White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/447/

Article:
Lewis, R J, Krzywda, S, Brannigan, J A et al. (2000) The trans-activation domain of the sporulation response regulator Spo0A revealed by X-ray crystallography. Molecular Microbiology. pp. 198-212. ISSN: 0950-382X

https://doi.org/10.1046/j.1365-2958.2000.02134.x

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Molecular Microbiology (2000) **38**(2), 198–212

The *trans*-activation domain of the sporulation response regulator Spo0A revealed by X-ray crystallography

Richard J. Lewis,[1] Szymon Krzywda,[1] James A. Brannigan,[1] Johan P. Turkenburg,[1] Katarína Muchová,[2] Eleanor J. Dodson,[1] Imrich Barák[2] and Anthony J. Wilkinson[1]*
[1]*Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK.*
[2]*Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, 842 51 Bratislava, Slovak Republic.*

Summary

Sporulation in *Bacillus* involves the induction of scores of genes in a temporally and spatially co-ordinated programme of cell development. Its initiation is under the control of an expanded two-component signal transduction system termed a phosphorelay. The master control element in the decision to sporulate is the response regulator, Spo0A, which comprises a receiver or phosphoacceptor domain and an effector or transcription activation domain. The receiver domain of Spo0A shares sequence similarity with numerous response regulators, and its structure has been determined in phosphorylated and unphosphorylated forms. However, the effector domain (C-Spo0A) has no detectable sequence similarity to any other protein, and this lack of structural information is an obstacle to understanding how DNA binding and transcription activation are controlled by phosphorylation in Spo0A. Here, we report the crystal structure of C-Spo0A from *Bacillus stearothermophilus* revealing a single α-helical domain comprising six α-helices in an unprecedented fold. The structure contains a helix–turn–helix as part of a three α-helical bundle reminiscent of the catabolite gene activator protein (CAP), suggesting a mechanism for DNA binding. The residues implicated in forming the σ^A-activating region clearly cluster in a flexible segment of the polypeptide on the opposite side of the structure from that predicted to interact with DNA. The structural results are discussed in the context of the rich array of existing mutational data.

Accepted 1 August, 2000. *For correspondence. E-mail ajw@ysbl.york.ac.uk; Tel. (+44) 1904 432519; Fax (+44) 1904 410519.

© 2000 Blackwell Science Ltd

Introduction

In its natural habitat, the soil, *Bacillus subtilis* experiences drastic fluctuations in its environment. As a result, it has developed a series of strategies to survive physical stresses and starvation. It can synthesize antibiotics to kill competitors, it can secrete degradative enzymes to scavenge vital nutrients and it can develop motility and move towards new sources of nutrients or away from sources of stress. More elaborate adaptation involves two mutually exclusive developmental pathways. First, it may develop competence to take up exogenous nucleic acid, perhaps conferring a genetic advantage over its rivals; alternatively, *B. subtilis* may form a resistant spore.

Sporulation commences with an asymmetric cell division producing two compartments of unequal size but each containing an identical chromosome. The smaller cell, the forespore, is subsequently engulfed by the larger mother cell, and the two cells collaborate in the construction of a thick proteinaceous shell around the developing spore. In the final stage, the mother cell lyses, releasing a mature spore that can lie dormant indefinitely and germinate when favourable conditions for growth are restored. Two aspects of this developmental cascade are key: (i) the cell must determine precisely when it is appropriate to form a spore; sporulation is costly in terms of time and energy and, once the asymmetric septum has formed, the process cannot easily be reversed; (ii) there must be a mechanism for activating the large array of hitherto silent genes required for the regulated assembly of a viable spore. The key molecule in transducing the signals indicative of a deteriorating cellular environment into the activation of genes required for sporulation is the response regulator, Spo0A.

The initiation of sporulation is under the control of an expanded two-component sensory signalling system termed a phosphorelay (Burbulys *et al*., 1991; Hoch, 1993). Two-component systems consisting of a sensor kinase and a response regulator are the most commonly used mechanism for signal transduction in microorganisms. Environmental signals trigger the ATP-dependent autophosphorylation of three or more sensor kinases on a specific histidine residue. The phosphoryl group then migrates via Spo0F and Spo0B to an aspartic acid residue in the response regulator Spo0A. Spo0A is the master control element in the decision to sporulate. If a threshold concentration of Spo0A~P is attained, sporulation

commences. Phosphorylation of Spo0A activates its latent transcription activation and repression properties by stimulating binding to 7 bp DNA consensus sequences 5′-TGTCGAA-3′, known as 0A-boxes, present in multiple copies at Spo0A-regulated promoters (Perego *et al*., 1988; Strauch *et al*., 1990; Baldus *et al*., 1994).

Spo0A~P activates transcription from the promoters of the *spoIIA* (Trach *et al*., 1991; Wu *et al*., 1991), *spoIIG* (Satola *et al*., 1991; Baldus *et al*., 1994) and *spoIIE* (York *et al*., 1992) operons, which encode forespore- (σ^F) and mother cell (σ^E)-specific RNA polymerase sigma factors and a protein phosphatase that localizes to the sporulation septum respectively. It also activates transcription of genes encoding the phosphorelay components Spo0F (Strauch *et al*., 1993) and Spo0A itself (Strauch *et al*., 1992). Transcription of *spoIIA* depends on RNA polymerase containing σ^H (E-σ^H), whereas *spoIIG* and *spoIIE* transcription depends on σ^A, the primary vegetative phase sigma factor of *B. subtilis* (Kenney *et al*., 1988; Wu *et al*., 1991). Spo0A can thus activate transcription from promoters controlled by two different sigma factors. Spo0A~P is also a negative regulator of transcription of *abrB* (Strauch *et al*., 1990), *kinA* (Hoch, 1993) and, at higher concentrations, *spo0F* and *spo0A* (Strauch *et al*., 1992). The overall pattern of gene regulation is more complex because AbrB is a repressor of transcription of *spo0H* that encodes σ^H (Weir *et al*., 1991). The importance of Spo0A as a global regulator of transcription is highlighted in a recent study, which reveals over 500 genes whose transcript levels are Spo0A dependent (Fawcett *et al*., 2000).

Spo0A from *B. subtilis* consists of a single polypeptide chain of 29.5 kDa, which forms two domains of similar size, an N-terminal regulatory domain and a C-terminal *trans*-activation domain. The regulatory domain contains all the signature residues characteristic of the large family of response regulators. In contrast, the sequence of the C-terminal domain, which is predicted to contain a helix–turn–helix (HTH) DNA-binding motif, is conserved only among Spo0A homologues from endospore-forming bacteria (Brown *et al*., 1994). A proteolytic fragment encompassing the N-domain (N-Spo0A) remains a substrate for the phosphorelay, whereas the C-terminal fragment (C-Spo0A) is able to bind to DNA and activate transcription (Ireton *et al*., 1993; Grimsley *et al*., 1994). This implies that the N-domain is an inhibitor of the function of the C-domain in intact Spo0A and that this inhibition is overcome by phosphorylation, at Asp-56. In a previous study, we determined the crystal structure of N-Spo0A from *Bacillus stearothermophilus*, revealing the stereochemical basis of aspartic acid phosphorylation and suggesting a general mechanism of activation in response regulators (Lewis *et al*., 1999). Here, we present the crystal structure of the effector domain of Spo0A, which defines the surfaces on the protein responsible for DNA binding and *trans*-activation and provides a basis for discussing the effects of the numerous and diverse mutations accrued on this molecule.

Results

Structure solution and overall structure

Partial tryptic digestion of Spo0A from *B. stearothermophilus* produces two stable fragments (Muchová *et al*., 1999). Amino-terminal sequencing of these fragments identified their N-termini as Met-1 and Asp-139, and this information coupled with that from mass spectrometry suggests that the fragments comprise residues 1–130 of the N-terminal receiver domain and 139–259 of the C-terminal effector domain. Although these particular fragments have not been characterized in terms of their biochemical functions, analogous experiments with similar tryptic fragments of Spo0A from *B. subtilis* have shown that N-Spo0A is a substrate for the phosphorelay and that C-Spo0A supports transcription activation (Grimsley *et al*., 1994). For high-level expression in *Escherichia coli*, fragments of the *spo0A* gene encoding the N- and C-terminally truncated proteins were amplified by polymerase chain reaction (PCR) methods and cloned into pET-based plasmid vectors. The cloning, purification and crystallization of the C-terminal fragment of Spo0A from *B. stearothermophilus* comprising residues 139–259 has been described (Muchová *et al*., 1999). The amino acid sequences of Spo0A from *B. subtilis* and *B. stearothermophilus* are very similar in the C-terminal domain where they share 111 identities and six conservative substitutions over 117 residues (Fig. 1). Spo0A from *B. subtilis* is eight residues longer than the *B. stearothermophilus* homologue; two of the extra residues are in the receiver domain, and the other six are in the linker region connecting the domains. To facilitate the interpretation of the structural results on C-Spo0A in terms of what is known of Spo0A function, in the following discussion we will use the numbering of residues for the much better characterized Spo0A homologue from *B. subtilis* (Fig. 1).

The structure was solved from crystals of selenomethionine-substituted protein using the method of multiwavelength anomalous dispersion to determine the crystallographic phases. The asymmetric unit of the crystal contains three molecules of C-Spo0A, denoted A, B and C. The intermolecular contacts are not extensive; instead, they are typical of lattice interactions in crystals and consistent with measurements in solution, which suggest that C-Spo0A is a monomer. Residues 228–245 are very poorly defined in the electron density maps for one of the three molecules (molecule B), and they are assumed to be disordered. Similarly, the first

residue at the amino-terminus and the last five residues at the carboxyl-terminus are disordered in one or more of the molecules. Otherwise, the structure is very well defined by the electron density maps, and this is reflected in the quality of the refinement statistics (Fig. 2A; Table 1). The A, B and C molecules can be superimposed with root mean squared deviations of 0.4–0.5 Å in main-chain C_α atomic positions following pairwise least squares overlap of the main-chain atoms of residues 152–225 and 246–262. C-Spo0A is a single-domain molecule with the shape of an ellipsoid of approximate dimensions 45 Å × 35 Å × 30 Å (Fig. 2B). The N- and C-termini trail from the same face of the molecule, the former being connected to the receiver domain in intact Spo0A by a protease-sensitive linker that is highly variable in sequence among Spo0A homologues (Fig. 1). The structure is a largely helical assembly with six α-helices connected by generally short segments of polypeptide. These helices are labelled αA–αF to distinguish them from the helices in N-Spo0A, which are referred to by numerals α1–α5.

In their comparison of the sequences of the effector domains of Spo0A from various spore-forming strains of *Bacillus* and *Clostridium*, Brown *et al.* (1994) noted three regions, I, II and III, which were exceptionally well conserved (Fig. 1). These regions correspond to the AB loop and the N-terminal segment of helix B (region I), the CD loop and helix D (region II) and the EF loop and the N-terminal segment of helix F (region III). These disparate strings of sequence form a cluster towards the upper surface of C-Spo0A, with the AB and EF loops lying parallel to and, in some sense, buttressing helix D (Fig. 2B). Helices C and D form a HTH DNA-binding motif, as was predicted from the sequence (Brown *et al.*, 1994). The HTH is a recurring substructure in DNA-binding proteins and has been observed in the context of a variety of differently folded domains. The first and second of these helices are commonly referred to as 'scaffolding' and 'recognition' helices, respectively, because in the first crystal structures of protein DNA complexes, the first helix seemed to play a structural role as a platform for the second helix, which lies in the major groove of the DNA and mediates sequence-specific binding. We will use these terms even though subsequent structures have shown these descriptions to be imprecise, because both helices can make base-specific contacts to the DNA. The recognition helix in C-Spo0A is longer than was anticipated (Brown *et al.*, 1994). The short side-chains of alanine residues 202 on helix C and 209 on helix D facilitate the close packing of these helices, forming apolar contacts to Val-212 and Pro-199 respectively. Polar residues Asn-206–Thr-207–Thr-208 form the turn of the HTH, the hydroxyl of Thr-207 forming an N-cap for helix D.

The HTH has been observed within other all α-helical DNA-binding domains, most notably in the bacteriophage 434 CI and Cro proteins. However, the arrangement of the other helices in the bacteriophage repressors is quite different from that of C-Spo0A. A more extensive structural similarity is seen within the DNA-binding domain of the catabolite gene activator protein (CAP), in which the overlap extends to helix B which precedes the HTH (PDB reference 1cgp). The three helices B, C and D form a compact, three-helical bundle, which packs against the N- and C-terminal α-helices in C-Spo0A, whereas the packing is against a β-sheet in CAP. To investigate the chain topology more systematically, we used the program DALI (Holm and Sander, 1993) to search for proteins in the protein structure database with structural similarity to C-Spo0A. The overall fold of Spo0A is not matched by any other protein whose structure is known, a result that is surprising given the large number of structures that are available (1300 non-redundant/representative entries in the protein structure database). On the other hand, it is consistent with the observation that C-Spo0A has no overall sequence similarity to other proteins.

The most extended structural similarities span segments of just three and sometimes four α-helices. The shared characteristic of the proteins identified is nucleic acid binding. The highest scoring match is with the DNA-binding domain of LexA, a repressor of genes that enable *E. coli* to survive UV irradiation (Fogh *et al.*, 1994). The closely overlapping segment is again the bundle of three helices, B, C and D. The most extended match is contained within the C-terminal domain of the site-specific Cre recombinase of bacteriophage P1 (Guo *et al.*, 1997). Here, the overlap spans four helices, B, C, D and F. Cre recombinase does not have a conventional HTH because 75 residues separate the helices that correspond to the HTH of C-Spo0A. Nevertheless, the orientation of the second of these helices with respect to the major groove of a *loxP* substrate DNA is similar to that seen in other HTH-containing DNA-binding proteins. Other proteins

Fig. 1. Alignment of complete sequences of Spo0A from endospore-forming bacteria. The secondary structure elements observed in the crystal structures of N-Spo0A and C-Spo0A are indicated above the alignment. β-Strands are denoted by arrows and α-helices by bars. The hypervariable linker connecting the two functional domains is denoted by a dashed line. Residues that cluster in the aspartyl pocket of N-Spo0A are marked with asterisks below the alignment. The three conserved regions of C-Spo0A described by Brown *et al.* (1994) are boxed as CRI, CRII and CRIII, as are the residues that comprise the HTH (αC and αD). Sequences are *Bacillus stearothermophilus*, translated from GenBank NID g2654215; *B. subtilis*, g143584; *Bacillus thuringiensis*, g520999; *Bacillus sphaericus*, g497958; *Clostridium pasteurianum*, g497970; *Clostridium difficile*, g1130608; *Clostridium innoculum*, g497968.

```
                          β1          α1            β2         α2
B. stearothermophilus  ------MSIKVCIADDNRELVSLLDEYISSQPDMEVIGTAYNGQDCLQML   44
B. subtilis            -----MEKIKVCVADDNRELVSLLSEYIEGQEDMEVIGVAYNGQECLSLF   45
B. thuringiensis       -----MEKIKVCLVDDNKELVSMLESYVAAQDDMEVIGTAYNGQECLNLL   45
B. sphaericus          -----MTKVKVAIADDNRELLKTMEHYFQGHPEIEIIATASNGKVCLQML   45
C. pasteurianum        ---MEYSKISVLIADDNKEFCNILNDYLLNQSDIVVVGIAKDGIEALKLI   47
C. difficile           MGGFLVEKIKIVLADDNKDFCQVLKEYLSNEDDIDILGIAKDGIEALDLV   50
C. innoculum           ----MNNKMNVYLVDDSMEMIHSMKEALTKSDGYQVIGSATNGEQCVNEL   46
                                     **
                            β3            α3             β4        α4
B. stearothermophilus  EEKRPDILLLDIIMPHLDGLAVLERIRAGF-EHQPNVIMLTAFGQEDVTK   93
B. subtilis            KEKDPDVLVLDIIMPHLDGLAVLERLRESDLKKQPNVIMLTAFGQEDVTK   95
B. thuringiensis       TDKQPDVLVLDIIMPHLDGLAVLEKMRHIERLKQPSVIMLTAFGQEDVTK   95
B. sphaericus          EEYTPDILLLDIIMPHLDGLAVLEAMYQNDRMSSIQVIMLTAFGQEDVMK   95
C. pasteurianum        EEKKPDLVILDIIMPNMDGLVVLEKLANINIDPVPNVIVLSAVGQDKITQ   97
C. difficile           KKTQPDLLILDVIMPHLDGLGVIEKLNTMDIPKMPKIIVLSAVGQDKITQ  100
C. innoculum           RGKHLDVLVLDLIMPKKDGISVLHDLKKNH-IQVDHIICTTPFVNDLIVA   95
                                  *                            *
                            β5           α5
B. stearothermophilus  KAVELGASYFILKPFDMENLAHHIRQVYGKTTPVVRKAAP-------APQ  136
B. subtilis            KAVDLGASYFILKPFDMENLVGHIRQVSGNASSVTHRAPSSQ-SSIIRSS  144
B. thuringiensis       KAVDLGASYFILKPFDMENLTSHIRQVSGKANAMIKRPLP----SFRSAT  141
B. sphaericus          QAVDLGASYFMLKPFEFDQLVQKILHCAGQKASIPKKASV-----LQPTT  140
C. pasteurianum        RAITLGADYYVVKPFDMDVFTKRIRQMFNNTILDSETKKTMPIS-EKAAD  146
C. difficile           SAINLGADYYIVKPFDFVVFINRIRELVSNRVTQVEPKPRPVQETQMTRS  150
C. innoculum           GVQNYRIDYILMKPFEIEQLCDKLNFIVGYTRKEHISSNVLQVN-LDDDE  143
                                    *
                                  αA                αB
B. stearothermophilus  VRDNKPKNLDASITSIIHEIGVPAHIKGYLYLREAIAMVYHDIELLGSIT  186
B. subtilis            QPEPKKKNLDASITSIIHEIGVPAHIKGYLYLREAISMVYNDIELLGSIT  194
B. thuringiensis       TVDGKPKNLDASITSIIHEIGVPAHIKGYMYLREAISMVYNDIELLGSIT  191
B. sphaericus          PQKLNQHQLDSTITAIIKEIGVPAHIKGYSYLREAIQMVFEDIELLGSVT  190
C. pasteurianum        VKISQSVPLDLEDESIIHEIGVPAHIKGYMYLREAINMVVDNIELLSAVT  196
C. difficile           DFVKNVGNIETEITNIIHEIGVPAHIKGYLYLREAMKMVIDNVELLGAVT  200
C. innoculum           KERLQKLELESEITELLHEIGIPAHIKGYMYLRTAILETYLNVDFLGQIT  194
                                                  CR I
                          αC     HTH      αD                αE
B. stearothermophilus  KVLYPDIAKKYNTTASRVERAIRHAIEVAWSRGNLESISSLFGYTVSVSK  236
B. subtilis            KVLYPDIAKKFNTTASRVERAIRHAIEVAWSRGNIDSISSLFGYTVSMTK  244
B. thuringiensis       KVLYPDIAKKYNTTASRVERAIRHAIEVAWSRGNIDSISSLFGYTVSMSK  241
B. sphaericus          KILYPEIAKKFNTTPSRVERAIRHAIEVAWNRGNYEYISSMFGYTEHHLK  240
C. pasteurianum        KELYPSIAKKYNTTASRVERAIRHAIEVAWSRGQVDTINKLFGYTIHNGK  246
C. difficile           KELYPSIAKKFNTTPSRVERAIRHAIEVAWSRGKVDTINQLFGYTVHNTK  250
C. innoculum           KVLYPEIAKKYATTASRVERAIRHAIEVAWNRGNIDAIDDIFGYTISASK  244
                                       CR II
                              αF
B. stearothermophilus  AKPTNSEFIAMVADKLRLEHKAS----------                   259
B. subtilis            AKPTNSEFIAMVADKLRLEHKAS----------                   267
B. thuringiensis       AKPTNSEFIAMVADKLRLEHKAS----------                   264
B. sphaericus          SKPTNSEFIAMIADKIRIDMMAS----------                   263
C. pasteurianum        GKPTNSEFIAMIADKLRLKNKVKNVAQ------                   273
C. difficile           GKPTNSEFIAMIADKLRLEHSMVK---------                   274
C. innoculum           AKPTNSEFIAMISDKLRLEHRLRNKNNVIKTFR                   277
                         CR III
```

© 2000 Blackwell Science Ltd,

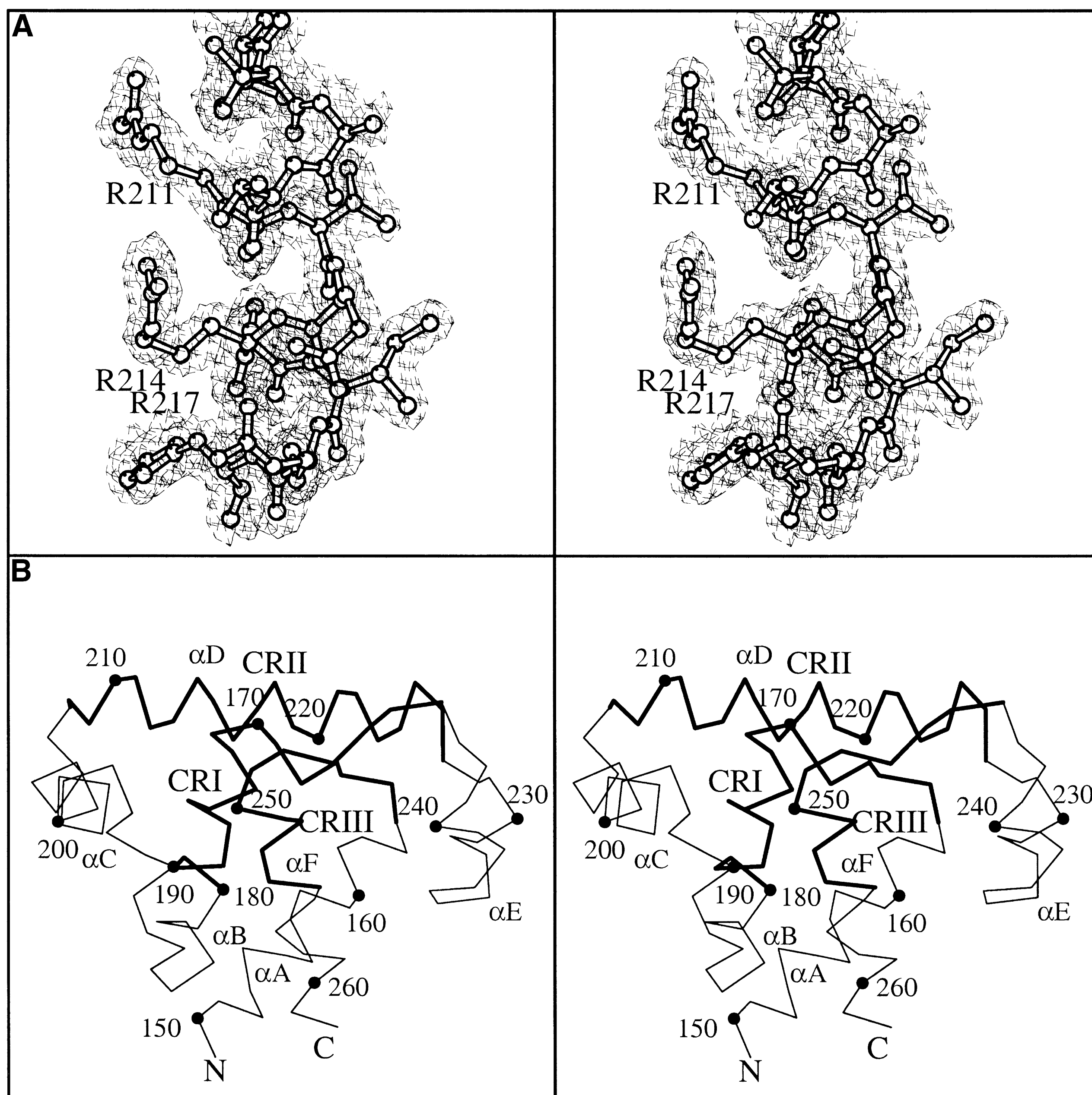


Fig. 2. A. Stereographic view of the final $2F_o-F_c$ refined electron density map displayed at a level of 1 σ. The residues shown, Thr-207 to His-218, are the first 11 residues from conserved region II, i.e. the N-terminus of the 'recognition' helix αD. Bound water molecules have been omitted for clarity. The cluster of three arginines (labelled) is striking.

B. Stereographic C_α trace of C-Spo0A with every 10th residue identified by a small sphere for its C_α atom and numbered according to the *B. subtilis* Spo0A sequence. A thicker line is used to delineate the three conserved regions defined by Brown *et al*. (1994), CRI (residues 164–180), CRII (207–227) and CRIII (244–255).

picked out by DALI as sharing structural similarities over 40 or more residues to C-Spo0A include transposases (1tc3, 2ezk), a helicase (1pjr), a DNA polymerase (1taq), a topoisomerase (1d3y), a nucleotide exchange factor (1pbv) and a Z-DNA-binding protein (1qbj). Most, if not all, of these proteins interact with distorted forms of duplex DNA, which may be significant with regard to DNA binding by C-Spo0A.

The known function of Spo0A as a DNA-binding protein strongly suggests that the predicted HTH is used in DNA binding, although there is no direct experimental evidence to support this assertion. The similarity of the three-helical cluster (helices B, C and D) in C-Spo0A to a similar motif in CAP is consistent with a common mode of DNA binding (Fig. 3). Several mutations in *spo0A* that lead to an

Table 1. Data collection and model refinement statistics.

Data collection statistics	Native data	MAD data		
		λ_1	λ_2	λ_3
Resolution (Å)	2.0	3.2	3.2	3.2
Wavelength (Å)	0.9330	0.9790	0.9792	0.8855
Number of unique reflections	24 473	12 977	12 820	13 011
Redundancy	3.6	2.0	2.0	1.8
Completeness (%)	87.6 (55.6)	98.6 (95.4)	98.6 (95.0)	98.9 (96.9)
$I/(\sigma)I$	19.4 (3.2)	22.8 (13.8)	21.9 (14.4)	19.5 (17.1)
R_{sym} (%)	6.0 (31.3)	3.3 (4.7)	3.4 (5.0)	3.8 (4.9)

Values in parentheses refer to data in the highest resolution shell, 2.07–2.00 and 3.31–3.20 Å for the native and MAD data respectively: the native data are 98.6% complete in the resolution range 20–2.25 Å.

Model refinement statistics	
Number of protein atoms	2659
Number of solvent atoms	323
R_{factor} (no. of reflections)	19.7 (23 085)
R_{free} (no. of reflections)	25.8 (1221)
r.m.s.d. bond lengths (Å)	0.019
r.m.s.d. bond angles (°)	1.9

inability of *B. subtilis* to sporulate map in the region encoding the HTH. In some cases, it has been shown that these mutations severely affect the ability of Spo0A~P to activate and repress transcription (Table 2). Although this is consistent with impaired DNA binding, these mutations have not been characterized fully, and other explanations are possible. A noticeable feature in the structure of C-Spo0A is the three arginine residues (211, 214 and 217) on successive turns of helix D (Fig. 2A). Calculation of the surface electrostatic potential reveals that the C-terminal end of the scaffolding helix, C, and the N-terminal end of the recognition helix, D, are significantly positively charged (Fig. 4). This is again consistent with this surface forming a prominent DNA binding determinant allowing favourable electrostatic contacts to be made to the negatively charged ribose-phosphate backbone of DNA.

The juxtaposition of the HTH with respect to the DNA duplex is variable in crystal structures of protein–DNA complexes. Homeodomain proteins, for example, recognize their B-form DNA targets with residues situated exclusively in the middle of their long 'recognition' helices, whereas λ repressor uses residues from both the 'scaffolding' and the 'recognition' helices to bind in the major groove of DNA, but at markedly different angles with respect to the helical axis of DNA. CAP and λ Cro bend their DNA targets by 90° and 40° respectively. It has been noted that the 0A-box sequences recognized by Spo0A are often flanked by short A+T tracts. AT-rich tracts promote compression of the minor groove and thus DNA bending (Wu and Crothers, 1984). Asayama *et al.* (1995) described how the *spo0F* promoter DNA might be bent by up to 80° on binding of Spo0A~P. Given the different modes of DNA binding by HTH-containing proteins and the likelihood of distortion of the duplex upon Spo0A binding, the structural basis of sequence-specific recognition of 0A-boxes by C-Spo0A cannot be predicted with confidence. Nevertheless, the probable mode of binding of Spo0A to DNA can be deduced from Fig. 3, in which structures of C-Spo0A and the DNA-binding domains of CAP, CI and Cro have been oriented similarly, with the last three illustrated in their complexes with their DNA recognition sequences.

Comparison with other response regulator output domains

The crystal structure of C-Spo0A brings to five the number of response regulator effector domain structures that are known. The other structures are of NarL, which mediates nitrate-dependent regulation of many anaerobic electron transport- and fermentation-related genes (Baikalov *et al.*, 1996), OmpR, which regulates expression of the outer membrane proteins OmpC and OmpF (Kondo *et al.*, 1997; Martínez-Hackert and Stock, 1997), CheB, a methylesterase involved in chemotaxis (Djordjevic *et al.*, 1998), and PhoB, which regulates genes required under conditions of phosphate starvation (Okamura *et al.*, 2000). For NarL and CheB, the structures are of the intact response regulator and illustrate the juxtaposition of receiver and effector domains. The tertiary structures of the *trans*-activation domains of OmpR (C-OmpR) and PhoB (C-PhoB) are very similar, each belonging to the winged-helix family of DNA-binding proteins and consisting of three α-helices in an assembly with two β-sheets (Fig. 5). A similar arrangement of three α-helices occurs

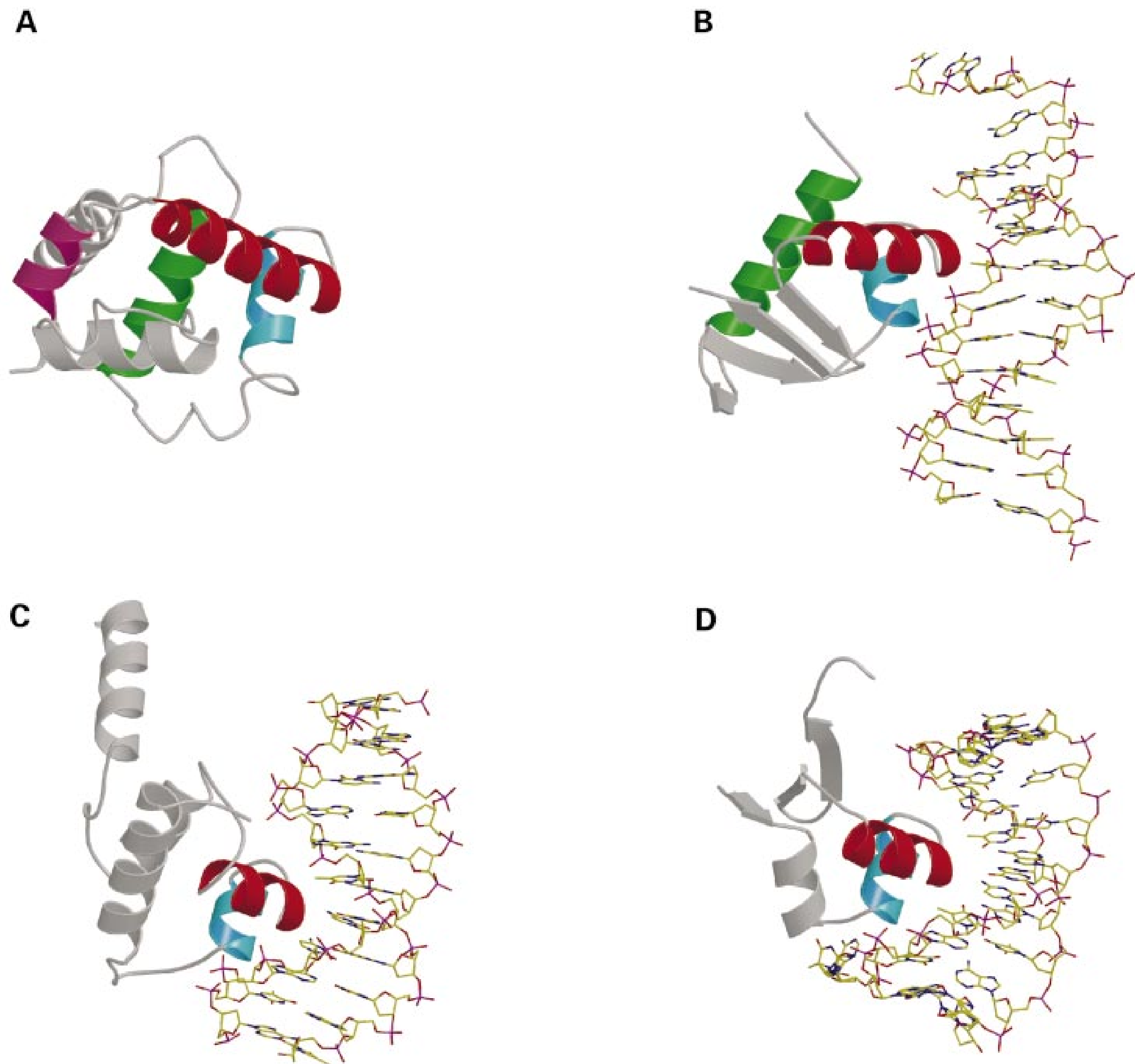


Fig. 3. Domains of representative DNA-binding proteins. (A) C-Spo0A and (B) CAP bound to DNA (1cgp); (C) bacteriophage λ CI repressor bound to DNA (1lmb); and (D) bacteriophage λ Cro bound to DNA (6cro). The figure illustrates how the N-terminus of the recognition helix of the HTH in CAP, Cro and CI dips into the major groove of DNA. It is envisaged that C-Spo0A will bind DNA in a similar manner; however, the precise interactions Spo0A would make with DNA and any resulting distortions of the DNA cannot be predicted with confidence. The proteins are represented as ribbon drawings with side-chains and water molecules omitted. The helices that comprise the HTH are coloured blue (scaffolding helix) and red (recognition helix). The first helix of the three-helical bundle common to C-Spo0A and CAP is in green. DNA is represented as a stick model. The figure suggests that the sigma A-activating region (SAAR) in C-Spo0A, which is coloured pink, is distal to the DNA-binding surface.

in Spo0A and NarL, although the three domains are otherwise quite different (Fig. 5).

Discussion

The study of mutant alleles of *spo0A* has been important in defining the relationship between sequence and function in Spo0A (Table 2). As the ultimate phosphoryl group acceptor of the phosphorelay, a substrate for the protein phosphatase Spo0E and a DNA-binding protein that activates transcription directed by RNA polymerase containing either σ^H (E-σ^H) or σ^A (E-σ^A), the biochemical consequences of mutations in *spo0A* can be complex to unravel. The crystal structures of C-Spo0A and N-Spo0A provide a framework for further understanding Spo0A function and interpreting the rich array of mutational data (Lewis *et al*., 1999; 2000).

Sporulation depends on a dramatic alteration in the profile of gene expression, orchestrated initially by Spo0A. This involves first the downregulation of transcription of genes associated with stationary phase and, secondly, the activation of genes required for sporulation. The whole process depends on the achievement of critical concentrations of Spo0A~P, brought about by increased transcription of *spo0A* and its more efficient phosphorylation

Table 2. Phenotypes associated with *spo0A* mutations.

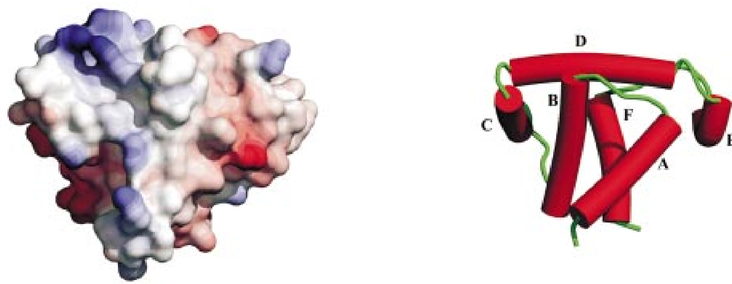
Alteration	Location	Phenotype	Citation
N12K	$\beta1-\alpha1$ loop	*sof102,* suppressor of *spo0F*	Spiegelman *et al.* (1990)
E14K	$\alpha1$	*sof107,* suppressor of *spo0F*	Spiegelman *et al.* (1990)
E14A	$\alpha1$	*sof115,* suppressor of *spo0F*	Spiegelman *et al.* (1990)
P60S	$\beta3-\alpha3$ loop	*sof118,* suppressor of *spo0F*	Spiegelman *et al.* (1990)
D92Y	$\alpha4$	*sof114,* suppressor of *spo0F* and *spoA9V*	Spiegelman *et al.* (1990)
Q121R	$\alpha5$	*sof108,* suppressor of *spo0F*	Spiegelman *et al.* (1990)
F105S	$\beta5$	*sos118,* suppressor of *sof118* (Spo^-)	Spiegelman *et al.* (1990)
P60S	$\beta3-\alpha3$ loop	*coi1,* inappropriate sporulation	Olmedo *et al.* (1990)
A87V	$\beta4-\alpha4$ loop	*coi2,* inappropriate sporulation	Olmedo *et al.* (1990)
Q90K	$\beta4-\alpha4$ loop	*coi15,* inappropriate sporulation	Olmedo *et al.* (1990)
Δ(H61-D75)	$\Delta(\alpha3)$	*sad57,* suppressor of D56N (constitutive)	Ireton *et al.* (1993)
Δ(L62-N81)	$\Delta(\alpha3-\beta4)$	*sad67,* suppressor of D56N (constitutive)	Ireton *et al.* (1993)
Δ(H61-N81)	$\Delta(\alpha3-\beta4)$	*sad76,* suppressor of D56N (constitutive)	Ireton *et al.* (1993)
Δ(D75)	$\alpha3-\beta4$ loop	*sad54,* suppressor of D56N (constitutive)	Ireton *et al.* (1993)
D10Q	$\beta1$	Spo^-	Green *et al.* (1991)
D56Q	$\beta3$	Spo^-	Green *et al.* (1991)
Δ(E20-D28)	$\Delta(\alpha1-\beta2)$	*Δ209*, phosphorylation independent	Green *et al.* (1991)
Δ(E14-E47)	$\Delta(\alpha1-\alpha2)$	*Δ267*, phosphorylation independent	Green *et al.* (1991)
H162R	αA	*suv4,* suppressor of *spo0A9V*	Perego *et al.* (1991)
		Suppressor of S250H	Schmeisser *et al.* (2000)
L174F	αB	*suv3,* suppressor of *spo0A9V*	Perego *et al.* (1991)
		Suppressor of S250H	Schmeisser *et al.* (2000)
D200N	αC	Fails to repress *abrB* and to activate *spoIIA*. Spo^-	F. Schmeisser and I. Barák, unpublished
E213D	αD	Fails to repress *abrB* and to activate *spoIIA*. Spo^-	F. Schmeisser and I. Barák, unpublished
G227R	αE	Fails to activate from σ^A dependent promoters. Spo^-	Hatt and Youngman (1998)
I229A	αE	Lower activation from σ^A dependent promoters	Buckner *et al.* (1998)
D230A	αE	Lower activation from σ^A dependent promoters	Buckner *et al.* (1998)
S231F	αE	Suppresses Spo^- phenotype caused by σ^A H359R	Buckner *et al.* (1998)
I232A	αE	Lower activation from σ^A dependent promoters	Buckner *et al.* (1998)
S233P	αE	Fails to activate from σ^A dependent promoters. Spo^-	Hatt and Youngman (1998)
F236S	αE	Fails to activate from σ^A dependent promoters. Spo^-	Hatt and Youngman (1998)
V240A	αE	Fails to activate from σ^A dependent promoters. Spo^-	Hatt and Youngman (1998)
V240G/K265R	αE, αF	Fails to activate from σ^A dependent promoters. Spo^-	Hatt and Youngman (1998)
S250H	αF	Fails to repress *abrB* and to activate *spoIIA*. Spo^-	Schmeisser *et al.* (2000)
A257V	αF	*spo0A9V*, represses *abrB*, fails to activate *spoIIA*	Perego *et al.* (1991)
A257E	αF	*spo0A153,* represses *abrB*, fails to activate *spoIIA*	Perego *et al.* (1991)
D258V	αF	Represses *abrB*, fails to activate *spoIIA* or *spoIIG* Spo^-	Rowe-Magnus *et al.* (2000)
L260V	αF	Represses *abrB*, fails to activate *spoIIA* or *spoIIG* Spo^-	Rowe-Magnus *et al.* (2000)
Δ(I252-S267)	$\Delta(\alpha F)$	Spo^-	Ferrari *et al.* (1985)

by the phosphorelay. Spo0A~P influences stationary phase gene expression mainly through repression of *abrB* transcription and developmental gene expression through activation at multiple loci. The threshold concentration of Spo0A~P required for these two actions is different (Chung *et al*., 1994). Moderate concentrations of Spo0A~P allow effective repression of *abrB*, which in turn leads to elevated expression of the first sporulation-specific σ-factor, σ^H. Transcription directed by E-σ^H leads to an increase in the concentration of the phosphorelay components, such that a higher proportion of the cellular Spo0A can become phosphorylated. This facilitates the attainment of a second threshold because Spo0A~P stimulates E-σ^H-directed transcription of *spo0A*. At this point, stage II *spo* genes can be activated, and the cell is primed for sporulation.

Transcription activation

The initiation of transcription by bacterial RNA polymerase involves the binding of the enzyme to the promoter to form a closed complex, which is converted to an open complex by unwinding of the duplex at the transcription start site. Consensus sequences 6 bp in length and centred at −35 and −10 facilitate promoter binding and duplex unwinding respectively. Transcriptional activation by auxiliary factors is required at promoters where the −35 and/or −10 sequences have a poor match to the consensus or where the separation of these signature sequences differs significantly from the optimal spacing of 17 bp. Activators of transcription bind to their cognate operator sequences and either facilitate the binding of RNA polymerase or increase the efficiency with which the closed complex is converted to the transcriptionally competent open complex.

The promoters where Spo0A-dependent transcription activation takes place contain multiple 0A-boxes with generally poorer obedience to the consensus (Spiegelman *et al*., 1995), perhaps accounting for the higher threshold concentrations of Spo0A~P needed for activation (Chung *et al*., 1994). Examination of the *spoIIG* promoter reveals



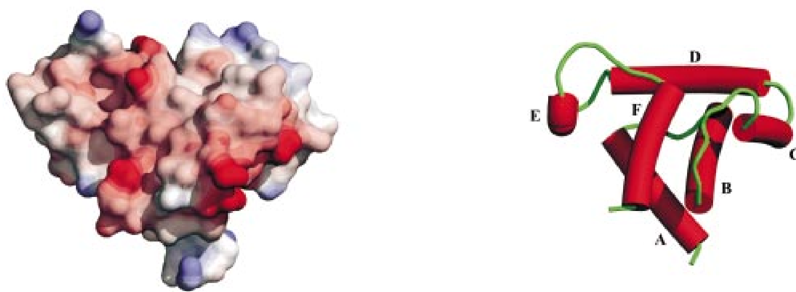
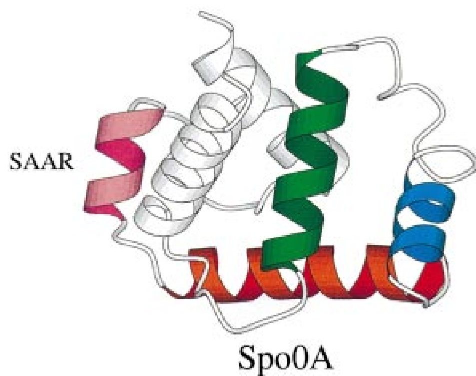


Fig. 4. Electrostatic surface and cartoon representations of C-Spo0A; (top) in the same orientation as Fig. 2 and (bottom) rotated by approximately 180° around a vertical axis. The cartoons are provided to enable the orientation of the molecule to be deduced. Helices appear as red cylinders. In the surface representation, positive charge is coloured blue, negative charge red. The most significant patch of positive charge is located at the N-terminus of the recognition and the C-terminus of the scaffolding helices. The figure was produced using the default settings in MOLVIEWER, a locally written program.



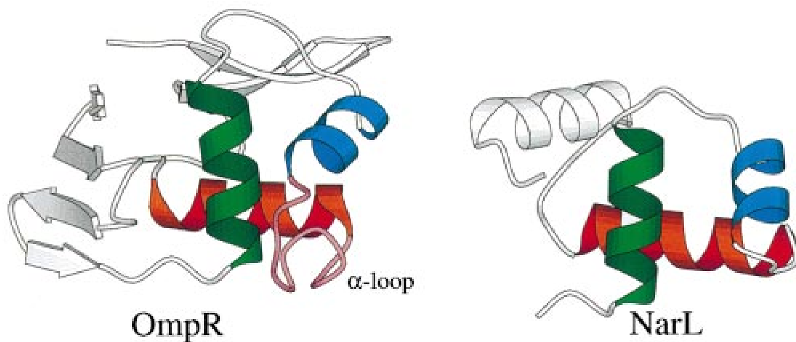


Fig. 5. Comparison of DNA-binding domains from the response regulators Spo0A, OmpR (1opc) and NarL (1rnl). In each structure, the recognition helix runs horizontally and is coloured red, whereas the scaffolding helix runs approximately vertically and is coloured blue. The first helix of the three-helical bundle is coloured green. The α-loop of OmpR and the SAAR of Spo0A are coloured pink.

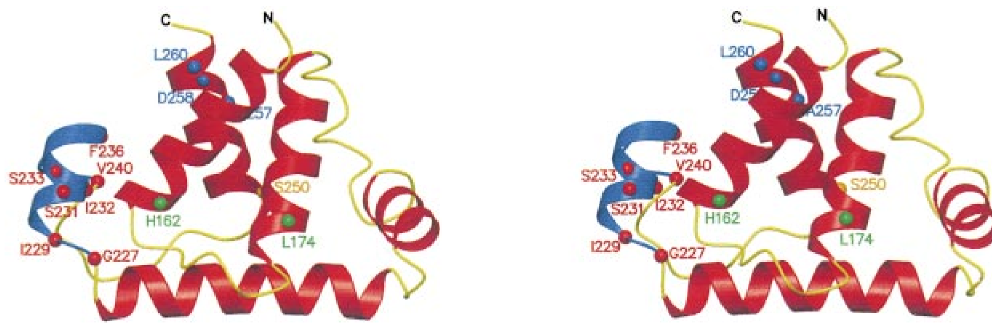


Fig. 6. Ribbon diagram of C-Spo0A showing positions of residues that, when mutated, affect the function of Spo0A. The SAAR region is coloured blue; the rest of the structure is red (helices) and yellow (turns). The C_α atoms of residues whose mutation affects σ^A-dependent activation are coloured red; those that affect both σ^A- and σ^H-dependent activation are in blue. The C_α atoms at the sites of the two *suv* mutations are in green. The C_α atom of residue 250, at which mutation has been reported to affect both activation and repression, is coloured orange.

two tandemly repeated 0A-boxes upstream of the transcription start site. One of these 0A-boxes overlaps the −35 site, part of the promoter that is specifically recognized by conserved residues of RNA polymerase sigma factors (in a region termed 4.2; Busby and Ebright, 1994). A conserved 0A-box is also situated near the −35 site of the σ^A-dependent *spoIIE* promoter. Multiple 0A-boxes are present adjacent to and upstream of the −35 site of the σ^H-dependent *spoIIA* promoter. Spo0A appears to be a class II transcription activator binding at or near to the −35 site to make direct contact with the sigma subunit of RNA polymerase. Instead of recruiting RNA polymerase to the promoter region, class II activators modify prebound RNA polymerase–DNA complexes (Busby and Ebright, 1994).

Activation from σ^A-dependent promoters

Transcription of *spoIIE* and *spoIIG* is directed by E-σ^A and requires Spo0A. RNA polymerase will bind to DNA fragments spanning the −35 region of the *spoIIG* promoter in the absence of Spo0A~P (Bird *et al.*, 1996). However, the −10 region is not protected in these complexes from DNase I cleavage, nor does unwinding of DNA take place adjacent to the start site (Rowe-Magnus and Spiegelman, 1998). The gap between the −35 and −10 consensus sequences is 22 bp, longer than the optimal 17 bp. Transcription initiation does not take place from this complex unless Spo0A~P is added (Rowe-Magnus and Spiegelman, 1998). Binding of Spo0A~P to the proximal −35 0A-box induces stronger binding of RNA polymerase to the −10 site and unwinding of the duplex in the −13 to −3 region.

To investigate the transcription–activation properties of Spo0A further, the coding region of *spo0A* has been subjected to random PCR mutagenesis (Hatt and Youngman, 1998). Mutations of eight residues clustered in a contiguous 14-residue segment (227 and 240) led to defective transcription from the *spoIIE* promoter (Table 2). These mutations had no effect on transcription from the *spoIIA* promoter, which is directed by E-σ^H, nor was repression of *abrB* affected (Buckner *et al.*, 1998; Hatt and Youngman, 1998). Several aspects of this σ^A-activating region (SAAR) are striking in the crystal structure (Figs 3 and 6). First, it is remote from the putative surface used for DNA binding, being situated at the opposite end of the DNA recognition helix (Figs 2B, 3 and 6). Residues 229–236, which form helix E, are followed by a sharp turn at Gly-237 enabling the polypeptide to fold back onto itself and giving the impression that the SAAR forms a loosely tethered loop. The C_α atoms of two residues flanking the SAAR, Trp-224 and Lys-246, are only 6.5 Å apart, suggesting how an insertion in *spo0A* could adapt an ancestral repressor protein into an activator of transcription.

Secondly, these residues form a somewhat negatively charged surface, suggesting the possibility of favourable electrostatic interactions between Spo0A and σ^A in the transcription initiation complex (Fig. 4). This is significant, because two mutations in *sigA*, which lead to defects in transcription at Spo0A-dependent, σ^A-dependent promoters, but not at Spo0A-independent, σ^A-dependent promoters, map to a region encoding a highly positively charged sequence, 352–KALRKLRH–359 (Schyns *et al.*, 1997; Buckner *et al.*, 1998; Hatt and Youngman, 1998). These mutations, K356E and H359R, are located in region 4.2 of σ^A and have a sporulation-deficient phenotype. The mutation S231F in Spo0A suppresses the Spo$^-$ phenotype caused by the H359R mutation in σ^A (Schyns *et al.*, 1997). This suppression is not allele specific: S231F also partially suppresses the effects of H359A and K356E substitutions in σ^A on Spo0A-dependent transcription. Spo0A(S231F) also efficiently stimulates wild-type E-σ^A. This suggests that phenylalanine at position 231 directly or indirectly establishes a new interaction with E-σ^A.

Thirdly, the SAAR is clearly highly mobile; it is disordered in one of the three molecules of the asymmetric unit of the crystals. In the other two molecules, these residues have higher than average temperature

factors, and the SAAR appears to be rather loosely attached to the rest of the domain. It seems that structural plasticity, either in the *trans*-acting element or in RNA polymerase itself, is important in the positive regulation of transcription. In this context, rigidity introduced into the SAAR by proline may account for defective *trans*-activation in the S233P variant of Spo0A; the more conservative S233A mutation has no effect on transcription (Hatt and Youngman, 1998).

Mutagenesis studies have also delineated regions of C-OmpR and C-PhoB that interact with RNA polymerase. These cluster in the extended loop that connects the scaffolding helix to the recognition helix (Fig. 5). This has been termed the activation domain, or α-loop, and is not present in other members of the winged-helix family. In the case of C-OmpR, interactions of the α-loop with the α-subunit of RNA polymerase are required for transcription activation from OmpR-dependent promoters (Igarashi *et al.*, 1991). However, residues in the corresponding region of PhoB contact region 4.2 of σ^{70} (Makino *et al.*, 1993). In contrast, the SAAR in C-Spo0A is adjacent to the C-terminus of the recognition helix αD (Fig. 5). Like the SAAR in Spo0A, the α-loop of OmpR and the corresponding segment in PhoB are flexible in comparison with the rest of their structures.

The σ^H-activating region

The interactions of Spo0A with E-σ^H have similarities to, and differences from, those with E-σ^A. As for σ^A, mutations in region 4.2 of σ^H specifically decrease Spo0A~P-dependent stimulation of transcription directed by E-σ^H. Thus, the Q201A and R205A substitutions in σ^H reduce Spo0A-dependent transcription by E-σ^H from the *spoIIA* promoter without affecting transcription by E-σ^H from the *spoVG* promoter, which is Spo0A-independent (Buckner and Moran, 1998). It seems therefore that Spo0A-dependent transcription requires homologous regions of σ^A and σ^H. A matching region of *E. coli* σ^{70} is implicated in transcription activation by λ repressor (Kuldell and Hochschild, 1994), FNR and AraC (Lonetto *et al.*, 1998). However, none of the mutations identified in the SAAR affected E-σ^H-dependent transcription, nor was S231F, which restores stimulation of transcription to region 4.2 mutants of σ^A, able to suppress the effects of the region 4.2 mutants in σ^H. This suggests the possibility that different regions of Spo0A are involved in activation of transcription by E-σ^H and E-σ^A. It has proved difficult to identify mutants in Spo0A that are specifically defective at σ^H-dependent promoters, probably because the two proteins regulate each other's expression (Weir *et al.*, 1991).

Deletion of the C-terminal 15 residues of Spo0A leads to defective sporulation (Ferrari *et al.*, 1985). The crystal structure shows that this deletion extends into the protein's hydrophobic core (Fig. 6); this will almost certainly lead to loss of Spo0A function because of protein instability. Within this portion of Spo0A are situated the sites of the point mutations *spo0A9V* (A257V) and *spo0A153* (A257E) (Perego *et al.*, 1991). These mutations prevent σ^H-dependent transcription of *spoIIA* without affecting repression of *abrB*, implying that DNA binding is unaffected but that transcription activation is lost. Ala-257 is situated in the hydrophobic protein interior surrounded by residues from helices B and F (including Ile-180, Val-183, Tyr-184, Ile-187, Leu-190 and Ile-253, Ala-254 and Leu-260). To accommodate the valine replacement, some reorganization of the hydrophobic core will be required, probably affecting the packing between these helices. The glutamate substitution would be expected to have more drastic consequences. The role of the C-terminal 10 residues has recently been explored systematically by valine-scanning mutagenesis (Rowe-Magnus *et al.*, 2000). Valine substitution of residues 259 and 261–267 had no effect on sporulation, perhaps consistent with the structural data, which show these residues to be either disordered (263–267) or having side-chains extending away from the core of the molecule (residues K259, R261 and L262). In contrast, the D258V and L260V substitutions severely reduced the sporulation frequency. As noted earlier, Leu-260 packs into the protein core, although Asp-258 is solvent exposed. The A257V, D258V and L260V substitutions were shown in this study to block both σ^A- and σ^H-dependent transcription activation without affecting *abrB* repression (Rowe-Magnus *et al.*, 2000).

Two suppressors of *spo0A9V* (A257V) have been isolated, *suv-4* (H162R) and *suv-3* (L174F) (Perego *et al.*, 1991). No suppressors of *spo0A153* (A257E) were found, suggesting that the introduction of a negative charge at this position cannot be compensated for by other mutations (Fig. 6). *suv-3* and *suv-4* also suppress the Spo$^-$ phenotype caused by the mutation S250H, which neither activates transcription from *spoIIA* nor represses *abrB* (Schmeisser *et al.*, 2000). Surprisingly, the sites of the suppressor mutations (162 and 174) are on the opposite face of C-Spo0A from the sites of the primary mutations. One explanation is that suppression may arise from alteration of intermolecular contacts in a Spo0A dimer. Alternatively, the *suv* mutations may simply generate additional contacts with RNA polymerase (Perego *et al.*, 1991). Accordingly, stimulation of transcription by Spo0A harbouring either of the *suv* mutations from a *spoIIE–lacZ* fusion is somewhat higher than that achieved by wild-type Spo0A (Schmeisser *et al.*, 2000).

Transcription repression

Transcription repression is usually regarded as a simpler

phenomenon than activation. The DNase I footprints of repressor proteins frequently overlap with those of RNA polymerase and, in many instances, it has been shown that repressor and polymerase binding to the promoter are mutually exclusive events. At the *abrB* promoter where Spo0A acts a repressor, there is a tandem orientation of two well-conserved 0A-boxes overlapping the promoter downstream of the start site of transcription (Strauch *et al*., 1990). Steric hindrance is not the key to repression, however, as Spo0A and RNA polymerase (E-σ^A) can bind simultaneously at this promoter (Greene and Spiegelman, 1996). Instead, it appears that Spo0A prevents RNA polymerase from inducing strand denaturation, implying that the two proteins form interactions when bound to *abrB* promoter DNA. These interactions are likely to be different from those involved in activation, as the I229A substitution in the SAAR region of Spo0A has no effect on the repression of *abrB* transcription (Buckner *et al*., 1998).

Phosphorylation-dependent activation

A key question is how does phosphorylation of the receiver domain switch on the *trans*-activation functions of the effector domain? Two plausible mechanisms are: (i) that phosphorylation unmasks surfaces required for gene activation/repression hitherto occluded by the receiver domain; and (ii) that phosphorylation promotes Spo0A dimer formation enabling co-operative binding to pairs of 0A-boxes on the DNA. These possibilities are not mutually exclusive.

At the structural level, our present view of the activation mechanism in Spo0A is as follows. In the presence of the phosphodonor, Spo0B~P, Spo0A mediates its own phosphorylation on Asp-56 in a reaction requiring magnesium cations. The acid phosphate is co-ordinated to the nearby Mg^{2+}, and both species are stabilized by interactions with protein functional groups and water molecules. These include the side-chains of five of the most highly conserved residues in response regulator receiver domains. In particular, the side-chain of Thr-84 moves to form a charge–dipole interaction with the newly arrived phosphoryl moiety, a movement that is tracked by the closely packed side chain of Phe-105, hitherto exposed on the surface of the protein (Lewis *et al*., 1999). This 'aromatic switch' may be a general mechanism for aspartate-phosphate signalling in response regulators (Birck *et al*., 1999; Cho *et al*., 2000). The mechanism by which these structural changes are propagated to the effector domain, however, may be specific to each individual response regulator. There is little or no direct evidence to suggest how interdomain communication takes place in Spo0A. The structures of the receiver and the effector domains do not reveal the conformation of the dozen or so residues connecting the two domains. This segment of the polypeptide is sensitive to proteases, and it is almost certainly flexible.

Despite the availability of high-resolution crystal structures of both domains of Spo0A, several important long-standing questions remain unanswered. First, how do structural changes in the phosphorylated receiver domain bring about activation in the effector domain? In FixJ, the response regulator of symbiotic nitrogen fixation, the switch involves a movement of the aromatic side-chain (Phe-101) to establish a dimer-forming surface (Birck *et al*., 1999). Dimers of *B. subtilis* Spo0A~P have also been reported (Asayama *et al*., 1995), although there are contrary reports (Grimsley *et al*., 1994). In the crystal, N-Spo0A~P exists as a monomer, although this may be due to the influence of the crystallization conditions (Lewis *et al*., 1999). The oligomeric state of Spo0A~P is thus not yet convincingly established. Secondly, how does Spo0A recognize DNA? Thirdly, what is the nature of the interactions between Spo0A and RNA polymerase that bring about activation and repression of transcription? Answers to these questions demand structures of intact Spo0A and Spo0A~P and, ultimately, of larger assemblies including DNA and RNA polymerase holoenzyme.

Experimental procedures

The structure of C-Spo0A was determined by MAD phasing from a single crystal of selenomethionyl-C-Spo0A. To prepare selenomethionyl-C-Spo0A, *E. coli* B834 (DE3) – a methionine auxotroph – was transformed with pETC0ABst, a plasmid that directs the overexpression of C-Spo0A (Muchová *et al*., 1999). Suitable transformants were grown in 25 ml of LB media containing 30 μg ml^{-1} kanamycin, until the A_{600} reached 0.8. These cells were harvested by centrifugation and washed twice with SeMet media before the pellet was used to inoculate 1 l of SeMet media. SeMet media comprises 2 × M9 salts, 2 mM $MgSO_4$, 25 μg ml^{-1} $FeSO_4.7H_2O$, 0.4% glucose, 1 μg ml^{-1} riboflavin, niacinamide, thiamine and pyridoxine monohydrochloride, the 19 common amino acids (except methionine) and seleno-L-methionine, each at a concentration of 40 μg ml^{-1} and, finally, 30 μg ml^{-1} kanamycin. This culture was grown for a further 4–5 h before the addition of IPTG at a final concentration of 1 mM to induce production of C-Spo0A. Four hours later, the cells were harvested by centrifugation, and the cell pellets were stored at −80°C overnight. Purification and crystallization were carried out essentially as reported previously (Muchová *et al*., 1999). The near-complete incorporation of selenomethionine was confirmed by mass spectrometry.

MAD data were collected at beamline BM14 of the ESRF, Grenoble, France, from one crystal at three different wavelengths to 3.2 Å spacing: each data set was individually integrated and reduced using the HKL suite (Otwinowski and Minor, 1997) scaling the Bijvoet pairs separately. Native data to 2.0 Å resolution were collected separately on ID14-EH4.

Data collection statistics are summarized in Table 1. The structure was determined with the program SOLVE in its automatic mode (Terwilliger and Berendzen, 1999). A total of six heavy-atom sites were found corresponding to the two internal methionines 174 and 247 in the three crystallographically independent molecules. MAD phasing from these sites produced an electron density map with a clear boundary between protein and solvent. These phases were improved by solvent flattening, histogram matching and phase extension to 2.2 Å in DM (Cowtan, 1994). This electron density map was of sufficient quality for the NCS operators between the three molecules present in the asymmetric unit to be defined. The phases were improved further by threefold NCS averaging in DM. This electron density map was readily interpretable, allowing us to build almost the entire structure for one molecule, except for residues 228–245 and those at the extreme N- and C-termini. The other two molecules in the asymmetric unit were generated using the NCS operators, and the model was then subjected to refinement in REFMAC (Murshudov *et al.*, 1997). Successive rounds of rebuilding and refinement were interspersed until refinement converged. Statistics for the final model are presented in Table 1.

Acknowledgements

This work has been supported variously by the Wellcome Trust, the BBSRC, the Foundation for Polish Science and the Slovak Academy of Sciences. We thank Drs Garib Murshudov and Marek Brzozowski for helpful discussion, and Dr George Spiegelman for providing data before publication. We are grateful for access to the ESRF, France, and the support of the beamline staff. Atomic co-ordinates and structure factors have been deposited at the RCSB with accession code 1fc3.

References

Asayama, M., Yamamoto, A., and Kobayashi, Y. (1995) Dimer form of phosphorylated Spo0A, a transcriptional regulator, stimulates the *spo0F* transcription at the initiation of sporulation in *Bacillus subtilis*. *J Mol Biol* **250**: 11–23.

Baikalov, I., Schröder, I., Kaczor-Grzeskowiak, M., Grzeskowiak, K., Gunsalus, R.P., and Dickerson, R.E. (1996) Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* **35**: 11053–11061.

Baldus, J.M., Green, B.D., Youngman, P., and Moran, C.P., Jr (1994) Phosphorylation of *Bacillus subtilis* transcription factor Spo0A stimulates transcription from the *spoIIG* promoter by enhancing binding to weak 0A boxes. *J Bacteriol* **176**: 296–306.

Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., *et al.* (1999) Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure* **7**: 1505–1515.

Bird, T.H., Grimsley, J.K., Hoch, J.A., and Spiegelman, G.B. (1996) The *Bacillus subtilis* response regulator Spo0A stimulates transcription of the *spoIIG* operon through modification of RNA polymerase promoter complexes. *J Mol Biol* **256**: 436–448.

Brown, D.P., Ganova-Raeva, L., Green, B.D., Wilkinson, S.R., Young, M., and Youngman, P. (1994) Characterization of Spo0A homologs in diverse *Bacillus* and *Clostridium* species identifies a probable DNA-binding domain. *Mol Microbiol* **14**: 411–426.

Buckner, C.M., and Moran, C.P., Jr (1998) A region in *Bacillus subtilis* σ^H required for Spo0A-dependent promoter activity. *J Bacteriol* **180**: 4987–4990.

Buckner, C.M., Schyns, G., and Moran, C.P., Jr (1998) A region in the *Bacillus subtilis* transcription factor Spo0A that is important for *spoIIG* promoter activation. *J Bacteriol* **180**: 3578–3583.

Burbulys, D., Trach, K.A., and Hoch, J.A. (1991) Initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay. *Cell* **64**: 545–552.

Busby, S., and Ebright, R. (1994) Promoter structure, promoter recognition and transcription activation in prokaryotes. *Cell* **79**: 743–746.

Cho, H.S., Lee, S.-Y., Yan, D., Pan, X., Parkinson, J.S., Kustu, S., *et al.* (2000) NMR structure of activated CheY. *J Mol Biol* **297**: 543–551.

Chung, J.D., Stephanopoulos, G., Ireton, K., and Grossman, A.D. (1994) Gene expression in single cells of *Bacillus subtilis*: evidence that a threshold mechanism controls the initiation of sporulation. *J Bacteriol* **176**: 1977–1984.

Cowtan, K. (1994) 'DM': an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newsl Protein Crystallogr* **31**: 34–38.

Djordjevic, S., Goudreau, P.N., Xu, Q., Stock, A.M., and West, A.H. (1998) Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. *Proc Natl Acad Sci USA* **95**: 1381–1386.

Fawcett, P., Eichenberger, P., Losick, R., and Youngman, P. (2000) The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **97**: 8063–8068.

Ferrari, F.A., Trach, K., Le Coq, D., Spence, J., Ferrari, E., and Hoch, J.A. (1985) Characterisation of the *spo0A* locus and its deduced product. *Proc Natl Acad Sci USA* **82**: 2647–2651.

Fogh, R.H., Ottleben, G., Ruterjans, H., Schnarr, M., Boelens, R., and Kaptein, R. (1994) Solution structure of the LexA repressor DNA-binding domain determined by H-1-NMR spectroscopy. *EMBO J* **13**: 3936–3944.

Green, B.D., Bramucci, M.G., and Youngman, P. (1991) Mutant forms of Spo0A that affect sporulation initiation: a general model for phosphorylation mediated activation of bacterial signal transduction proteins. *Semin Dev Biol* **2**: 21–29.

Greene, E.A., and Spiegelman, G.B. (1996) The Spo0A protein of *Bacillus subtilis* inhibits transcription of the *abrB* gene without preventing binding of the polymerase to the promoter. *J Biol Chem* **271**: 11455–11461.

Grimsley, J.K., Tjalkens, R.B., Strauch, M.A., Bird, T.H., Spiegelman, G.B., Hostomsky, Z., *et al.* (1994) Subunit composition and domain structure of the Spo0A sporulation transcription factor of *Bacillus subtilis*. *J Biol Chem* **269**: 16977–16982.

Guo, F., Gopaul, D.N., and Van Duyne, G.D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **389**: 40–46.

Hatt, J.K., and Youngman, P. (1998) Spo0A mutants of *Bacillus subtilis* with sigma factor-specific defects in transcription activation. *J Bacteriol* **180**: 3584–3591.

Hoch, J.A. (1993) Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* **47**: 441–465.

Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**: 123–138.

Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Aiba, H., Mizono, T., *et al.* (1991) Functional map of the α subunit of *Escherichia coli* RNA polymerase: two modes of transcription activation by positive factors. *Proc Natl Acad Sci USA* **88**: 8958–8962.

Ireton, K., Rudner, D.Z., Siranosian, K.J., and Grossman, A.D. (1993) Integration of multiple developmental signals in *Bacillus subtilis* through the Spo0A transcription factor. *Genes Dev* **7**: 283–294.

Kenney, T.J., Kirshman, P.A., and Moran, C.P., Jr (1988) Gene encoding σ^E is transcribed from a σ^A-like promoter in *Bacillus subtilis*. *J Bacteriol* **170**: 3058–3064.

Kondo, H., Nakagawa, A., Nishihira, J., Nishimura, Y., Mizuno, T., and Tanaka, I. (1997) *Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site. *Nature Struct Biol* **4**: 28–31.

Kuldell, N., and Hochschild, A. (1994) Amino acid substitutions in the −35 recognition motif of σ^{70} that results in defects on phage λ repressor stimulated transcription. *J Bacteriol* **176**: 2991–2998.

Lewis, R.J., Brannigan, J.A., Muchová, K., Barák, I., and Wilkinson, A.J. (1999) Phosphorylated aspartate in the structure of a response regulator protein. *J Mol Biol* **294**: 9–15.

Lewis, R.J., Muchová, K., Brannigan, J.A., Barák, I., Leonard, G., and Wilkinson, A.J. (2000) Domain swapping in the sporulation response regulator Spo0A. *J Mol Biol* **297**: 757–770.

Lonetto, M.A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J Mol Biol* **284**: 1353–1365.

Makino, K., Amemura, M., Kim, S.-K., Nakata, A., and Shinagawa, H. (1993) Role of the σ^{70} subunit of RNA polymerase in transcriptional activation by activator protein PhoB in *Escherichia coli*. *Genes Dev* **7**: 149–160.

Martínez-Hackert, E., and Stock, A.M. (1997) The DNA-binding domain of OmpR: crystal structure of a winged helix transcription factor. *Structure* **5**: 109–124.

Muchová, K., Lewis, R.J., Brannigan, J.A., Offen, W.A., Brown, D.P., Barák, I., *et al.* (1999) Crystallization of the regulatory and effector domains of the key sporulation response regulator Spo0A. *Acta Crystallogr* **D55**: 671–676.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr* **D53**: 240–255.

Okamura, H., Hanaoka, S., Nagadoi, A., Makino, K., and Nishimura, Y. (2000) Structural comparison of the PhoB and OmpR DNA-binding/transactivation domains and the arrangement of PhoB molecules on the phosphate box. *J Mol Biol* **295**: 1225–1236.

Olmedo, G., Ninfa, E.G., Stock, J., and Youngman, P. (1990) Novel mutations that alter regulation of sporulation in *Bacillus subtilis*: evidence that phosphorylation of regulatory protein Spo0A controls the initiation of sporulation. *J Mol Biol* **215**: 359–372.

Otwinowski, Z., and Minor, W. (1997) Processing of diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307–326.

Perego, M., Spiegelman, G.B., and Hoch, J.A. (1988) Structure of the gene for the transition state regulator *abrB*: regulator synthesis is controlled by the *spo0A* sporulation gene in *Bacillus subtilis*. *Mol Microbiol* **2**: 689–699.

Perego, M., Wu, J.-J., Spiegelman, G.B., and Hoch, J.A. (1991) Mutational dissociation of the positive and negative regulatory properties of the Spo0A sporulation transcription factor of *Bacillus subtilis*. *Gene* **100**: 207–212.

Rowe-Magnus, D.A., and Spiegelman, G.B. (1998) Contributions of the domains of the *Bacillus subtilis* response regulator Spo0A to transcription stimulation of the *spoIIG* operon. *J Biol Chem* **273**: 25818–25824.

Rowe-Magnus, D.A., Richer, M.J., and Spiegelman, G.B. (2000) Identification of a second region of the Spo0A response regulator of *Bacillus subtilis* required for transcription activation. *J Bacteriol* **182**: 4352–4355.

Satola, S.W., Kirshman, P.A., and Moran, C.P., Jr (1991) Spo0A binds to a promoter used by σ^A RNA polymerase during sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **88**: 4533–4537.

Schmeisser, F., Brannigan, J.A., Lewis, R.J., Wilkinson, A.J., Youngman, P., and Barák, I. (2000) A new mutation in Spo0A with intragenic suppressors in the effector domain. *FEMS Microbiol Lett* **185**: 123–128.

Schyns, G., Buckner, C.M., and Moran, C.P., Jr (1997) Activation of the *Bacillus subtilis spoIIG* promoter requires interaction of Spo0A and the sigma subunit of RNA polymerase. *J Bacteriol* **179**: 5605–5608.

Spiegelman, G.B., Van Hoy, B., Perego, M., Day, J., Trach, K., and Hoch, J.A. (1990) Structural alterations in the *Bacillus subtilis* Spo0A regulatory protein which suppresses mutations at several *spo0* loci. *J Bacteriol* **172**: 5011–5019.

Spiegelman, G.B., Bird, T.H., and Voon, V. (1995) Transcription regulation by the Spo0A protein of *Bacillus subtilis*. In *Two-Component Signal Transduction*. Hoch, J.A., and Silhavy, T.J. (eds). Washington DC: American Society for Microbiology Press, pp. 159–179.

Strauch, M.A., Webb, V., Spiegelman, G.B., and Hoch, J.A. (1990) The Spo0A protein of *Bacillus subtilis* is a repressor of the *abrB* gene. *Proc Natl Acad Sci USA* **87**: 1801–1805.

Strauch, M.A., Trach, K.A., Day, J., and Hoch, J.A. (1992) Spo0A activates and represses its own synthesis by binding at its dual promoters. *Biochemie* **74**: 619–626.

Strauch, M.A., Wu, J.-J., Jonas, R.H., and Hoch, J.A. (1993) A positive feedback loop controls transcription of the *spo0F* gene, a component of the sporulation phosphorelay in *Bacillus subtilis*. *Mol Microbiol* **7**: 967–974.

Terwilliger, T.C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr* **D55**: 849–861.

Trach, K., Burbulys, D., Strauch, M.A., Wu, J.-J., Dhillion, N., Jonas, R., *et al.* (1991) Control of the initiation of sporulation in *Bacillus subtilis* by a phosphorelay. *Res Microbiol* **142**: 815–823.

Weir, J., Predich, M., Dubnau, E., Nair, G., and Smith, I. (1991) Regulation of *spo0H*, a gene coding for the *Bacillus subtilis* sigma-H factor. *J Bacteriol* **173**: 521–529.

Wu, H.-M., and Crothers, D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature* **308**: 509–513.

Wu, J.-J., Piggot, P.J., Tatti, K.M., and Moran, C.P., Jr (1991) Transcription of the *Bacillus subtilis spoIIA* locus. *Gene* **101**: 113–116.

York, K., Kenney, T.J., Satola, S., Moran, C.P., Jr, H. Poth and Youngman, P. (1992) Spo0A controls the σ^A dependent activation of *Bacillus subtilis* sporulation specific transcription unit *spoIIE*. *J Bacteriol* **174**: 2648–2658.