Genetic and Biochemical Analysis of CodY-Binding Sites in Bacillus subtilis

Boris R. Belitsky* and Abraham L. Sonenshein

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 9 November 2007/Accepted 6 December 2007

CodY is a global transcriptional regulator that is known to control directly the expression of at least two dozen operons in Bacillus subtilis, but the rules that govern the binding of CodY to its target DNA have been unclear. Using DNase I footprinting experiments, we identified CodY-binding sites upstream of the B. subtilis ylmA and yurP genes. The protected regions overlapped versions of a previously proposed CodY-binding consensus motif, AATTTTCWGAAAATT. Multiple single mutations were introduced into the CodY-binding sites of the ylmA, yurP, dppA, and ilvB genes. The mutations affected both the affinity of CodY for its binding sites in vitro and the expression in vivo of lacZ fusions that carry these mutations in their promoter regions. Our results show that versions of the AATTTTCWGAAAATT motif, first identified for Lactococcus lactis CodY, with up to five mismatches play an important role in the interaction of B. subtilis CodY with DNA.

CodY is a global transcriptional regulator in Bacillus subtilis that controls, directly or indirectly, the expression of several hundred genes. CodY acts predominantly as a repressor of transcription, though a number of genes are positively regulated by CodY. Many of the CodY-regulated genes are involved in nitrogen or carbon metabolism and in adaptation to nutritional conditions unfavorable for growth (43, 58, 60).

The activity of CodY is increased by interaction with two types of effectors, branched-chain amino acids (isoleucine, leucine, and valine [ILV]) (11, 21, 46, 54) and GTP (24, 43, 47, 54). The pools of these molecules apparently reflect the nutritional status of the cells (32, 33, 59), allowing the bacteria to change the pattern of CodY-dependent gene expression in response to nutrient limitation. As a result, expression of CodY-repressed genes generally correlates inversely with the growth rate and is low in media enriched with amino acids that include ILV, is highly elevated if ILV is omitted or exhausted, and is further derepressed in the absence of other amino acids (2, 17, 54, 57), when the pools of all CodY effectors are presumably low. Glucose also contributes to the enhancement of CodY activity (17, 57), apparently by increasing the pools of CodY effectors, probably through the positive effect of carbon-catabolite protein CcpA on ILV synthesis (55, 56).

CodY homologs occur in most low-G+C gram-positive bacteria and have been shown to play a similar global regulatory role in the cases studied to date (6, 12, 21, 22, 28, 36, 37). In several pathogenic bacteria, CodY regulates functions associated with their virulence (6, 13, 36, 37, 61, 63). B. subtilis CodY was the first member of the group to be discovered (58), remains the best-studied protein of the family, and is the only member whose three-dimensional structure is known (31) (A. J. Wilkinson, personal communication).

B. subtilis CodY is a dimeric DNA-binding protein that uses a winged helix-turn-helix motif to interact with DNA (26, 31, 52). The CodY effectors, ILV and GTP, act additively to increase the affinity of CodY for DNA (24, 54). Several binding sites for B. subtilis CodY have been characterized by DNase I footprinting (7, 26, 51–54). In all cases, CodY interacts with an A+T-rich DNA sequence, but CodY-binding sites display low sequence similarity, and the general rules governing B. subtilis CodY binding to DNA have remained elusive. It has been proposed that CodY may recognize some distinct topological feature of DNA structure rather than a specific DNA sequence (52). Recently, however, two groups working with Lactococcus lactis suggested a consensus 15-bp motif, AATTTTGWGGAAAATT or AATTTTGNGAAAATT, for CodY binding in L. lactis and other bacteria (12, 22). The validity of this consensus in L. lactis cells was confirmed by the identification of several binding sites by footprinting (12) and by showing that placement of a perfect consensus sequence upstream of a CodY-independent gene brought that gene under CodY control (12). Mutations in several other putative CodY-binding motifs had an effect on the mobility of corresponding DNA fragments in gel shift experiments (22). However, only in the case of the L. lactis oppD gene has the role of the natural CodY-binding motif been confirmed by mutational analysis and expression analysis in vivo (11). The predictive power of the consensus sequence has been difficult to assess because both in L. lactis and especially in B. subtilis, regulatory regions of many CodY-controlled genes do not seem to possess a sequence that is very similar to the proposed consensus (12, 22). On the other hand, the proposed consensus motif is A+T rich and variants of the sequence can be frequently found in the regulatory regions of bacterial genes that appear not to be regulated by CodY. Our analysis of the B. subtilis genome (29) showed that more than 200 regulatory regions, many more than the apparent number of direct CodY targets (43), contain this sequence if three mismatches are allowed and that the sequence appears more than 1,000 times if four mismatches...
are allowed. At the same time, the *B. subtilis* genome does not contain even one copy of the proposed consensus sequence without mismatches.

Despite high sequence conservation in the DNA-binding region, *L. lactis* CodY is functionally different from *B. subtilis* CodY in that it does not interact with GTP as an effector and appears to be primarily dedicated to the regulation of nitrogen metabolism genes (11, 12, 22, 46). Thus, we sought to test rigorously the role of the proposed CodY-binding motif in CodY-DNA binding regulation.*B. subtilis*. Our results showed that various versions of this sequence, one of which has five mismatches, do serve as efficient CodY-binding sites for all four of the CodY-regulated genes that we have tested. Moreover, by combining the knowledge about the presence of CodY-binding motifs with data from expression and chromatin immunoprecipitation (ChIP)-to-chip analyses, it is possible to generate more accurate predictions of direct CodY targets.

**MATERIALS AND METHODS**

**Bacterial strains and culture media.** The *B. subtilis* strains used in this study are described in Table 1 and in the text and were grown at 37°C in DS nutrient broth medium or in TSS minimal medium (18) with 0.5% glucose as the carbon source and 0.2% NH₄Cl as the nitrogen source. The same media with the addition of agar were used for growth of bacteria on plates. The TSS medium was supplemented as indicated with a mixture of 16 amino acids (aa) (2) or a mixture of 13 aa, i.e., the same amino acids minus isoleucine, leucine, and valine (when present, the concentration of these 3 aa was 200 µg/ml each). The mixture of 16 aa contained all of the amino acids commonly found in proteins, except for glutamine, asparagine, histidine, and tyrosine. Escherichia coli strain JM101 (67) and its pcosB90 cod::Tn10 derivative (34) was used for isolation of plasmids and was grown in LB medium (41). The antibiotics used, when appropriate, were chloramphenicol, 2.5 µg/ml; tetracycline, 10 to 15 µg/ml; spectinomycin, 50 µg/ml; a combination of erythromycin, 0.5 µg/ml, and lincomycin, 12.5 µg/ml, for *B. subtilis* strains; and ampicillin, 50 µg/ml, for *E. coli* strains.

**DNA manipulations and transformation.** Methods for plasmid isolation, agarose and polyacrylamide gel electrophoresis, use of restriction and DNA modification enzymes, DNA ligation, PCR, and electroporation of *E. coli* cells were as described by Sambrook et al. (48). For the oligonucleotides used in this work, see Table S1 in the supplemental material. Isolation of chromosomal DNA and transformation of *B. subtilis* by chromosomal or plasmid DNA was as previously described (5). All cloned PCR-generated fragments were verified by sequencing.

**Construction of transcriptional lacZ fusions.** To create a *yurP* transcriptional fusion, the 0.31-kb PCR product containing the entire intergenic region upstream of the *yurP* gene was synthesized by using chromosomal DNA of *B. subtilis* strain SMY as the template and, respectively, and cloned between the XbaI and HindIII sites of pMW22 to create PMW22 (3) in the HindIII restriction site of the polylinker had been removed by blunt ending. HpaII contains a polylinker with XbaI, BamHI, EcoRI, XbaI, BglII, XbaI, BglII, and HindIII sites available for cloning and creating transcriptional fusions to the *E. coli lacZ* gene. A derivative of the polylinker was created as described above for pBM1487. The modified 0.31-kb PCR product containing the 3′- and 5′-coding sequences of the *yurP* gene preceded by the ribosomal binding site of the *B. subtilis spoVG* gene (H.-J. Kim, personal communication).

**Mutations in the CodY-binding sites.** Plasmids pBB1485 (yurP-yurP-lacZ) and pBB1494 (yurP-yurP-lacZ) were constructed as described above for pBB1487, by cloning the 0.25-kb yurP PCR product created with pBB135 and pBB336 and the 0.23-kb dppA PCR product created with pBB301 and pBB302. Plasmids pBB1501 (yurP-DppA-lacZ) and pBB1502 (yurP-DppA-lacZ) were created as described previously for pRS10 (54), by using pRS5 or chromosomal DNA of strain SMY as the template, respectively.

**B. subtilis strains carrying lacZ fusions at the *yurP* locus (Table 1) were isolated after transforming strain BB2511 (amyE-lacA) with appropriate plasmids, selecting for resistance to erythromycin conferred by the plasmids, and screening for loss of the spectinomycin resistance marker, which indicated a double-crossover homologous recombination event. Strain BB2511 and all of its derivatives have very low endogenous β-galactosidase activity due to a null mutation in the *lacA* gene (9).**

---

Table 1. *B. subtilis* strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMY</td>
<td>Wild type</td>
<td>P. Schaeffer</td>
</tr>
<tr>
<td>JH12586</td>
<td>ΔarbB::cat trpC2 phe-1</td>
<td>45</td>
</tr>
<tr>
<td>PJJ84</td>
<td>ΔamyE::F(lycB ΔTsp-lacZ) erm trpC2</td>
<td>26</td>
</tr>
<tr>
<td>PS251</td>
<td>codY::(erm::Spe) trpC2</td>
<td>P. Serror</td>
</tr>
<tr>
<td>SG82</td>
<td>lacA::tet trpC2</td>
<td>9</td>
</tr>
<tr>
<td>BB1888</td>
<td>lacA::tet</td>
<td>SMY plus SG82 DNA</td>
</tr>
<tr>
<td>BB2511</td>
<td>ΔamyE::spe lacA::tet</td>
<td>BB1888/pDG1730;23</td>
</tr>
<tr>
<td>BB2675</td>
<td>ΔamyE::F(yurP::lacZ) lacA::tet</td>
<td>BB2511/pBB1468</td>
</tr>
<tr>
<td>BB2676</td>
<td>ΔamyE::F(dppA::lacZ) lacA::tet</td>
<td>BB2511/pBB1469</td>
</tr>
<tr>
<td>BB2677</td>
<td>ΔamyE::F(dppA::lacA) lacA::tet</td>
<td>BB2511/pBB1470</td>
</tr>
<tr>
<td>BB2696</td>
<td>ΔamyE::F(dppA54::lacZ) lacA::tet</td>
<td>BB2511/pBB1478</td>
</tr>
<tr>
<td>BB2697</td>
<td>ΔamyE::F(dppA55::lacZ) lacA::tet</td>
<td>BB2511/pBB1479</td>
</tr>
<tr>
<td>BB2698</td>
<td>ΔamyE::F(yurP1::lacZ) lacA::tet</td>
<td>BB2511/pBB1480</td>
</tr>
<tr>
<td>BB2700</td>
<td>ΔamyE::F(yurP3::lacZ) lacA::tet</td>
<td>BB2511/pBB1482</td>
</tr>
<tr>
<td>BB2716</td>
<td>ΔamyE::F(yurP2::lacZ) lacA::tet</td>
<td>BB2511/pBB1483</td>
</tr>
<tr>
<td>BB2718</td>
<td>ΔamyE::F(yurP::lacZ) lacA::tet</td>
<td>BB2511/pBB1485</td>
</tr>
<tr>
<td>BB2719</td>
<td>ΔamyE::F(lycB ΔTsp::lacZ) lacA::tet</td>
<td>BB2511/pBB1486</td>
</tr>
<tr>
<td>BB2720</td>
<td>ΔamyE::F(lycB ΔTsp2::lacZ) lacA::tet</td>
<td>BB2511/pBB1487</td>
</tr>
<tr>
<td>BB2729</td>
<td>ΔamyE::F(yurP1::lacZ) lacA::tet</td>
<td>BB2511/pBB1493</td>
</tr>
<tr>
<td>BB2730</td>
<td>ΔamyE::F(yurP2::lacZ) lacA::tet</td>
<td>BB2511/pBB1494</td>
</tr>
<tr>
<td>BB2736</td>
<td>ΔamyE::F(dppA50::lacZ) lacA::tet</td>
<td>BB2511/pBB1495</td>
</tr>
<tr>
<td>BB2753</td>
<td>ΔamyE::F(lycB ΔTsp::lacZ) lacA::tet</td>
<td>BB2511/pBB1501</td>
</tr>
<tr>
<td>BB2759</td>
<td>ΔamyE::F(lycB ΔTsp2::lacZ) lacA::tet</td>
<td>BB2511/pBB1502</td>
</tr>
</tbody>
</table>
lapping PCR. In the first step, 0.27-kb products containing the 5’ part of the ilvB6ΔT regulatory region were synthesized by using pRPS10 (54) as the template, oligonucleotide oRPS6 as the forward primer, and mutagenic oligonucleotide oBB332 or oBB334 as the reverse primer. In a similar manner, 0.27-kb products containing the 3’ part of the ilvB6ΔT regulatory region were synthesized by using mutagenic oligonucleotide oBB331 or oBB333 as the forward primer and oligonucleotide oRPS8 as the reverse primer. The appropriate pairs of PCR products were used in the second splicing step of PCR mutagenesis as overlap- ping templates to generate modified 0.50-kb fragments containing the entire ilvB6ΔT regulatory region; oligonucleotides oRPS6 and oRPS8 served as the forward and reverse PCR primers, respectively.

To create pBB1486 (ilvB6ΔTp2-lacZ) and pBB1487 (ilvB6ΔT2p-lacZ), the corresponding spliced PCR products were digested with EcoRI and BamHI and cloned into pSK23 that had been digested with EcoRI and BglII.

**RNA isolation and primer extension.** Cells of *B. subtilis* grown in TSS glucose-ammonium medium with 16 aa until mid-exponential phase. Total RNA from a 6-m1 culture was isolated with TRIZol (Invitrogen). Primer extension experiments were performed by a modification of a previously described protocol (48) with Superscript II reverse transcriptase (Invitrogen) and two lacZ-specific oligonucleotides, oBB102 and oBB253, as primers.

**CodY** overexpression and purification. The *B. subtilis* CodY protein containing five additional histidine residues at its C terminus was purified from *E. coli* strain TK1 overproducing CodY (29). The cells were induced in L-broth cultures (optical density at 600 nm = 0.35) by addition of 0.2% L-arabinose and incubated for 4 more h. The cells were pelleted, washed in 50 mM Tris-Cl (pH 7.9)–500 mM NaCl–5 mM imidazole–5% glycerol–1 mM phenylmethylsulfonyl fluoride–0.1% Nonidet P-40. The supernatant was clarified by centrifugation, and CodY was purified to virtual homogeneity on a Ni2+ affinity column (His- Bind resin; Novagen) as described by the manufacturer using 385 mM imidazole for elution. The protein concentration was determined with the Bio-Rad protein assay reagent with bovine serum albumin as the standard.

**Labeling of DNA fragments.** The 424- and 428-bp PCR products containing the regulatory regions of the *ylmA* and *ylmAp* genes were synthesized by using vector-specific oligonucleotide oBB67 as the forward primer and oligonucleotides oBB253 and oBB102 as the reverse primers and plasmids pBB1485 (ylm-lacZ) and pBB1486 (ylmAp-lacZ) or their mutant derivatives as the templates, respectively. The reverse primer for each PCR (which would prime the synthesis of the template strand of the PCR product) was labeled with T4 polynucleotide kinase and [γ-32P]ATP. oBB67 starts 96 bp upstream of the XbaI site used for cloning, and oBB102 and oBB253 start 36 or 89 bp, respectively, downstream of the HindIII site that serves as a junction between the promoters and the lacZ regulatory region. The Patser program at http://rsat.ulb.ac.be/rsat/ was used to analyze the sequence. The proposed 15-bp CodY-binding consensus motif AATTTT appears to be 54 bp upstream of the *ylmA* gene appears to encode the 16-aa-containing medium) was about 100-fold higher in the wild-type strain BB2718 than in the *codY* null mutant strain BB2724 than in the wild-type strain BB2718 (Table 2). In the wild-type strain, the activity of the fusion was 14-fold derepressed if ILV was omitted from the 16-aa-containing medium in a strain containing a *lacZ* null mutation that eliminated endogenous *β*-galactosidase activity (9).

**RESULTS**

We tested the roles of putative CodY-binding sites similar to the proposed 15-bp CodY-binding consensus motif AATTTT CWWGAATT (12, 22) in the regulation of four different *B. subtilis* operons. All of the promoters tested have been previously shown to be highly repressed by CodY as detected by DNA microarray analysis and other approaches (35, 43, 54, 58).

**The *ylmA* promoter.** The *ylmA* gene appears to encode the ATP-binding component of an ABC-type transporter of unknown specificity. We constructed a *ylmA-lacZ* transcriptional fusion and showed that it was highly repressed by CodY: expression of the fusion in glucose-ammonium minimal medium containing ILV and a mixture of 13 other amino acids (referred to here as the 16-aa-containing medium) was about 100-fold higher in *codY* null mutant strain BB2724 than in the wild-type strain BB2718 (Table 2). In the wild-type strain, the activity of the fusion was 14-fold derepressed if ILV was omitted, i.e., in cells growing in 13-aa-containing medium, and was further increased 2.7-fold when glucose-ammonium medium was unsupplemented (Table 2). This pattern of expression is common for CodY-repressed genes (2, 17, 54, 57). The effect of ILV was previously described during adaptation to addition or deprivation of ILV (54); here we have quantitated the effects of the presence of ILV and other amino acids under steady-state growth conditions. No effect of the medium composition on *ylmA* expression was observed in the *codY* null mutant strain (Table 2).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Fusion genotype</th>
<th>Addition to medium (no. of aa)</th>
<th>β-Galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB2718</td>
<td>Wild type</td>
<td>ylmAp</td>
<td>None</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>0.22</td>
</tr>
<tr>
<td>BB2724</td>
<td><em>codY</em></td>
<td>None</td>
<td>20.2</td>
<td>94</td>
</tr>
<tr>
<td>BB2729</td>
<td>Wild type</td>
<td>ylmAp1</td>
<td>None</td>
<td>&lt;0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>0.03</td>
</tr>
<tr>
<td>BB2733</td>
<td><em>codY</em></td>
<td></td>
<td>17.7</td>
<td>100</td>
</tr>
<tr>
<td>BB2730</td>
<td>Wild type</td>
<td>ylmAp2</td>
<td>None</td>
<td>25.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>20.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>9.8</td>
</tr>
<tr>
<td>BB2734</td>
<td><em>codY</em></td>
<td></td>
<td>27.9</td>
<td>100</td>
</tr>
</tbody>
</table>

* Cells were grown in TSS glucose-ammonium medium with or without a mixture of 13 aa or 16 aa and ILV (16 aa) (see Materials and Methods). β-Galactosidase activity was assayed and expressed in Miller units (MU). All values are averages of at least two experiments, and the mean errors did not exceed 30%.

a All strains contained a *lacZ*-null mutation that eliminated endogenous *β*-galactosidase activity (9).

b *β*-Galactosidase activity of each fusion in the 16-aa-containing medium in a strain containing a *codY* null mutation was normalized to 100%.

c *β*-Galactosidase activity was similar to the background activity in a strain without a *lacZ* fusion.
FIG. 1. (A) Primer extension analysis of the ybmA mRNA. Primer oBB102 annealing to the lacZ gene of the ybmA-lacZ fusion was extended with reverse transcriptase by using as the template total RNAs from fusion-containing strains BB2718 (codY+, lane 1) and BB2724 (codY, lane 2) grown in glucose-ammonium medium containing 16 aa. The sequence of the template strand of pBB1485 determined from reactions primed with oBB102 is shown to the right. The apparent transcription start site of the ybmA gene is in bold and marked by the +1 notation. The bent arrow indicates the direction of transcription. (B) Primer extension analysis of the yurP mRNA. Primer oBB102 annealing to the lacZ gene of the yurP-lacZ fusion was extended with reverse transcriptase by using as the template total RNA from fusion-containing strains BB2675 (codY+, lane 1) and BB2687 (codY, lane 2) grown in glucose-ammonium medium containing 16 aa. The sequence of the template strand of pBB1468 determined from reactions primed with oBB102 is shown to the left. The apparent transcription start site of the yurP gene is in bold and marked by the +1 notation. The bent arrow indicates the direction of transcription.

TATCCT with two mismatches each to the −35 and −10 consensus sequences for σ^A-type promoters, respectively, and a 17-bp spacer region can be identified upstream of the transcription start point (Fig. 2D).

A DNase I footprinting experiment showed that, in the presence of ILV and GTP, CodY protected a region of the template DNA strand from position +1 to position +33 with respect to the transcription start point of ylmA and extending into the vector part of the labeled PCR product (that had been synthesized by using as the template the fusion construct used for in vivo studies) (Fig. 2A and D). As a result, the true downstream end of the protected region in the wild-type DNA context was not determined. The protected region overlaps a 15-bp sequence that has two mismatches with respect to the proposed CodY-binding consensus (Fig. 2D and Table 3). No intrinsic binding of CodY was detected in the absence of the effectors, even at a CodY concentration of 800 nM (Fig. 3). GTP alone stimulated the binding of CodY (some protection was obtained at 400 nM CodY). In the presence of ILV alone, CodY binding to the ylmA promoter was detected at 50 to 100 nM, and when both effectors were present, binding was apparent at 25 to 50 nM CodY (Fig. 2A and 3). The additive effects of GTP and ILV on the binding of *B. subtilis* CodY were previously reported for the ibxB promoter (24, 54).

We introduced separately two single substitution mutations into the putative CodY-binding motif of the ylmA-lacZ fusion (Fig. 4A). The mutations did not affect the expression of the fusion in the absence of CodY, suggesting that they had no effect on the intrinsic activity of the promoter itself (Table 2, strains BB2733 and BB2734). The ylmAp1-lacZ fusion, with an improved match of the putative CodY-binding motif with the proposed consensus, was fully repressed in codY^+ cells under all of the growth conditions tested and highly derepressed in the absence of CodY (Table 2, strains BB2729 and BB2733). The ylmAp2 mutation, which decreased the similarity of the putative CodY-binding motif to the proposed consensus, caused 45-fold derepression of the ylmA promoter (Table 2, strain BB2730). In accord with these results, the ylmAp1 mutation increased the affinity of CodY for its binding site in the ylmA regulatory region in vitro about eightfold, and the ylmAp2 mutation diminished CodY binding about fourfold (Fig. 2B). (The pattern of ylmAp2 digestion by DNase I was slightly altered due apparently to the nucleotide change introduced by the mutation.) Similar changes in the affinity of CodY for the entire ylmA regulatory region were detected in gel shift experiments (data not shown).

Despite constitutive repression of the ylmAp1 promoter in vivo, interaction of CodY with the ylmAp1 regulatory region in vitro was still responsive to the addition of ILV and GTP (compare Fig. 2B and C). It is possible that CodY binding at the ylmAp1 promoter, in contrast to binding at the ylmAp2 promoter, is activated by the low effector concentrations that are presumed to be present in cells grown in the absence of ILV or other amino acids. An alternative, but unlikely, possibility is that the small increase in the affinity of CodY for the mutant promoter in the absence of effectors (compare Fig. 2C and 3) could account for the strong repression seen in vivo.

The *yurP* promoter, *yurP* is a gene of unknown physiological significance that encodes an enzyme involved in fructosamino acid metabolism (64). It is coregulated and apparently cotranscribed with the downstream *yurONML* genes (43) that probably participate in the same biochemical pathway (64). Using a newly constructed *yurP-lacZ* transcriptional fusion, we confirmed that the fusion construct is highly repressed by CodY (43). Expression of the fusion in 16-aa-containing medium in codY null mutant strain BB2687 was more than 1,000-fold higher than in wild-type strain BB2675 (Table 4). In the wild-type strain, the activity of the fusion was derepressed 13-fold in 13-aa-containing medium, i.e., when ILV was omitted, and was further increased 4-fold in glucose-ammonium medium (Table 4).

A primer extension experiment established that the transcription start point of *yurP* appears to be 47 bp upstream of the initiation codon (Fig. 1B). The sequences TTAGCA and TAATAT, with no and two mismatches with the −35 and −10 regions of σ^A-dependent promoters, respectively, and a 17-bp
spacer region, can be identified upstream of the transcription start point (Fig. 5D). A mutation in the yurP regulatory region, yurPp4, that we isolated inadvertently turned out to be an A-to-G substitution in the /H11002 region of the promoter (TTGACA to TTGGCA). This mutation completely abolished expression from the promoter (data not shown), supporting our localization of the transcription start point.

A DNase I footprinting experiment showed that CodY protected a region of the template DNA strand stretching from position −64 to position −39 with respect to the transcription start point (Fig. 5A). The protected region overlaps a 15-bp sequence with two mismatches with respect to the proposed CodY-binding consensus (Table 3 and Fig. 5D).

We introduced separately three single substitution mutations into the putative CodY-binding motif within the yurP-lacZ fusion (Fig. 4A). One of them, yurPp1, improved the match of this region with the CodY-binding consensus motif; two other mutations, yurPp2 and yurPp3, decreased the similarity of this region to the consensus motif. The yurPp1 mutation increased by 4-fold the ability of CodY to repress yurP expression in 16-aa-containing medium, and both yurPp2 and yurPp3 caused 19- to 64-fold derepression of the promoter.
CodY-regulated genes

Some other CodY-regulated genes

<table>
<thead>
<tr>
<th>Promoter region</th>
<th>Putative CodY-binding site (no. of mismatches)</th>
<th>Score</th>
<th>Location</th>
<th>Verification by footprinting and/or mutations</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yurP</td>
<td>AAAATTTCAGAAATAAT (2)</td>
<td>10.6</td>
<td>−53 to −39</td>
<td>Yes</td>
<td>This work</td>
</tr>
<tr>
<td>yhGD box 1</td>
<td>AtTTTTTCTAAGCTT (3)</td>
<td>8.0</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhGD box 2</td>
<td>tgaATCTGAAATAAT (4)</td>
<td>8.1</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yxbB</td>
<td>ATTTaTCAAGgATTT (3)</td>
<td>8.7</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhC</td>
<td>aTaaTTTCAGACATTT (3)</td>
<td>9.4</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dppA box 1</td>
<td>tATTTATATAtATTT (5)</td>
<td>4.1</td>
<td>−19 to −5</td>
<td>Yes</td>
<td>52</td>
</tr>
<tr>
<td>dppA box 2</td>
<td>tATTTTATATATTT (4)</td>
<td>7.8</td>
<td>−7 to +8</td>
<td>Yes</td>
<td>52</td>
</tr>
<tr>
<td>dppA box 3</td>
<td>AATTTTATATATT (3)</td>
<td>5.6</td>
<td>+3 to +17</td>
<td>Yes</td>
<td>52</td>
</tr>
<tr>
<td>dppA box 4</td>
<td>AATTTTATATATT (5)</td>
<td>1.4</td>
<td>+12 to +26</td>
<td>Yes</td>
<td>52</td>
</tr>
<tr>
<td>yuN box 1</td>
<td>AATTTTATATATT (2)</td>
<td>10.1</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yuN box 2</td>
<td>AATTTTATATATT (2)</td>
<td>12.0</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ivB</td>
<td>AATTTTATATATT (3)</td>
<td>8.9</td>
<td>−65 to −51</td>
<td>Yes</td>
<td>54</td>
</tr>
<tr>
<td>ybcC</td>
<td>AATTTTATATATT (3)</td>
<td>4.7</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rocA</td>
<td>AATTTTATATATT (2)</td>
<td>&lt;2.0f</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>appD</td>
<td>AATTTTATATATT (3)</td>
<td>8.1</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ykwB box 1</td>
<td>AATTTTATATATT (3)</td>
<td>7.9</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ykwB box 2</td>
<td>AATTTTATATATT (2)</td>
<td>11.1</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ywAB</td>
<td>AATTTTATATATT (2)</td>
<td>12.0</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ivH</td>
<td>AATTTTATATATT (3)</td>
<td>9.6</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ykgE</td>
<td>AATTTTATATATT (3)</td>
<td>10.3</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ylmA</td>
<td>AATTTTATATATT (2)</td>
<td>9.9</td>
<td>+3 to +17</td>
<td>Yes</td>
<td>This work</td>
</tr>
</tbody>
</table>

Some other CodY-regulated genes

| ackA            | AcTTCAAGAAATTT (4)                          | 6.9   | −90 to −76 | Yes                                         | 53           |
| citB            | ATTTTTTTCAGAAATAAT (2)                      | 10.6  | −4 to +11  | Yes                                         | 27           |
| flagBpD3       | ATTTTTTTCAGAAATAAT (3)                      | 7.7   | −58 to −44 | Yes                                         | 7            |
| hudB            | AgTTTCAAGAAATAAT (4)                        | 8.6   | +8 to +22  | Yes                                         | 14, 66       |
| acsA            | AgTTTCAAGAAATAAT (4)                        | 7.1   | −48 to −34 | 20                                         |
| gabAP1          | AgTTTCAAGAAATAAT (3)                        | 9.8   | −16 to −2  | 15                                         |
| ivA             | AAATTTCAGAAATAAT (3)                        | 9.3   | ND       | 35                                         |
| yurB            | AAATTTCAGAAATAAT (3)                        | 8.0   | +15 to +29 | 65                                         |
| yurH            | AAATTTCAGAAATAAT (3)                        | 8.5   | −50 to −36 | 65                                         |
| yurF            | AAATTTCAGAAATAAT (2)                        | 10.3  | ND       | 25                                         |

Table 4. Strains BB2675, BB2687, and BB2700. In accord with this result, the yurP promoter is derepressed in CodY repression in the absence of ILV and other amino acids (2, 17, 54, 57), presumably due to low pools of CodY effectors. However, the yurP gene is expressed in glucose-ammonium minimal medium in the wild-type strain at a level of only 7% compared to the level of yurP expression in the codY mutant (Table 4, strains BB2675 and BB2687). Apparently, the yurP regulatory region is able to interact with CodY very efficiently when CodY is relatively inactive, similar to the situation with the mutant ylmAp1 promoter. Nevertheless, CodY binding to the yurP regulatory region is still responsive to CodY effectors both in vivo and in vitro: yurP-lacZ expression was decreased 89-fold by addition of the 16-aa mixture (Table 4, strain BB2675), and no binding of CodY (even at 800 nM) was detected in the absence of the effectors (Fig. 5C). Making the CodY-binding site of the yurP gene even stronger by introduction of the yurP promoter did not abolish the dependence of yurP expression on CodY effectors (Table 4, strain BB2698).

The dppA promoter. The dppA promoter, encoding d-alanyl-aminopeptidase and a dipeptide permease (8, 40), was the first described target of CodY in B. subtilis (58). Derepression of the dppA operon apparently leads to readthrough transcription of the ykfA/BCD operon, which likely encodes...
enzymes involved in peptidoglycan recycling (43). Three mutations in the experimentally determined CodY-binding site of the \textit{dppA} promoter that disrupted regulation by CodY were previously isolated (52, 56). As pointed out previously (12, 22), two of these mutations, a single substitution mutation, \textit{dppAp50}, and a 12-bp deletion, \textit{dppAp129} (formerly \textit{dcs-50} and \textit{dcs-129}, respectively), affect two partially overlapping 15-bp sequences, \textit{dppA} box 3 and \textit{dppA} box 4 (Table 3) (Fig. 4B), that are contained within the actual CodY-binding site and have five mismatches each with respect to the proposed CodY-binding consensus motif. The third mutation, \textit{dppAp49} (formerly, \textit{dcs-49}), is a 4-bp insertion located immediately upstream of the \textit{dppA} box 3 sequence. However, two more 15-bp sequences with four or five mismatches with respect to the CodY-binding consensus site of the \textit{dppA} operon overlaps the binding site of AbrB, another global regulator in \textit{B. subtilis} that provides about twofold repression of \textit{dpp} in the absence of CodY but has little effect when CodY is active (52, 57). To exclude any AbrB contribution to the effects of the regulatory mutations analyzed above, we performed all of the in vivo experiments described above with strains containing an \textit{abrB} null mutation. However, repeating some of these experiments in an \textit{abrB}+/H11001 background revealed that the \textit{dppAp50} and \textit{dppAp51} mutations in the CodY-binding motif not only affected CodY-dependent regulation but also relieved most of AbrB-dependent regulation (Table 6).

The actual CodY-binding site of the \textit{dpp} operon overlaps the binding site of AbrB, another global regulator in \textit{B. subtilis} that provides about twofold repression of \textit{dpp} in the absence of CodY but has little effect when CodY is active (52, 57). To exclude any AbrB contribution to the effects of the regulatory mutations analyzed above, we performed all of the in vivo experiments described above with strains containing an \textit{abrB} null mutation. However, repeating some of these experiments in an \textit{abrB}+/H11001 background revealed that the \textit{dppAp50} and \textit{dppAp51} mutations in the CodY-binding motif not only affected CodY-dependent regulation but also relieved most of AbrB-dependent regulation (Table 6).

The \textit{ilvB} promoter. The \textit{ilvB} promoter of the \textit{ilvBHC leuABCD} operon is one of the best-studied targets of CodY (35, 54, 55, 62). The experimentally determined high-affinity CodY-binding region (region II) of the \textit{ilvB} gene (54) partially overlaps a 15-bp sequence with three mismatches with respect
of a 4 bp-insertion within the dppAp49 allele. The 12-bp deletion affects nucleotides in the positions from T(215)C, A(+82)G, and C(+84)T]. None of these mutations affected the high-affinity CodY-binding site. To make sure that these preexisting mutations did not affect the interaction between CodY and the ilvB regulon, we constructed two fusions analogous to ilvBΔT-lacZ: the ilvBΔT-lacZ fusion that had the four mutations originating from pRPS5 and the ilvBΔT-lacZ fusion that had no mutations. Though the expression level of the ilvBΔT-lacZ fusion was about 1.7-fold higher than that of the other two fusions, regulation of all three fusions by CodY was very similar (Table 7, compare strains BB2723, BB2753, and BB2759 and their codY derivatives). Additionally, it was shown that CodY has very similar affinities for and protects identical regions of DNA in the ilvBΔT and ilvBΔT DNA probes (24).

The ilvBp1 mutation, which increased the similarity of the putative CodY-binding motif to the proposed consensus, increased the ability of CodY to repress the expression of the ilvBΔT-lacZ fusion in 16-aa-containing medium sevenfold, and the ilvBp2 mutation, which decreased the similarity of the putative CodY-binding motif to the proposed consensus, caused threefold derepression of the promoter (Table 7, strains BB2719 and BB2720).

**DISCUSSION**

A DNA microarray analysis of *B. subtilis* genes regulated by CodY revealed several hundred genes organized in dozens of operons; 67 such operons were also identified as potential direct targets of CodY in ChIP-DNA microarray (ChIP-to-chip) experiments (43). Guédon et al. (22) reported previously that among the CodY-dependent promoters identified by the latter approach, only 25 (40%) contained a DNA sequence conserved between the regions protected by CodY in DNase I footprinting experiments with the wild-type CodY (lack of the T-box terminator-antiterminator regulatory element (19, 39)). Therefore, we renamed this fusion ilvBΔT-lacZ. Furthermore, we found that plasmid pRPS5, the precursor of pRPS10 and the template for PCR products for all of the footprinting experiments performed previously (26, 54, 55), contained four of the mutations [T(215)C, A(+82)G, and C(+84)T].

### TABLE 4. Expression of yurP-lacZ fusions

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Fusion genotype</th>
<th>Addition to medium (no. of aa)</th>
<th>β-Galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB2675</td>
<td>Wild type</td>
<td>yurPp</td>
<td>None</td>
<td>13 143.0 7.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16 34.8 1.8</td>
<td></td>
</tr>
<tr>
<td>BB2687</td>
<td>codY</td>
<td>None</td>
<td>16 2,170.0 112</td>
<td></td>
</tr>
<tr>
<td>BB2698</td>
<td>Wild type</td>
<td>yurPp1</td>
<td>None</td>
<td>13 15.3 0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16 3.2 0.2</td>
<td></td>
</tr>
<tr>
<td>BB2709</td>
<td>codY</td>
<td>None</td>
<td>16 1,720.0 100</td>
<td></td>
</tr>
<tr>
<td>BB2716</td>
<td>Wild type</td>
<td>yurPp2</td>
<td>None</td>
<td>13 442.0 24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16 180.0 9.8</td>
<td></td>
</tr>
<tr>
<td>BB2717</td>
<td>codY</td>
<td>None</td>
<td>16 1,840.0 100</td>
<td></td>
</tr>
<tr>
<td>BB2700</td>
<td>Wild type</td>
<td>yurPp3</td>
<td>None</td>
<td>13 426.0 22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16 133.0 6.9</td>
<td></td>
</tr>
<tr>
<td>BB2711</td>
<td>codY</td>
<td>None</td>
<td>16 1,930.0 100</td>
<td></td>
</tr>
</tbody>
</table>

| a | Cells were grown and β-galactosidase activity (Miller units [MU]) was assayed as described for Table 2.  
| b | All strains contained a lacA null mutation that eliminated endogenous β-galactosidase activity (9).  
| c | The β-galactosidase activity of each fusion in the 16-aa-containing medium in a strain containing a codY null mutation was normalized to 100%.

FIG. 4. CodY-binding motifs and mutations. (A) Mismatches with the consensus sequence are indicated by lowercase letters; the number of mismatches is in parentheses. The motifs are shown below the sequence lanes as substituting nucleotides. The score for each substitution is shown as a multiple of the background score (24). (B) Locations of putative CodY-binding motifs in the dppA regulatory region. The –10 and –35 promoter regions and transcription start sites are in bold. dppA box I, box II, box III, and box IV are underlined. The uppercase letters correspond to the region protected by CodY in DNase I footprinting experiments with the wild-type dppA promoter (26, 53). The vertical arrow indicates the site of a 4 bp-insertion within the dppAp129 allele. The 12-bp dppAp129 deletion affects nucleotides in the positions from +5 to +16.

to the proposed CodY-binding consensus motif (Table 3). When we sought to introduce substitution mutations (Fig. 4A) into the previously constructed ilvBΔT-lacZ fusion (54) that lacks the T-box terminator-antiterminator regulatory element (19, 39), we found that the ilvBΔT fragment of the fusion contains six previously untested substitution mutations [T(–215)C, T(–159)C, T(–24)C, A(+2)T, A(+82)G, and C(+84)T] introduced during the fusion construction and originating from plasmid pRPS10 (54). Therefore, we renamed this fusion ilvBΔT-lacZ. Furthermore, we found that plasmid pRPS5, the precursor of pRPS10 and the template for PCR products for all of the footprinting experiments performed previously (26, 54, 55), contained four of the mutations [T(–215)C, A(+2)T,
differing from the proposed 15-bp CodY-binding consensus motif, AATTTTCWGAAAATT, at three or fewer positions. Our further examination of the regulatory regions of CodY-dependent genes revealed that among the 16 transcription units most highly repressed by CodY (43) (http://www.genome.jp/kegg/expression/), 14 contain in their regulatory regions sequences with only two or three mismatches with respect to the proposed consensus (Table 3). In the present work, we demonstrated by genetic and biochemical approaches that three of these putative CodY-binding motifs, within the yurP, ylmA, and ilvB promoters, play an important role in the regulation of these genes by CodY. Additionally, a sequence with five mismatches with the proposed CodY-binding consensus motif plays an equally important role in the CodY-dependent regulation of the dpp operon. (The rocA promoter, the only other one among the 16 highly regulated CodY targets that does not have a recognizable CodY-binding motif, is, according to our unpublished results, regulated by CodY indirectly.)

The 15-bp consensus-like sequences in the yurP, ylmA, and ilvB promoters overlap actual CodY-binding sites, as determined in DNase I footprinting experiments in this work and previously (26, 52, 54). All 11 single substitution mutations that reduced or improved the match of the 15-bp sequences with the proposed consensus sequence significantly decreased...
or increased, respectively, the ability of CodY to repress gene expression in vivo. In the case of the yurP and ylmA promoters, we also showed by DNase I footprinting that mutations in the 15-bp sequences changed the affinity of CodY for the corresponding DNA regions in the expected way. We conclude that, at least for B. subtilis promoters that are strongly regulated by CodY, the proposed 15-bp consensus sequence, AATTTTCW, is a motif that is strongly associated with CodY binding and regulation. In contrast to the yurP, ibv, and dppA promoters, the ylmA promoter was not detected as a direct CodY target by ChIP-to-chip analysis (43). Thus, CodY directly regulates more genes than inferred previously.

Using a set of 27 established and likely binding motifs for B. subtilis CodY, we created a position-specific weight matrix (Fig. 6) and used it to calculate scores for the motifs analyzed in this work and to more accurately search the regulatory regions of the B. subtilis genome. The scores for individual CodY-binding motifs reflect their relative similarities to the entire set of motifs used to create the matrix: the possible theoretical range of scores was +14 to −36. All of the highly regulated genes presented in Table 3, excluding dppA and

or increased, respectively, the ability of CodY to repress gene expression in vivo. In the case of the yurP and ylmA promoters, we also showed by DNase I footprinting that mutations in the 15-bp sequences changed the affinity of CodY for the corresponding DNA regions in the expected way. We conclude that, at least for B. subtilis promoters that are strongly regulated by CodY, the proposed 15-bp consensus sequence, AATTTTCW, is a motif that is strongly associated with CodY binding and regulation. In contrast to the yurP, ibv, and dppA promoters, the ylmA promoter was not detected as a direct CodY target by ChIP-to-chip analysis (43). Thus, CodY directly regulates more genes than inferred previously.

Using a set of 27 established and likely binding motifs for B. subtilis CodY, we created a position-specific weight matrix (Fig. 6) and used it to calculate scores for the motifs analyzed in this work and to more accurately search the regulatory regions of the B. subtilis genome. The scores for individual CodY-binding motifs reflect their relative similarities to the entire set of motifs used to create the matrix: the possible theoretical range of scores was +14 to −36. All of the highly regulated genes presented in Table 3, excluding dppA and

TABLE 6. Effect of dppA promoter mutations on AbrB-dependent regulation of dppA-lacZ fusions

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Fusion genotype</th>
<th>Addition to medium (no. of aa)</th>
<th>β-Galactosidase activity MU</th>
<th>%b</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB2681</td>
<td>codY</td>
<td>dppAp</td>
<td>None</td>
<td>155.0</td>
<td>41</td>
</tr>
<tr>
<td>BB2786</td>
<td>codY abrB</td>
<td>dppAp50</td>
<td>None</td>
<td>344.0</td>
<td>110</td>
</tr>
<tr>
<td>BB2787</td>
<td>codY</td>
<td>dppAp51</td>
<td>None</td>
<td>366.0</td>
<td>108</td>
</tr>
<tr>
<td>BB2788</td>
<td>codY</td>
<td>dppAp52</td>
<td>None</td>
<td>360.0</td>
<td>108</td>
</tr>
<tr>
<td>BB2789</td>
<td>codY</td>
<td>dppAp53</td>
<td>None</td>
<td>309.0</td>
<td>82</td>
</tr>
<tr>
<td>BB2790</td>
<td>codY</td>
<td>dppAp54</td>
<td>None</td>
<td>300.0</td>
<td>100</td>
</tr>
</tbody>
</table>

a All strains were grown in TSS glucose-ammonium medium with a mixture of 16 amino acids. β-Galactosidase activity (Miller units [MU]) was assayed as described for Table 2.

b All strains contained a lacA null mutation that eliminated endogenous β-galactosidase activity (9).

c The β-galactosidase activity of each fusion in a strain containing codY and abrB null mutations was normalized to 100%.

CodY, the proposed 15-bp consensus sequence, AATTTTCW GAAAATT, is a motif that is strongly associated with CodY binding and regulation. In contrast to the yurP, ibv, and dppA promoters, the ylmA promoter was not detected as a direct CodY target by ChIP-to-chip analysis (43). Thus, CodY directly regulates more genes than inferred previously.

Using a set of 27 established and likely binding motifs for B. subtilis CodY, we created a position-specific weight matrix (Fig. 6) and used it to calculate scores for the motifs analyzed in this work and to more accurately search the regulatory regions of the B. subtilis genome. The scores for individual CodY-binding motifs reflect their relative similarities to the entire set of motifs used to create the matrix: the possible theoretical range of scores was +14 to −36. All of the highly regulated genes presented in Table 3, excluding dppA and

FIG. 6. Position-specific weight matrix of CodY-binding motifs. Shown is the position-specific weight matrix for 27 established and likely B. subtilis CodY-binding motifs showing the relative frequency of occurrence of each of the four nucleotides (A, T, C, or G) at each of the 15 positions of the motif. The logo presentation of the matrix (50) was created at http://weblogo.berkeley.edu/.
rocA, had a motif with a highly significant score equal to or above 8. Only about 150 B. subtilis genes had putative CodY-binding motifs with such high scores in their regulatory regions.

The B. subtilis genome has no sequence identical to the CodY-binding consensus. It is possible that such sequences would be such strong binding sites for CodY that no regulation would be possible. That is, CodY would bind to these sites whether the pools of its effectors were high or low. Indeed, the ylmAp1 mutant promoter has a CodY-binding motif with only one mismatch with respect to the consensus (score = 12.4) and is fully repressed by CodY under all of the growth conditions tested. Interestingly, this motif is different from its counterpart containing the yurPp1 mutation and from four naturally occurring sequences in the regulatory regions of B. subtilis genes that have one mismatch with the CodY-binding consensus motif and similarly high scores (the CodY-dependent regulation of these genes has not been analyzed in detail).

Putative CodY-binding motifs with two or three mismatches with respect to the consensus and scores above 6.9 can also be found within the actual binding sites of the CodY-repressed cibB and fgbBPpD promoters (7, 27) (Table 3). Such motifs can also be found in the regulatory regions of the CodY-repressed gsbP, ureA, ilvB, ilvD, ybgE, ywFB, and ywfH genes (15, 25, 35, 65) (Table 3). The regulatory regions of the CodY-repressed hutP, rnap, rapC, bkdR, and acsA genes (10, 16, 17, 30) and the positively regulated ackA gene (53) contain putative motifs that differ from the consensus at at least four positions, and only some of these motifs have scores above 6.9 (the total number of B. subtilis genes with such scores is about 300) (Table 3 and data not shown). Among these genes, only ackA (score = 6.9) has been shown to be a direct target of CodY by DNase I footprinting experiments (53). Deletion or substitution mutations away from the consensus in the putative CodY-binding motifs of ackA and hutP significantly decreased or increased expression from the corresponding promoters, respectively (14, 42, 66).

Actual CodY-binding sites of several CodY-regulated promoters, such as comK, srA, hac, and fgbBPp1 (7, 51), do not have any 15-bp sequence that differs from the consensus at fewer than five positions and has a score above 5 (for comparison, there are more than 1,000 B. subtilis genes that are preceded by a motif with such a score; data not shown). It is still possible that such degenerate versions of the sequences in these and other CodY-dependent genes where the consensus motif cannot be confidently recognized are sufficient for interaction with CodY, as is the case for the dppA promoter.

The score of a CodY-binding motif may generally correlate with the strength of CodY binding, but if some feature of CodY-binding sites other than similarity to the consensus sequence, such as the nature of adjacent residues or local topology, contributes to the strength of CodY binding, the 15-bp sequence may be of variable importance from promoter to promoter. In fact, the scores for the wild-type and mutant promoters analyzed in this work do not correlate well with the observed efficiency of regulation mediated by CodY (Fig. 4A). However, the effects of multiple mutations making the yurP and dppA CodY-binding motifs less similar to the consensus do correlate with the decrease in their scores, confirming the predictive value of this approach (Fig. 4A). It is possible that alternative putative CodY-binding motifs in the dppA promoter contribute to the regulation provided by low-scoring dppA box 3 and increase the efficiency of dppA regulation.

The difficulty in identifying CodY-binding sites with low similarity to the consensus could partially explain why no perfect correlation was found previously and in this work between the presence of putative CodY-binding motifs and the number of genes directly regulated by CodY in L. lactis or B. subtilis (12, 22), especially considering that not all of these genes are preceded by a highly scoring CodY-binding motif (data not shown). On the other hand, many genes preceded by highly scoring CodY-binding motifs are apparently not subject to CodY-dependent regulation. In the latter cases, the motifs may be positioned improperly with respect to the corresponding promoter or be masked from interaction with CodY by some features of local DNA structure or binding of other regulatory proteins. Thus, while the presence of an identifiable version of the CodY-binding consensus motif in the regulatory region of a gene does not necessarily indicate that the gene is regulated by CodY, the presence of such a sequence, combined with the knowledge that expression of the gene is responsive to CodY, as determined by DNA microarray or lacZ fusion experiments or other approaches, seems to be a very good, albeit not perfect, predictor of direct regulation by CodY. Similarly, the reliability of ChIP-to-chip analysis is greatly improved if a recognizable CodY-binding motif can be found in the corresponding DNA fragment.

We do not know yet which features of the CodY-binding consensus motif are critical for interaction with CodY. Such knowledge would be useful for identifying poorly conserved CodY-binding motifs and eliminating higher-scoring motifs that are unable to interact efficiently with CodY. In this work, we mutated the 15-bp sequence at six different positions; every mutation had a significant effect on gene expression, though most of these positions are not more conserved than others (Table 3 and data not shown). Our results, as well as the previously described mutation in the putative CodY-binding motif of hutP (14), might suggest that the central C and G nucleotides of the consensus motif are very important for CodY-dependent regulation. However, one of these nucleotides is frequently missing from actual CodY-binding sites and its absence does not prevent efficient regulation of the corresponding genes by CodY; some of the putative motifs lack both of these nucleotides (Table 3). Alteration of the completely conserved adenine nucleotide (Table 3) by the dppAp50 mutation almost wiped out CodY repression. The absence of information on positively regulated direct targets of CodY, other than ackA, does not allow us to approach the question of whether CodY-binding sites of such genes have any features that distinguish them from CodY-binding sites of the negatively regulated genes.

The molecular mechanism of CodY-dependent repression or activation of transcription remains unknown. The CodY-binding motifs of CodY-repressed genes lie almost immediately upstream of the −35 promoter region or almost immediately downstream of the transcription start point or overlap the transcription start point (Table 3). In all of these cases, it is easy to imagine how CodY bound so close to the promoter can interfere with important steps of transcription initiation or serve as a roadblock for transcription elongation.
Only one binding site for CodY was detected experimentally in the regulatory regions of the ilvBA and ywrP genes. In contrast, multiple binding sites with different affinities for CodY have been previously detected in the ilvB and dppA regulatory regions (26, 54). The ilvB and dppA sites characterized in this work correspond to high-affinity CodY-binding sites detected previously. No highly scoring CodY-binding motif can be found in the low-affinity binding sites within the ilvB and dppA regulatory regions. By deletion analysis, no physiological role for the upstream low-affinity CodY binding sites in the expression of the ilvB promoter was detected (62).

ACKNOWLEDGMENTS

We are grateful to L. D. Handke and P. Jern for helpful discussions and to P. Serror and H.-J. Kim for the gift of strains. This work was supported by a research grant (GM042219) from the U.S. Public Health Service.

REFERENCES


Downloaded from http://jb.asm.org/ on September 12, 2017 by guest