The Structure of CodY, a GTP- and Isoleucine-responsive Regulator of Stationary Phase and Virulence in Gram-positive Bacteria

Vladimir M. Levdkov, Elena Blagova, Pascale Joseph, Abraham L. Sonenshein, and Anthony J. Wilkinson

From the Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, United Kingdom and the Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111

CodY is a global regulator of transcription in Gram-positive bacteria. It represses during growth genes required for adaptation to nutrient limitation, including virulence genes in some human pathogens. CodY activity is regulated by GTP and branched chain amino acids, metabolites whose intracellular concentrations drop as cells enter stationary phase. Although CodY has a highly conserved sequence, it has no significant similarity to proteins of known structure. Here we report crystal structures of two fragments of CodY from Bacillus subtilis that clearly constitute its cofactor and DNA binding domains and reveal that CodY is a chimera of previously observed folding units. The N-terminal cofactor-binding fragment adopts a fold reminiscent of the GAF domains found in cyclic nucleotide phosphodiesterases and adenylate cyclases. It is a dimer stabilized by an intermolecular six α-helical bundle that buries an extensive apolar surface rich in residues invariant in CodY orthologues. The branched chain amino acid ligands reside in hydrophobic pockets of each monomer distal to the dimer-forming surface. The structure of the C-terminal DNA binding domain belongs to the winged helix-turn-helix family. The implications of the structure for DNA binding by CodY and its control by cofactor binding are discussed.

The chromosomes of bacteria are packed with genes that enable the organism to exploit diverse ecological niches and to survive adversity. The key to adaptation and survival is to interpret indicators of a changing environment and respond by modulating the repertoire of genes being expressed. A fundamental challenge to survival is posed by the depletion of nutrients and the onset of starvation. Bacillus subtilis has at its disposal an array of adaptations to poor growth conditions, including virulence genes involved in peptide and amino acid utilization (4). In pathogenic bacteria such as Streptococcus pyogenes, Listeria monocytogenes, Enterococcus faecalis, Bacillus anthracis, and Clostridium difficile, accumulating evidence points to a role for CodY in the regulation of virulence gene expression (6).

CodY was first discovered as a repressor of the dipeptide permease operon in B. subtilis (1). Subsequently it has been recognized as having a much wider role in controlling the expression of stationary phase genes. It is now known to regulate over 100 genes distributed across some 70 or so operons (2). The CodY regulon encodes extracellular degradative enzymes, transporter proteins, catabolic enzymes, factors involved in genetic competence, antibiotic synthesis pathways, chemotaxis proteins, and sporulation proteins. During rapid growth, CodY thus represses genes whose products allow adaptation to nutrient depletion. The repressor function of CodY is activated by two different effectors, GTP and isoleucine or valine, which may be viewed as sensors of the energetic and metabolic status of the cell, respectively (3–5). These co-repressors act independently and additively to increase the affinity of CodY for its target sites on DNA. CodY binds its ligands selectively but with moderate affinities (in the millimolar range), consistent with its need to monitor the concentrations of GTP and isoleucine/valine obtaining in rapidly growing cells.

CodY is highly conserved in the low G + C Gram-positive bacteria. In Lactococcus lactis, CodY regulates the expression of genes encoding extracellular peptidases, peptide transport proteins, and intracellular enzymes involved in peptide and amino acid utilization (4). In pathogenic bacteria such as Streptococcus pyogenes, Listeria monocytogenes, Enterococcus faecalis, Bacillus anthracis, and Clostridium difficile, accumulating evidence points to a role for CodY in the regulation of virulence gene expression (6).

CodY from B. subtilis is a dimer of 29-kDa subunits. Its 259-amino acid residue sequence is unrelated to that of any proteins other than CodY orthologues. A putative helix-turn-helix (HTH) motif has been identified close to the C terminus of the protein and specific residues in this motif are required for high affinity DNA binding (7, 8). The only other clue to structure provided by the sequence is a set of possible GTP binding motifs again located in the C-terminal half of the molecule (3). To investigate the structural basis of GTP- and branched chain amino acid (BCAA)-dependent regulation of stationary phase gene expression, we embarked on crystallographic studies of CodY. Although we have been able to crystallize the intact protein (9), the crystals were weakly diffracting and unsuitable for structure determination. We therefore sought to identify stable proteolytic fragments of CodY as alternative targets for structure determination. Here we report the crystal structures of N- and C-terminal fragments constituting residues 1–155 and 168–259, respectively.

The abbreviations used are: HTH, helix-turn-helix; BCAA, branched chain amino acid; bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxyethyl)propane-1,3-diol; SeMet, selenium methionine; BSA, bovine serum albumin; PDE-2A, phosphodiesterase 2A; Gpp(NH)p, guanosine 5’-[β,γ-imido]triphosphate.

* This work was supported by the Biotechnology and Biological Science Research Council, United Kingdom, Grant BBS/B/1213X, European Union SPINE Contract QLG2-CT-2002-00988, and United States Public Health Service Grant GM042219. The costs of publication of this article were defrayed in part by the payment of page charges. This advertisement must therefore be hereby marked "advertisement" in accordance with 18 U.S.C.

11366 JOURNAL OF BIOLOGICAL CHEMISTRY VOLUME 281 • NUMBER 16 • APRIL 21, 2006
TABLE 1

| X-ray data collection and refinement statistics |
|-----------------------------------------------|
| **Data collection**                            |
| Crystalline form | Native | SeMet CodY | SeMet CodY peak | SeMet CodY inflection point | SeMet CodY remote |
| Crystal | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| X-ray source | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Wavelength (Å) | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Resolution range (Å) | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Space group | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Unit cell parameters (Å) | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Unique reflections, overall/outer<sup>a</sup> | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Completeness (%), overall/outer shell<sup>b</sup> | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| Redundancy, overall/outer shell<sup>c</sup> | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| I(>1σ), overall/outer shell<sup>d</sup> | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |
| R<sub>merge</sub><sup>e</sup>, (%), overall/outer shell<sup>f</sup> | 1.628 | 1.628 | 1.628 | 1.628 | 1.628 |

Refinement and model statistics

| Resolution range (Å) | 2.00–1.80 | 2.00–1.80 | 2.00–1.80 | 2.00–1.80 |
| R factor (R-free)<sup>g</sup> | 0.153 (0.212) | 0.164 (0.2440) | 0.151 (0.318) | 0.168 (0.227) |
| Reflections (working/free)<sup>h</sup> | 14,653/775 | 7,231/350 | 819 / 64 | 479 / 14 |
| Outer shell reflection (working/free)<sup>i</sup> | 1386 | 1375 | 175 | 131 |
| Molecules/asymmetric unit | 1 | 1 | 1 | 1 |
| No. of protein non-H atoms | 34.72, 168.90 | 54.85, 273.90 | 25.94, 164.90 | 54.93, 273.90 |
| Number of water molecules | 34.72, 168.90 | 54.85, 273.90 | 25.94, 164.90 | 54.93, 273.90 |
| Ramachandran plot<sup>j</sup> | 94.4/18.43 | 95.6/66.5 | 92.6/45.4 | 95.9/67.8 |
| Bond lengths (Å) | 7.2/56.5 | 7.0/57.2 | 6.4/57.8 | 6.7/56.4 |
| Ramachandran plot<sup>g</sup> | 92.6/74.0/0.0/0.0 | 92.6/74.0/0.0/0.0 | 92.6/74.0/0.0/0.0 | 92.6/74.0/0.0/0.0 |

**Experimental Procedures**

Protein Preparation and Crystallization—CodY was purified as a C-terminal histidine-tagged protein (9) and partially digested with a series of proteases. Mass spectrometry of the digestion products led us to conclude that a region containing a cluster of 12 highly charged residues (156–167) was particularly sensitive to proteolysis. Using ligand-independent cloning methods, we prepared pET28a derivative constructs encoding CodY fragments spanning residues 1–155 and 168–259, each with an N-terminal His<sub>6</sub> tag. The recombinant CodY fragments were overexpressed in Escherichia coli BL21 and purified by (i) Ni<sup>2+</sup>-chelation chromatography on a 5-ml high-performance chelating-Sepharose column (Amersham Biosciences) charged with nickel ions and (ii) gel filtration chromatography. For CodY-(1–155) a HiLoad Superdex 200 preparation grade column equilibrated in 50 mM Tris, pH 8.0 was used.
7.5, 150 mM sodium chloride was used. For CodY-(168–259) a Superdex 75 preparation grade column equilibrated with 0.1M sodium citrate, pH 5.6, 200 mM sodium chloride was used. SeMet derivatives were prepared by growth of E. coli B834 cells harboring the expression plasmids in ZYP-5052 autoinduction medium (10).

Crystals of CodY-(1–155) were grown in hanging drops containing a 1:1 volume ratio of 20 mg ml\(^{-1}\) protein in gel filtration buffer and 18% monomethyl ethyl polyethylene glycol 5000 in 0.1M Tris-HCl, pH 7.0, 0.2 M calcium acetate, 10 mM Gpp(NH)p, and either 10 mM isoleucine (native protein) or 10 mM valine (SeMet protein). Crystals of CodY-(168–259) were grown similarly from 1:1 mixtures of 20 mg ml\(^{-1}\) protein in 100 mM sodium citrate, pH 5.6, 0.2 M NaCl and 32% saturated ammonium sulfate in 0.1 M bis-Tris, pH 6.5, 5% glycerol, and 5% Tacsimate (Hampton Research).

**Partial Proteolysis of CodY**—CodY-His\(_6\) protein (84 μg, purified as described above) or bovine serum albumin, Fraction V (BSA; 84 μg; Sero- logicals Products, Inc.), were preincubated at 25 °C for 10 min with or without BCAAs (10 mM each of isoleucine, leucine, and valine). We then added trypsin (0.26 μg, Sigma) or α-chymotrypsin (0.52 μg, Sigma) and incubated the samples at 25 °C in 0.12 ml of a buffer containing 20 mM Tris-HCl, pH 8, 1 mM MgCl\(_2\), 1 mM dithiothreitol, and 0.5 mM EDTA (for trypsin) or 2 mM CaCl\(_2\) (for α-chymotrypsin). After various times, 0.01-ml samples were removed, denatured, and subjected to SDS-PAGE. After electrophoresis, the proteins were electrotransferred to sheets of nitrocellulose. Bands that stained with Coomassie Blue were excised from the nitrocellulose and subjected to N-terminal sequencing by successive Edman degradation by the Tufts Protein and Nucleic Acid Core Facility.

**Structure Determination**—Native data taken from crystals of CodY-(1–155) to 1.7-Å resolution were collected on beamline 10.1 at the SRS (Daresbury, UK), using a MAR Research CCD165 detector. Three-wavelength CodY-(1–155) SeMet data were collected to 2.3-Å resolution at beamline BM14 at the ESRF (Grenoble, France). Data were processed and scaled using the program HKL2000 (11). The crystals were assigned to space group P2\(_1\)2\(_1\)2 with one molecule per asymmetric unit. The structure was solved by MAD phasing using the program HKL2MAP (12). The model was built using the program ARP/WARP (13) with manual refitting in the program COOT (14) and refined with REFMAC (15). The electron density maps were of excellent quality and all 155 CodY residues of the single molecule in the asymmetric unit of the crystal were defined together with an isoleucine ligand and 175 solvent molecules. The final free \(R\) factor is 21.2%, \(R\) work is 15.3%, and the root mean square deviations of bond lengths and bond angles from ideal geometry are small (Table 1).

The crystals of CodY-(168–259) were in space group P422. The crystals were weakly scattering and the x-ray diffraction data were limited to 2.8–3.0 Å spacing. The structure was solved using three wavelength MAD data and the model was built and refined as described above. The model contains a complete A chain and all but residues 234–235 and
FIGURE 2. Alignment of the sequences of representative CodY orthologues. The alignment was generated in the program ClustalW and produced in ESPript (35). The secondary structure elements of the subtilis CodY were indicated above the alignment. Residues whose side chains in CodY from B. subtilis form prominent interactions with the isoleucine ligand are indicated by asterisks below the alignment, whereas residues whose side chains form prominent dimer-forming interactions are indicated by filled triangles. The species are: Bacsu, B. subtilis; Bacan, B. anthracis; Clodi, C. difficile; Clope, C. perfringens; Entfa, E. faecalis; Lisino, L. monocytogenes; Staau, Staphylococcus aureus; Strpy, Streptococcus pyogenes; Strag, Streptococcus agalactiae; Strth, Streptococcus thermophilus; Strpn, S. pneumoniae; Lacla, L. lactis.
Structure of the Global Regulator CodY

257–259 in chains B and C. The temperature factors of chain C are significantly higher than for chains A and B. The final free R factor is 24.8%, Rwork is 20.0% and the geometry is satisfactory (Table 1).

RESULTS

Tertiary Structure of CodY-(1–155)—CodY-(1–155) has a three-layered globular structure. At the base of the molecule, as shown in Fig. 1A, is a three-helix bundle. Two of the helices are at the N terminus of the chain, whereas the third is at the C terminus of the chain. Helices α2 and α5 pack against a central 5-stranded anti-parallel β-pleated sheet, with a strand order β2–β1–β3–β4–β3 (Fig. 1A and B). The top layer, which does not lend itself to such simple description, is formed by two extended loops that connect strands β2 and β3, and β3 and β4. The β2–β3 segment contains two α-helices α3 and α4, whereas the β2–β3 loop lacks any recognizable secondary structure. These loops form the walls of a cavity whose base is formed by the β-sheet itself.

Amino Acid Binding—The electron density maps clearly define the presence of an isoleucine ligand consistent with the inclusion of this amino acid in the crystallization drops (Fig. 1D). No evidence for the binding of the GTP analogue, Gpp(NH)p was seen, however. The isoleucine is bound above the β-sheet in such a way that the β2–β3 and β3–β4 loops effectively clasp the ligand. The isobutyl group of the ligand projects downwards toward the sheet so that it is enclosed in a hydrophobic cavity circumscribed by the side chains of Met62, Met65, Phe71, Pro72, Tyr75, and Pro99 (Fig. 1E) and main chain atoms from residues 97–99. Valine binds to CodY in a very similar manner to isoleucine, the only significant difference in binding is caused by the absence of a methylene group in valine relative to isoleucine, which may be associated with a small rotation of the phenyl ring of Phe71 (Fig. 1F). The α-amino and α-carboxylate groups of the amino acid ligand form polar interactions with the protein and solvent. The former forms charge-dipole interactions with the main chain carbonyl groups of Thr60 and Phe98 and a further hydrogen bond to the surface water molecule WAT24. The carbonylate group forms a two-pronged ion-pairing interaction with the guanidinium group of Arg61 and further polar interactions with the amide group of Val100 and the water molecule WAT143.

The residues making up the isoleucine-binding pocket are strongly conserved in the Bacillus, Listeria, Enterococcus, Staphylococcus, Streptococcus, and Lactococcus spp., although surprisingly they are less well conserved in the Clostridium spp. (Fig. 2). Arg61, which binds the ligand carbonylate, is replaced by apolar side chains in CodY from C. difficile and Clostridium perfringens, whereas the apolar Met65, which contacts the apolar side chain of the ligand, is replaced by Glu or Lys in these species. This suggests differences in the mode of ligand binding.

Quaternary Structure—Analysis of the molecular packing suggests an obvious dimer consistent with observations made from gel filtration of CodY-(1–155) during purification, and with the knowledge that intact CodY is a dimer (9). In this dimer, the subunits are related by a 2-fold crystallographic symmetry axis. Intermolecular contacts are mediated principally through the first and the last α-helices of the respective domains (Fig. 3). These helices come together to form a four-helix bundle that is extended to a six-helix bundle by the adjacent helices α2 from the respective subunits (Fig. 3). Each subunit contributes ~1050 Å² of its surface area to the contact interface, corresponding to 13% of the total accessible surface area and typical of a dimer interface (16).

Highly conserved and/or invariant residues among CodY orthologues cluster at the dimer interface. There are two intermolecular salt bridges formed between the pairs of Arg6 and Glu144 residues, which engage in two-pronged interactions with each other. These interactions would appear to be crucial for CodY structure and/or function as both residues are invariant in the CodY orthologues (Fig. 2). There is an additional pair of hydrogen bonding interactions between Gln15 and Thr148, the former residue being strongly conserved whereas the latter is invariant.

In the dimer, the isoleucine ligands of the respective subunits are distal to the dimer interface and their Cα atoms are separated by 57 Å (Fig. 3). The C-terminal Arg155 residues that in the intact protein are connected to the DNA binding domains are close in space.

CodY Has a GAF Domain—Despite the absence of any substantial sequence homology, a number of structures in the Protein Data Bank contain domains that can be superimposed onto CodY-(1–155) to yield positional root mean squared deviations of 2.5–4.0 Å for 80–120 matching Cα atoms (17, 18). The shared characteristic of these proteins is a GAF domain, named after its discovery, on the basis of sequence analysis, in cGMP-stimulated phosphodiesterases, adenylate cyclases and a bacterial transcription regulator EhlA (19). It is now recognized that GAF domains are present in numerous signaling and sensory proteins.

The structure of the GAF domain was first defined in a study of the yeast protein YKG9 (Protein Data Bank entry code 1fsm) (18) and subsequently other GAF domain structures in complex with ligands have been determined (Protein Data Bank entry codes 1mco and 1ykd (20, 21)). The superposition of CodY and YKG9 shows clearly that CodY belongs to the GAF domain family (Fig. 4, A and B). The β-sheet and two of the α-helices that pack against it are superimposable in all GAF domain proteins, with family members distinguished by differences in the loops connecting β-strands on the face of the sheet distal to these helices. In the ligand-responsive members of the family (20, 21), these loops shape the binding pocket.

Structure of the C-terminal Domain—CodY-(168–259) has a compact globular structure with an anti-parallel β-sheet of topology β6–β8–β7 and five α-helices arranged with respect to the strands in the order α6–α7–β6–α8–α9–β7–β8–α10 (Fig. 1C). Helices α8 and α9 correspond to the predicted helix-turn-helix spanning residues 203–226. The sequence in this region is exceptionally highly conserved in the CodY orthologous set (Fig. 2). The role of this HTH in DNA binding has been established by site-directed mutagenesis studies, which have shown that non-conservative substitutions of Ala207, Arg214, Ser215, and Val218 dramatically lower the affinity of CodY for ilvB and dpp promoter DNAs (8). The structure presented here shows that the side chains of Arg214, Ser215, and Val218 are situated at the beginning of the putative recognition helix with their side chains pointing outwards from the molecule, consistent with a role in DNA recognition. Ala207 resides on the preceding helix at a position where a small side chain would be required if the HTH helices are to be arranged in the current manner.

Structural comparisons reveal that the DNA binding domain of CodY belongs to the winged HTH family of nucleic acid-binding
proteins (17, 22). These domains consist of three $\alpha$-helices ($\alpha_7$, $\alpha_8$, and $\alpha_9$ in CodY) packing against a three-stranded anti-parallel $\beta$-sheet. The so-called “wings,” W1 and W2, are constituted by the loop connecting the second and third strands ($\beta 7$–$\beta 8$) and a looping segment following the third strand ($\beta 8$), respectively. In three dimensions, these segments flank the second helix of the HTH apparently like the wings of a butterfly.

BCAA-induced Conformational Change—Because crystals prepared in the absence of BCAAs did not diffract well, we sought an alternative way of testing whether interaction with BCAAs induces a conformational change in CodY. We observed that trypsin has a preferred cleavage site after Lys$^{169}$. Cleavage at other lysine and arginine residues also occurs, albeit with much lower efficiency (Fig. 5A). Isolation of each of the bands in Fig. 5A allowed us to determine the N-terminal sequence of each fragment and thereby deduce the sites of cleavage. Based on this analysis, we determined that BCAAs had no effect on cleavage after Lys$^{169}$, but almost completely inhibited cleavage after Lys$^{169}$, Arg$^{170}$, Arg$^{170}$, and Arg$^{170}$. As a control, we showed that the cleavage pattern of CodY was not affected by threonine or methionine (data not shown) and that cleavage of BSA was the same in the presence or absence of BCAAs (Fig. 5C and D). Incubation of CodY with $\alpha$-chymotrypsin yielded at least seven fragments (Fig. 5E). Of these, only cleavages after Tyr$^{149}$, after Phe$^{190}$, and near Tyr$^{191}$ occurred in the presence of BCAAs (Fig. 5F).

Cleavage at other sites, such as Tyr$^{149}$, Tyr$^{149}$, and Met$^{173}$ was completely inhibited by BCAAs. Again, cleavage of BSA was unaffected by BCAAs (Fig. 5, G and H). It is interesting to note that the presumed conformational change in CodY that accompanies binding of BCAAs can affect residues far removed in space from the co-repressor binding site.

DISCUSSION

DNA binding and transcriptional repression by CodY are activated by the binding of the effector molecule, be it GTP or a BCAA (3, 5). The proteolytic fragmentation of CodY and the resulting structural studies presented here lead to the conclusion that CodY is a modular protein made up of an N-terminal cofactor binding domain and a C-terminal DNA binding domain, belonging to the GAF and winged HTH domain families, respectively. These domains have been characterized in other systems, but their combination within a single polypeptide chain has not been reported before.

The exact mechanism by which binding of co-repressors increases the affinity of CodY for its target sites remains mysterious (see below), but it seems clear that interaction with BCAAs alters the conformation of CodY enough to change the susceptibility of certain residues to proteolytic enzymes (Fig. 5). These residues include some that are outside the N-terminal domain, suggesting that binding of BCAAs induces a conformational change that is translated to the C-terminal, DNA-binding domain.

Implications for GTP Binding—The GAF domain complexes described here define the BCAA binding site in CodY leaving open the question of how GTP is bound. In the absence of a structure of a GTP complex, we have examined other GAF domain proteins and CodY sequence alignments for other clues to nucleotide binding. Two of the known GAF domain proteins contain nucleotide ligands, cyclic guanosine monophosphate (cGMP) in mouse phosphodiesterase 2A (PDE-2A) and a pair of cyclic adenosine monophosphate (cAMP) ligands in the...
Structure of the Global Regulator CodY

tandem GAF domains of a cyanobacterial adenylyl cyclase (20, 21). The
topologies of the ligand-binding sites are very similar and the ligands
and the surrounding protein backbones are closely superimposable.

Compared with isoleucine in CodY, the cGMP in PDE-A2 is more
deeply embedded in the GAF domain (Fig. 4C). Following superposition,
the guanine base lies close to a pair of buried water molecules that
occupy the largest cavity in the CodY structure (Fig. 4D). These waters
are flanked by the side chains of Phe\(^{40}\) and Phe\(^{88}\), which are therefore
well placed to make ring-stacking interactions with the guanine base of
a putative GTP ligand. Phe\(^{88}\) is invariant, whereas Phe\(^{40}\) is conserved in
all of the CodYs except those from \(L.\) lactis and the Streptococcus spp.
Interestingly the CodYs from \(L.\) lactis and Streptococcus pneumoniae
(and by inference the other Streptococcus spp.) do not respond to GTP
(23, 24).\(^5\) The side chains of residues Gln\(^{58}\), Gln\(^{55}\), and Ser\(^{129}\) provide
hydrogen-bonding opportunities for the buried waters and potentially
for the polar groups of a guanine base. Ser\(^{129}\) is conserved as a small
residue (Ser, Gly, or Ala) in the CodY orthologues in Fig. 2, with the
exception of those from \(L.\) lactis and the Streptococcus spp., which may
again relate to their failure to respond to GTP. It is tempting to suggest
that the Trp replacements in these CodY proteins provide an indole
group to fill the volume that would otherwise be available to the guanine
base of GTP.

The cavity containing these waters is contiguous with the isoleucine-
binding pocket although it tapers considerably before opening out into
the latter. However, given that ligand binding is probably accompanied
by closure of the \(\beta_2-\beta_3\) and \(\beta_3-\beta_4\) loops by a claw-like action, it is possible
that these elements could adapt their conformation so as to
embrace the GTP ligand. Thus, the nucleotide base may bind in this
cavity with the ribose and triphosphate moieties extending toward the
surface where Arg\(^{61}\) may mediate important interactions with the
\(\gamma\)-phosphate, which is known to be a determinant of the action of GTP
because GDP is not an effector of CodY (3). This speculative model for
nucleotide binding predicts overlapping binding sites for GTP and iso-
leucine. The implied mutually exclusive binding of corepressors, how-
ever, is not easily reconcilable with the observation that GTP and iso-
leucine have independent and additive effects on DNA binding by
CodY. Although we have thus far been unable to grow crystals of CodY-
(1–155) in the absence of isoleucine/valine, the full-length protein has
been crystallized both in the absence of effectors and in the presence of
GTP as sole effector (9) leaving open the possibility that at least one of
the GTP binding determinants, such as the putative G motifs (3) resides
in the C-terminal domain.

Implications for Cofactor Control in CodY—In other transcriptional
regulatory systems, ligand binding promotes relative domain motions
and/or quaternary structural changes. In CodY, the isoleucine-binding
site is not linked to any obvious hinge that could mediate long-range
conformational change. As the isoleucines are bound in pockets on the
faces of the \(\beta\)-sheets distal to the dimer interface, it seems unlikely that
the effects of ligand binding would alter subunit interactions. Indeed there is no
evidence for ligand-dependent quaternary structural changes in CodY.

The organization of the CodY-(1–155) subunits in the dimer places the
C termini in close apposition, allowing for the possibility that the
emerging DNA binding domains reside side by side. However, 12 linker
residues between 155 and 168, a high proportion of which have ioniz-
able side chains (Fig. 2), are not present in either of the structures and an
appreciation of the juxtaposition of the domains in intact CodY will
have to await the solution of the structure of the intact protein.

\(^5\) R. P. Shivers and A. L. Sonenshein, unpublished data.

A system with possible structural and mechanistic analogy to
CodY is TraR. TraR in Agrobacterium tumefaciens is a quorum sensing
protein that responds to a homoserine lactone pheromone. The struc-
ture of TraR has been solved in complex with its cofactor and with a tra
box target DNA (Fig. 6A) (25, 26). The cofactor binding domain has a
GAF domain fold. Like CodY, TraR is a dimer with the monomer-
monomer interface constituted by the helical regions of the subunits
that are distal to the effector binding sites. The DNA binding domains
emerge from corresponding locations, but instead of the winged HTH
domain observed in CodY, the DNA binding domain in TraR belongs to
the LuxR family and has a GerE-type fold (27).

The homoserine lactone ligand is believed to stabilize TraR against
intracellular proteolysis (28). Consistent with this idea, it is even more
deeply recessed in the hydrophobic cavity of the GAF domain (25, 26)
than the isoleucine in CodY (Fig. 4C). This suggests that CodY too may
be susceptible to proteolysis as BCAA/GTP levels drop when the cells
enter the stationary phase. Although BCAA binding alters the suscepti-
bility to in vitro proteolysis (Fig. 5), in vivo measurements indicate that
CodY protein is present at the same concentrations throughout growth
and stationary phase (3).

Implications for DNA Binding—A winged HTH domain with similarity
to that observed in CodY occurs in the fatty acid-responsive tran-
scription factor FadR. The structure of FadR in its complex with FadR
operator DNA is shown in Fig. 6B (29, 30). The winged HTH DNA
binding domain is connected to a C-terminal seven-helix bundle
domain containing a hydrophobic pocket for acyl-CoA binding. FadR
binds to DNA as a dimer using its recognition helix to make contacts
with the major groove of the DNA and one of its wings, W1, to penetrate
the minor groove.

The winged HTH domains of FadR and CodY can be superposed on
one another to give a positional root mean square deviation of 2.7 \(\AA\) for
the 55 matching \(\text{Ca}\) atoms. This comparison shows that the invariant
residues among CodY orthologues lie on the face of the molecule cor-
responding to the DNA binding surface in FadR. These are located in
the segment L\(^{179}\)SYSE\(^{183}\) in the turn between the helices \(\alpha_6\) and \(\alpha_7\),
residues 211–228, which span the recognition helix and the preceding
turn, and residues 232–240, which constitute the wing, W1 (Fig. 2). In
the latter instance residues Arg\(^{233}\) and Lys\(^{238}\) appear to be well placed
for forming interactions with the backbone phosphates of the DNA.

It is interesting to note that there is no tertiary structural similarity
between CodY and the leucine-responsive regulatory protein, which

\(3\) J. M. Goldbach and J. E. Shearer, unpublished data.
performs in Gram-negative bacteria many of the functions attributed to CodY in B. subtilis (31, 32). It would appear that amino acid and guanine nucleotide regulation of stationary phase gene expression has evolved independently in these two systems. Finally in the process of manuscript review, a referee drew our attention to a bioinformatics analysis where it is shown that the GAF plus winged HTH domain composition of CodY is predictable by bioinformatics analysis.

Acknowledgments—We thank the Tufts Protein and Nucleic Acid Core Facility for assistance with peptide sequencing, and the staff at the Synchrotron Radiation Source at Daresbury, UK, and the European Synchrotron Radiation Facility in Grenoble, France, for help with data collection, and Dr. James A. Brannigan for helpful comments and discussion.

REFERENCES

1. Slack, F., Serror, P., Joyce, E., and Sonenshein, A. L. (1995) Mol. Microbiol. 15, 689–702
2. Molle, V., Nakaura, Y., Shivers, R. P., Yamaguchi, H., Losick, R., Fujita, Y., and Sonenshein, A. L. (2003) J. Bacteriol. 185, 1911–1922
3. Ratnayake-Lecamwasam, M., Serror, P., Wong, K. W., and Sonenshein, A. L. (2001) Genes Dev. 15, 1093–1103
4. Guedon, E., Serror, P., Ehrlich, S. D., Renault, P., and Delorme, C. (2001) Mol. Microbiol. 40, 1227–1239
5. Shivers, R. P., and Sonenshein, A. L. (2004) Mol. Microbiol. 53, 599–611
6. Sonenshein, A. L. (2005) Curr. Opin. Microbiol. 8, 203–207
7. Serror, P., and Sonenshein, A. L. (1996) Mol. Microbiol. 20, 843–852
8. Joseph, P., Ratnayake-Lecamwasam, M., and Sonenshein, A. L. (2005) J. Bacteriol. 187, 4127–4139
9. Blagova, E. V., Levdiakov, V. M., Tachikawa, K., Sonenshein, A. L., and Wilkinson, A. J. (2003) Acta Crystallogr. Sect. D 59, 155–157
10. Studier, F. W. (2005) Protein Expression Purif. 41, 207–234
11. Otwonowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
12. Pape, T., and Schneider, T. R. (2004) J. Appl. Crystallogr. 37, 843–844
13. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) Methods Enzymol. 374, 229–244
14. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. Sect. D 60, 2126–2132
15. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D 53, 240–255
16. Argos, P. (1988) Protein Eng. 2, 101–113
17. Holm, L., and Sander, C. (1993) J. Mol. Biol. 233, 123–138
18. Ho, Y.-W. J., Burden, L. M., and Hurley, J. H. (2000) EMBO J. 20, 5288–5299
19. Avarird, L., and Ponting, C. P. (1997) Trends Biochem. Sci. 22, 458–459
20. Martine, S. E., Wu, A. Y., Glavas, N. A., Tang, X.-B., Turley, S., Hol, W. G. J., and Brannigan, J. A. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 3260–3265
21. Martine, S. E., Bruder, S., Schultz, A., Zheng, N., Schultz, J. E., Beavo, J. A., and Linder, J. U. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 3082–3087
22. Gajiwala, K. S., and Burley, S. K. (2000) Curr. Opin. Struct. Biol. 10, 110–116
23. Petrovovic, D., Guedon, E., Ehrlich, S. D., Delorme, C., Ehrlich, P., and Renaut, P. (2004) Mol. Microbiol. 53, 613–621
24. den Hengst, C. D., Carley, P., Larsen, R., Buit, G., Nauta, A., van Sinderen, D., Kuipers, O. P., and Kok, J. (2005) J. Bacteriol. 187, 512–521
25. Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., Neddermann, P., and Di Marco, S. (2002) EMBO J. 21, 4393–4401
26. Zhang, R.-G., Pappas, T., Brace, J. L., Miller, P. C., Oulmassov, T., Molyneaux, J. M., Anderson, J. C., Bashkin, J. K., Winans, S. C., and Joehmiak, A. (2002) Nature 417, 971–974
27. Ducros, V. M.-A., Lewis, R. J., Verma, C. S., Dodson, E. J., Leonard, G., Turkenburg, J. P., Murshudov, G. N., Wilkinson, A. J., and Brannigan, J. A. (2001) J. Mol. Biol. 306, 739–771
28. Zhu, J., and Winans, S. C. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 1507–1512
29. Van Aalten, D. M. F., DiRusso, C. C., and Knudsen, J. (2001) EMBO J. 20, 2041–2050
30. Xu, Y., Heath, R. J., Li, Z., Rock, C. O., and White, S. W. (2001) J. Biol. Chem. 276, 17373–17379
31. Leonard, P. M., Smiths, S. H., Sedelnikova, S. E., Brinkman, A. B., de Vos, W. M., van der Oost, J., Rice, D. W., and Rafferty, J. B. (2001) EMBO J. 20, 990–997
32. Tani, T. H., Khodursky, A., Blumenthal, R. M., Brown, P. O., and Matthews, R. G. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 13471–13476
33. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Emsley, P., Cowtan, K., Murshudov, G. N., Cohen, S., Perrakis, A., and Noble, M. (2004) Acta Crystallogr. Sect. D 60, 2288–2294
34. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
35. Gouet, P., Courcelle, E., Stuart, D. L., and Metoz, F. (1999) Bioinformatics 15, 305–308