both promoter strength and selectivity.

In previous studies, we found that promoters with a CGTC −10 element could be transcribed by either the $E\sigma^X$ or $E\sigma^W$ holoenzyme, but with variable efficiency (9). In the present study, we found that in the context of P_X, the CGTC variant is much more active with $E\sigma^X$ than with $E\sigma^W$ and in vivo activity is completely eliminated in a *sigX* mutant. In contrast, when the CGTC −10 element occurs in the context of P_W, the resulting promoter is recognized about equally by the two holoenzymes in vitro and both appear to contribute to in vivo expression. Thus, it is difficult to predict which holoenzyme will play a dominant role in expression of candidate promoters with a CGTC −10 element, although in most cases it appears to be σ^X. Two promoters of this class, both dependent in vivo on $E\sigma^X$, are found upstream of the *dltABCD* and the *pssA* operons and control expression of genes involved in modification of teichoic acids and phospholipid biosynthesis, respectively (M. Cao, J. Qiu, and J. D. Helmann, unpublished data). Other examples include promoters preceding the *abh, divIC, ywbN,* and *yrhH* genes (9) and a promoter identified by Petersohn et al. upstream of the *yjbC* gene (10, 11, 24).

While our studies lend strong support to our model for DNA determinants that distinguish σ^X- from σ^W-dependent promoters, we were not successful in identifying the corresponding amino acid determinants in σ factor. A model, based on alignment of σ factor sequences (16) and analysis of mutations known to affect −10 region recognition, was developed in which two residues in an α-helical portion of conserved region 2 were proposed to contact the −10 element (28). However, replacement of these residues in σ^W with the corresponding residues from σ^X did not alter promoter selectivity in vitro. This may indicate that these residues play no role in promoter recognition. Alternatively, these residues may recognize the conserved portions of the −10 element, such as the initial CG dinucleotide, and other residues, not tested in this study, may discriminate between the CGAC (σ^X-specific) and CGTA (σ^W-specific) −10 elements.

Understanding the promoter selectivity of σ factors is essential for the use of consensus-directed approaches to defining σ factor regulons, which in turn provides important insights into biological function (10, 11, 24). One limitation of the consensus-directed methods is that they tend to identify those promoters, and only those promoters, that closely match a predefined consensus. This limitation can be circumvented, however, through the use of DNA arrays to define σ factor regulons. Indeed, knowledge of σ factor regulons will be key to the interpretation of genome-scale transcriptional profiling experiments in general, which often reveal the activation and repression of multiple regulons in response to stress conditions or environmental stimuli.

ACKNOWLEDGMENTS

We thank Megan Wilson and Iain Lamont for communication of unpublished results.

This work was supported by grant GM47446 from the NIH.

REFERENCES

1. **Cutting, S. M., and P. B. VanderHorn.** 1990. Genetic analysis, p. 27–74. *In* C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for bacillus. John Wiley and Sons, Ltd., Chichester, United Kingdom.
2. **Daniels, D., P. Zuber, and R. Losick.** 1990. Two amino acids in an RNA polymerase sigma factor involved in the recognition of adjacent base pairs in the −10 region of a cognate promoter. Proc. Natl. Acad. Sci. USA **87:**8075–8079.

3. **Gross, C. A., M. Lonetto, and R. Losick.** 1992. Bacterial sigma factors, p. 129–176. *In* S. L. McKnight and K. R. Yamamoto (ed.), Transcriptional regulation, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
4. **Haldenwang, W. G.** 1995. The sigma factors of *Bacillus subtilis*. Microbiol. Rev. **59:**1–30.
5. **Helmann, J. D.** 1994. Bacterial sigma factors, p. 1–17. *In* R. C. Conaway and J. Conaway (ed.), Transcription: mechanisms and regulation, vol. 3. Raven Press, New York, N.Y.
6. **Helmann, J. D., and P. L. deHaseth.** 1999. Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. Biochemistry **38:**5959–5967.
7. **Horsburgh, J. M., and A. Moir.** 1999. Sigma M, an ECF RNA polymerase sigma factor of *Bacillus subtilis* 168, is essential for growth and survival in high concentrations of salt. Mol. Microbiol. **32:**41–50.
8. **Huang, X., A. Decatur, A. Sorokin, and J. D. Helmann.** 1997. The *Bacillus subtilis* σ^X protein is an extracytoplasmic function sigma factor contributing to the survival of high temperature stress. J. Bacteriol. **179:**2915–2921.
9. **Huang, X., K. L. Fredrick, and J. D. Helmann.** 1998. Promoter recognition by *Bacillus subtilis* σ^W: autoregulation and partial overlap with the σ^X regulon. J. Bacteriol. **180:**3765–3770.
10. **Huang, X., A. Gaballa, M. Cao, and J. D. Helmann.** 1999. Identification of target promoters for the *Bacillus subtilis* extracytoplasmic function σ factor, σ^W. Mol. Microbiol. **31:**361–371.
11. **Huang, X., and J. D. Helmann.** 1998. Identification of target promoters for the *Bacillus subtilis* σ^X factor using a consensus-directed search. J. Mol. Biol. **279:**165–173.
12. **Hughes, J. D., P. W. Estep, S. Tavazoie, and G. M. Church.** 2000. Computational identification of cis-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae*. J. Mol. Biol. **296:** 1205–1214.
13. **Huynen, M., B. Snel, W. Lathe III, and P. Bork.** 2000. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. Genome Res. **10:**1204–1210.
14. **Juang, Y. L., and J. D. Helmann.** 1994. The delta subunit of *Bacillus subtilis* RNA polymerase. An allosteric effector of the initiation and core-recycling phases of transcription. J. Mol. Biol. **239:**1–14.
15. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin, et al.** 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. Nature **390:**249–256.
16. **Lonetto, A. M., K. L. Brown, K. E. Rudd, and M. J. Buttner.** 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial s factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. USA **91:**7573–7577.
17. **Lopez de Saro, F. J., A. Y. Woody, and J. D. Helmann.** 1995. Structural analysis of the *Bacillus subtilis* delta factor: a protein polyanion which displaces RNA from RNA polymerase. J. Mol. Biol. **252:**189–202.
18. **Malhotra, A., E. Severinova, and S. A. Darst.** 1996. Crystal structure of a σ^{70} subunit fragment from *E. coli* RNA polymerase. Cell **87:**127–136.
19. **McGuire, A. M., J. D. Hughes, and G. M. Church.** 2000. Conservation of DNA regulatory motifs and discovery of new motifs in microbial genomes. Genome Res. **10:**744–757.
20. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
21. **Missiakas, D., and S. Raina.** 1998. The extracytoplasmic function sigma factors: role and regulation. Mol. Microbiol. **28:**1059–1066.
22. **Moran, C. P., Jr.** 1993. RNA polymerase and sigma factors, p. 653–667. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria. ASM Press, Washington, D.C.
23. **Overbeek, R., M. Fonstein, M. D'Souza, G. D. Pusch, and N. Maltsev.** 1999. The use of gene clusters to infer functional coupling. Proc. Natl. Acad. Sci. USA **96:**2896–2901.
24. **Petersohn, A., J. Bernhardt, U. Gerth, D. Hoper, T. Koburger, U. Volker, and M. Hecker.** 1999. Identification of σ^B-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligonucleotide hybridization. J. Bacteriol. **181:**5718–5724.
25. **Siegele, D. A., J. C. Hu, W. A. Walter, and C. A. Gross.** 1989. Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. **206:**591–603.
26. **Slack, F. J., J. P. Mueller, and A. L. Sonenshein.** 1993. Mutations that relieve nutritional repression of the *Bacillus subtilis* dipeptide permease operon. J. Bacteriol. **175:**4605–4614.
27. **Waldburger, C., T. Gardella, R. Wong, and M. M. Susskind.** 1990. Changes in conserved region 2 of *Escherichia coli* sigma 70 affecting promoter recognition. J. Mol. Biol. **215:**267–276.
28. **Wilson, M. J.** 2000. Ph.D. thesis. University of Otago, Dunedin, New Zealand.
29. **Zuber, P., and R. Losick.** 1987. Role of AbrB and SpoOA- and SpoOB-dependent utilization of a sporulation promoter in *Bacillus subtilis*. J. Bacteriol. **169:**2223–2230.