This is a repository copy of *The trans-activation domain of the sporulation response regulator Spo0A revealed by X-ray crystallography.*

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

---

**Article:**
Lewis, R J, Krzywda, S, Brannigan, J A et al. (5 more authors) (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.
The trans-activation domain of the sporulation response regulator Spo0A revealed by X-ray crystallography

Richard J. Lewis,1 Szymon Krzywdzka,1
James A. Brannigan,7 Johan P. Turkenburg,1
Katarina Muchová,2 Eleanor J. Dodson,1
Imrich Barák2 and Anthony J. Wilkinson1
1Structural Biology Laboratory, Department of Chemistry,
University of York, Heslington, York, YO10 5DD, UK.
2Institute 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 phospho-acceptor 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 αA-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.

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 (Burbulis 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 O σ-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 spoI IA (Trach et al., 1991; Wu et al., 1991), spoII G (Satola et al., 1991; Baldus et al., 1994) and spoI IE (York et al., 1992) operons, which encode forespore- (σ F) and mother cell- (σ M)-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 spoI IA depends on RNA polymerase containing σ H (E-σ H), whereas spoII G and spoI IE 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), kia A (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 (Irelon 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 226–245 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 Cα 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 foxP substrate DNA is similar to that seen in other HTH-containing DNA-binding proteins. Other proteins
Fig. 2. A. Stereographic view of the final 2Fo–Fo 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 Cα trace of C-Spo0A with every 10th residue identified by a small sphere for its Cα 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, 2e2k), 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

© 2000 Blackwell Science Ltd, Molecular Microbiology, 38, 198–212
Table 1. Data collection and model refinement statistics.

<table>
<thead>
<tr>
<th>Data collection statistics</th>
<th>Native data</th>
<th>λ₁</th>
<th>λ₂</th>
<th>λ₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>2.0</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9330</td>
<td>0.9790</td>
<td>0.9792</td>
<td>0.8855</td>
</tr>
<tr>
<td>Number of unique reflections</td>
<td>24 473</td>
<td>12 977</td>
<td>12 820</td>
<td>13 011</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.6</td>
<td>2.0</td>
<td>2.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>87.6 (55.6)</td>
<td>98.6 (95.4)</td>
<td>98.6 (95.0)</td>
<td>98.9 (96.9)</td>
</tr>
<tr>
<td>R₁(α) (%)</td>
<td>19.4 (3.2)</td>
<td>22.8 (13.8)</td>
<td>21.9 (14.4)</td>
<td>19.5 (17.1)</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>6.0 (31.3)</td>
<td>3.3 (4.7)</td>
<td>3.4 (5.0)</td>
<td>3.8 (4.9)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Native data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of protein atoms</td>
<td>2659</td>
</tr>
<tr>
<td>Number of solvent atoms</td>
<td>323</td>
</tr>
<tr>
<td>Rfactor (no. of reflections)</td>
<td>19.7 (23 085)</td>
</tr>
<tr>
<td>Rfree (no. of reflections)</td>
<td>25.8 (1221)</td>
</tr>
<tr>
<td>r.m.s.d. bond lengths (Å)</td>
<td>0.019</td>
</tr>
<tr>
<td>r.m.s.d. bond angles (°)</td>
<td>1.9</td>
</tr>
</tbody>
</table>

inability of *B. subtilis* to sporulate 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 repressor bend their DNA targets by 90° and 40° respectively. It has been noted that the 3A-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 3A-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 H helices in an assembly with two β-sheets (Fig. 5). A similar arrangement of three H helices occurs...
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^A^) 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.

<table>
<thead>
<tr>
<th>Alteration</th>
<th>Location</th>
<th>Phenotype</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N12K</td>
<td>β1−α1 loop</td>
<td>sof102, suppressor of spo0F</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>E14K</td>
<td>α1</td>
<td>sof107, suppressor of spo0F</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>E14A</td>
<td>α1</td>
<td>sof115, suppressor of spo0F</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>P60S</td>
<td>β3−α3 loop</td>
<td>sof118, suppressor of spo0F</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>D92Y</td>
<td>α4</td>
<td>sof114, suppressor of spo0F and spoA9V</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>Q121R</td>
<td>α5</td>
<td>sof108, suppressor of spo0F</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>F105S</td>
<td>β5</td>
<td>sof118, suppressor of sof118 (Spo−)</td>
<td>Spiegelman et al. (1990)</td>
</tr>
<tr>
<td>P60S</td>
<td>β3−α3 loop</td>
<td>coi1, inappropriate sporulation</td>
<td>Olmedo et al. (1990)</td>
</tr>
<tr>
<td>A87V</td>
<td>β4−α4 loop</td>
<td>coi2, inappropriate sporulation</td>
<td>Olmedo et al. (1990)</td>
</tr>
<tr>
<td>Q90K</td>
<td>β4−α4 loop</td>
<td>coi15, inappropriate sporulation</td>
<td>Olmedo et al. (1990)</td>
</tr>
<tr>
<td>∆(H61-D75)</td>
<td>Δα3-x</td>
<td>sad57, suppressor of D56N (constitutive)</td>
<td>Iretin et al. (1993)</td>
</tr>
<tr>
<td>∆(L62-N81)</td>
<td>Δα1-β1</td>
<td>sad67, suppressor of D56N (constitutive)</td>
<td>Iretin et al. (1993)</td>
</tr>
<tr>
<td>∆(H61-N81)</td>
<td>Δα3-β4</td>
<td>sad76, suppressor of D56N (constitutive)</td>
<td>Iretin et al. (1993)</td>
</tr>
<tr>
<td>∆(D75)</td>
<td>Δα3-β4 loop</td>
<td>sad54, suppressor of D56N (constitutive)</td>
<td>Iretin et al. (1993)</td>
</tr>
<tr>
<td>D10Q</td>
<td>β1</td>
<td>Spo+</td>
<td>Green et al. (1991)</td>
</tr>
<tr>
<td>D56Q</td>
<td>β3</td>
<td>Spo+</td>
<td>Green et al. (1991)</td>
</tr>
<tr>
<td>∆(E20-D28)</td>
<td>Δα1-β2</td>
<td>Δ209, phosphorylation independent</td>
<td>Green et al. (1991)</td>
</tr>
<tr>
<td>∆(E14-E47)</td>
<td>Δα1-α2</td>
<td>Δ267, phosphorylation independent</td>
<td>Green et al. (1991)</td>
</tr>
<tr>
<td>H162R</td>
<td>αA</td>
<td>suv4, suppressor of spoA9V</td>
<td>Perego et al. (1991)</td>
</tr>
<tr>
<td>L174F</td>
<td>αB</td>
<td>Suppressor of S250H</td>
<td>Schmeisser et al. (2000)</td>
</tr>
<tr>
<td>D200N</td>
<td>αC</td>
<td>Fails to repress abrB and to activate spoI.A</td>
<td>F. Schmeisser and I. Barak, unpublished</td>
</tr>
<tr>
<td>E213D</td>
<td>αD</td>
<td>Fails to repress abrB and to activate spoI.A</td>
<td>F. Schmeisser and I. Barak, unpublished</td>
</tr>
<tr>
<td>G227R</td>
<td>αE</td>
<td>Fails to activate of α4 dependent promoters. spo−</td>
<td>Hatt and Youngman (1998)</td>
</tr>
<tr>
<td>I229A</td>
<td>αE</td>
<td>Lower activation of α4 dependent promoters</td>
<td>Buckner et al. (1998)</td>
</tr>
<tr>
<td>D283A</td>
<td>αE</td>
<td>Lower activation from α4 dependent promoters</td>
<td>Buckner et al. (1998)</td>
</tr>
<tr>
<td>S231F</td>
<td>αE</td>
<td>Suppresses Spo+ phenotype caused by α4 H359R</td>
<td>Buckner et al. (1998)</td>
</tr>
<tr>
<td>I232A</td>
<td>αE</td>
<td>Lower activation from α4 dependent promoters</td>
<td>Buckner et al. (1998)</td>
</tr>
<tr>
<td>S233P</td>
<td>αE</td>
<td>Fails to activate from α4 dependent promoters. spo−</td>
<td>Hatt and Youngman (1998)</td>
</tr>
<tr>
<td>F236S</td>
<td>αE</td>
<td>Fails to activate from α4 dependent promoters. spo−</td>
<td>Hatt and Youngman (1998)</td>
</tr>
<tr>
<td>V240A</td>
<td>αE</td>
<td>Fails to activate from α4 dependent promoters. spo−</td>
<td>Hatt and Youngman (1998)</td>
</tr>
<tr>
<td>V240G/K265R</td>
<td>αE, αF</td>
<td>Fails to activate from α4 dependent promoters. spo−</td>
<td>Hatt and Youngman (1998)</td>
</tr>
<tr>
<td>S250H</td>
<td>αF</td>
<td>Fails to repress abrB and to activate spoI.A</td>
<td>Schmeisser et al. (2000)</td>
</tr>
<tr>
<td>A257V</td>
<td>αF</td>
<td>spoO49V, represses abrB, fails to activate spoI.A</td>
<td>Perego et al. (1991)</td>
</tr>
<tr>
<td>A257E</td>
<td>αF</td>
<td>spoO4153, represses abrB, fails to activate spoI.A</td>
<td>Perego et al. (1991)</td>
</tr>
<tr>
<td>D258V</td>
<td>αF</td>
<td>Represses abrB, fails to activate spoI.A or spoIG Spo−</td>
<td>Rowe-Magnus et al. (2000)</td>
</tr>
<tr>
<td>L260V</td>
<td>αF</td>
<td>Represses abrB, fails to activate spoI.A or spoIG Spo−</td>
<td>Rowe-Magnus et al. (2000)</td>
</tr>
<tr>
<td>∆(I252-S267)</td>
<td>ΔαF</td>
<td>Spo-</td>
<td>Ferrari et al. (1985)</td>
</tr>
</tbody>
</table>

by the phosphoryl, 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 α-factor, αH. Transcription directed by E-αH leads to an increase in the concentration of the phosphoryl 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-αH-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 *spolIG* 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 (1ml). 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 Cα 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 Cα atoms at the sites of the two suv mutations are in green. The Cα 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 spoIE 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 spoIE 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 spoIE 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 Cα 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–KALRKLHR–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 spo0A promoter without affecting transcription by E-σ^H from the spo0VG 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 Hochchild, 1994), FNR and AraC (Lenton 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 spo0A 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 spo0A 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 spo0E-λAC 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
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 unmask 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$_2$O, 0.4% glucose, 1 µg ml$^{-1}$ riboflavin, niacinamide, thiamine, and pyridoxine monohydrochloride, the 19 common amino acids (except methionine) and seleno-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 nearly 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 Brozowski 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


© 2000 Blackwell Science Ltd, Molecular Microbiology, 38, 198–212


