This is to certify that the
dissertation entitled

TWO REGIONS OF TRANSCRIPTION FACTOR SPOIIID
ALLOW A MONOMER TO BIND DNA

presented by

Paul Richard Himes

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Microbiology and Molecular Genetics

Lee Kroos
Major Professor's Signature

8/19/09
Date

MSU is an Affirmative Action/Equal Opportunity Employer

TWO REGIONS OF TRANSCRIPTION FACTOR SPOIIID ALLOW A MONOMER TO BIND DNA

By

Paul Richard Himes

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Microbiology and Molecular Genetics

2009

ABSTRACT

TWO REGIONS OF TRANSCRIPTION FACTOR SPOIIID ALLOW A MONOMER TO BIND DNA

By

Paul Richard Himes

In response to nutrient limitation, *Bacillus subtilis* develops into two different cell types, a mother cell and a spore. The sporulation of *B. subtilis* is regulated by a cascade of sigma factors in both cell types and communication between the two cells. Regulation of transcription by each of the sigma factors is further modulated by transcription factors. SpoIIID is a key transcription regulator that affects, both positively and negatively, transcription of the regulons of both mother cell sigma factors. SpoIIID has been shown to bind directly to the promoters of a number of the genes it regulates and has been postulated to use a predicted helix-turn-helix motif to mediate binding. Sites bound by SpoIIID contain a 10 bp consensus sequence, and some strongly bound sites contain more than one copy to the consensus, suggesting that SpoIIID interacts with DNA as a dimer or binds cooperatively as monomers.

To better understand the interactions between SpoIIID and DNA, the requirements, both in SpoIIID and DNA, for binding have been characterized. SpoIIID was shown to be able to bind a single copy of its consensus sequence with high affinity as a monomer. There iwas little cooperativity when SpoIIID binds DNA containing multiple matches to its binding site consensus sequence. Analysis of the effects on assays of *in vivo* transcription and *in vitro* binding of DNA by SpoIIID of charge reversal substitutions to residues likely to be on the surface of SpoIIID led to the identification of

two regions essential for DNA binding. In addition to the putative helix-turn-helix motif, a second, C-terminal, basic region was shown to be required for binding to DNA. In having two distinct regions that allow high affinity binding as a monomer, SpoIIID appears to be unique among prokaryotic DNA-binding proteins with a single helix-turn-helix motif.

Further analysis of the results of *in vivo* transcription assays of charge reversal substitutions in SpoIIID in light of the NMR solution structure revealed 3 classes of substitutions that negatively impact transcriptional activation. The first class of substitutions, those affecting residues predicted to be involved in salt bridges or the hydrophobic core, likely decrease the stability of SpoIIID. The second class involves substitutions to residues that are proposed to be involved in DNA-binding. While the negative effects on *in vivo* transcriptional activation by substitutions in the first two classes are readily apparent, the reasons for the activation defect by the third class of substitutions are less clear. The 4 substitutions that make up the third class (D51K, H63E, H68E, and D82K) are in positions that are not believed to interact with DNA or be involved in the structural integrity of SpoIIID. These residues, then, are proposed to be involved specifically in interactions with RNA polymerase to activate transcription.

To my wife, Debbie, who inspires me, and my parents who taught me to do my best and leave the rest to God.

ACKNOWLEDGEMENTS

The path to completion of this thesis has taken many unexpected turns and has turned out to be considerably longer than it appeared at the outset. It could not have been completed at all without the assistance of a large and willing support network.

First, I would like to thank my mentor, Dr. Lee Kroos, who has shown much patience and provided guidance through all the twists and turns and setbacks of my project. He has taught me what it means to do good science and also that there comes a time when, although potential experiments may stretch out forever, you must say enough.

I would also like to thank the members of my committee, Dr. Rob Britton, Dr. Zach Burton, Dr. Wendy Champness, and Dr. Mike Thomashow, for all their helpful advice and for pushing me further in thinking about my project. Dr. Honggao Yan, Dr. Aizhou Liu, and Dr. James Geiger were also instrumental in providing advice regarding the structural aspects of SpoIIID, as were Zhenwei Lu and Stacy Hovde in their labs.

This project would not have been possible without the contributions of Bin Chen, Steve McBryant and William Wedemyer who analyzed the NMR data to solve the structure of SpoIIID, performed the analytical ultracentrifugation to show that SpoIIID bound DNA as a monomer, and modeled (and taught me to model) the structure of SpoIIID, respectively.

Current and previous members of the Kroos lab, who are too numerous to name here, provided their expertise and help in many ways. In particular, I would like to thank Lijuan Wang and John Perpich who also worked on a project involving SpoIIID and had many valuable discussions with me. Dr. Ruanbao Zhou provided many stimulating

conversations and always had a new way of looking at things that challenged how I was thinking.

I would also like to thank my family and friends for their love and support even if they did not always understand what I was doing. To my parents with their constant encouragement no matter what, I owe many thanks.

Last, but definitely not least, I must thank my wife Debbie without whom I would never have finished. Her patience and support has been nothing short of amazing. Words can never express my gratitude for your coming into my life, so I will just have to settle for these three, I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	adenine
Ap^r	ampicillin resistant
Bp	base pair
C	cytosine
cP	centipoise
DNA	deoxyribonucleic acid
Da	dalton
EDTA	(ethylenedinitriol) tetra-acetic acid
EMSAs	electrophoretic mobility shift assays
EQ	sedimentation equilibrium
G	guanine
H(6)	N-terminal hexa-histidine tag
HTH	helix-turn-helix
IPTG	isopropyl β-D-thiogalactopyranoside
K_d	dissociation constant
Kan^r	kanamycin resistant

kDa	kilodalton
LB	Luria-Bertani media
M	molar
Mb	megabases
ml	milliliter
μM	micromolar
N	A, T, G, or C
nM	nanomolar
nm	nanometer
Nm^r	neomycin resistant
NPG	4-nitrophenyl-β-D-glucuronide
PCR	polymerase chain reaction
psi	pounds per square inch
R	A or G
RNAP	RNA polymerase
rpm	revolutions per minute
$s_{20,w}$	sedimentation coefficient in water at 20°C

SDS	sodium dodecyl sulfate
Sp^r	spectinomycin resistant
SV	sedimentation velocity
T	thymine
T7	T7 RNA polymerase promoter and a translation initiation sequence
W	A or T
Y	T or C

Introduction

The Gram-positive model bacterium, *Bacillus* subtilis undergoes a complex development, in response to nutrient limitation, resulting in two different cell types, a mother cell that undergoes programmed cell death and a highly resistant spore. The stages of sporulation are highly regulated to allow proper temporal expression of a large number of genes resulting in a fully resistant spore. This regulation is the function of a cascade of cell-type-specific sigma factors, the activity of which is regulated by communication between cell compartments. Transcriptional regulation by the sigma factors is fine-tuned by temporally-regulated transcription factors in each compartment. A key transcription factor in the mother cell, SpoIIID regulates, both positively and negatively, the transcription of over 100 genes, including genes in controlled by both the early and the late mother cell sigma factors.

Chapter I of this thesis discusses the relevant background literature about sporulation with an emphasis on the events regulating the stages of sporulation in general and SpoIIID in particular.

Chapter II discusses the interactions between SpoIIID and DNA. SpoIIID was shown to interact with a single match to the DNA-binding consensus sequence with high affinity whether one or multiple matches are present and to do so as a monomer. Two regions of SpoIIID were identified that contribute to this binding and each is essential for interactions with DNA. Substitutions reversing the charge of charged residues or introducing a charge on noncharged residues predicted to be on the surface of SpoIIID were used to identify regions of SpoIIID important for activation of transcription *in vivo*.

Chapter III discusses the results of structural determination of SpoIIID and what meaning that provides to the effects of the mutational analysis of SpoIIID described in Chapter II. Three classes of substitutions causing a defect in *in vivo* transcriptional activation are identified, including 4 mutations that may be directly involved in transcriptional activation and may form a surface important for recruitment of RNA polymerase. Future directions for this project are also discussed.

Appendix 1 describes preparation of SpoIIID for the NMR studies used to solve its structure and a comparison between the spectra that result from analysis of SpoIIID in the presence or absence of DNA.

Appendix 2 describes unsuccessful attempts to create a strain bearing a *spoIIID* gene that is not dependent on SpoIIID for its expression.

Chapter I: Literature review

The question of how cells determine their fate by differentiating into distinct cell types has long been of interest in organisms ranging in orders of complexity from bacteria to humans. These decisions involve the coordinated control of complex genetic networks. The Gram-positive soil bacterium *Bacillus subtilis* provides an attractive model in which to examine the control of these decisions because, as it achieves stationary phase and the attendant nutrient limitation, individual cells within the population have the possibility to differentiate into one of five distinct cell types. Probably the most studied cell fate that can be chosen by *B. subtilis* is sporulation to produce a highly resistant dormant spore (Kroos, 2007, Piggot & Hilbert, 2004). Some cells can delay their entry into the sporulation pathway by secreting toxins to cannibalize their sibling cells and use the nutrients to allow for continued growth (Gonzalez-Pastor *et al.*, 2003). Alternatively, *B. subtilis* cells can respond to limited nutrients by becoming competent to uptake DNA, which might provide them with new abilities to thrive in their environment (Dubnau & Provvedi, 2000, Johnsen *et al.*, 2009). In a final option, cells can make a lipopeptide called surfactin which induces neighboring cells to secrete exopolysaccharide and other extracellular matrix materials to form a biofilm (Vlamakis *et al.*, 2008, Lopez *et al.*, 2009). Surfactin-producing cells and matrix producers do not inter convert, but rather co-exist in the biofilm. In each of these cell fates, only a subpopulation of cells becomes the differentiated cell, and entry into the pathway to each differentiation event is typically regulated by a bistable switch (Chai *et al.*, 2008, Dubnau & Losick, 2006). Additionally, vegetatively growing cells can also exist in two types, actively swimming individuals and nonmotile chains, within the same population (Kearns & Losick, 2005).

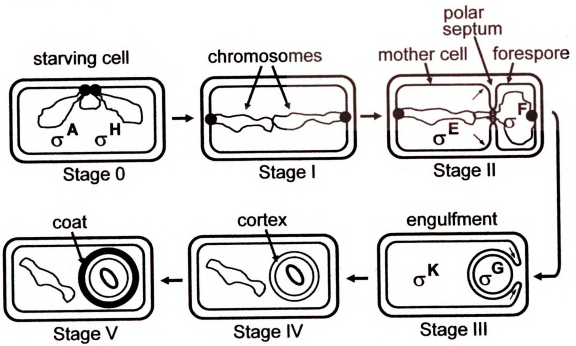
Amongst these cell fate decisions, sporulation provides an interesting example of communication between two different cell types as cells that enter the sporulation pathway divide into two progeny cells that have distinct fates(Hilbert & Piggot, 2004). The larger of the two, the mother cell, initially resides beside, later surrounds, and finally lyses to release the second cell type, the spore, which is highly resistant to a variety of environmental insults, while sensing environmental conditions so that it may germinate into a vegetative cell once favorable conditions return (Setlow, 2003). The process of sporulation has been divided into a series of distinct stages, each identified with a Roman numeral, as determined by their morphological characteristics (Ryter, 1965). Genes in which mutations cause a block in sporulation are identified by the three letters *spo* followed by the Roman numeral indicating the stage at which sporulation was blocked and then a letter to identify which mutation blocking that stage is being discussed (*i.e.*, the gene *spoIIID*, when mutated, results in a block at the third stage of sporulation) (Piggot & Coote, 1976). Similarly, genes which, when mutated, result in a germination defect are named with the three letters *ger*.

An overview of Sporulation

As depicted in Figure 1.1, in the naming scheme describing the stages of sporulation, vegetative growth is denoted as stage 0 (Hilbert & Piggot, 2004). During stage I, the cell has two copies of the chromosome from the previous round of DNA replication, which condense to form an axial filament. The origins of each copy are located at opposite

Figure 1.1. The morphological stages of *B. subtilis* sporulation. The first five stages of sporulation are depicted with the stage at which each sigma factor becomes active are indicated. Stages VI and V, in which the coat matures and the mother cell lyses to release the spore are not depicted. Adapted from (Kroos, 2007).

Fig. 1.1

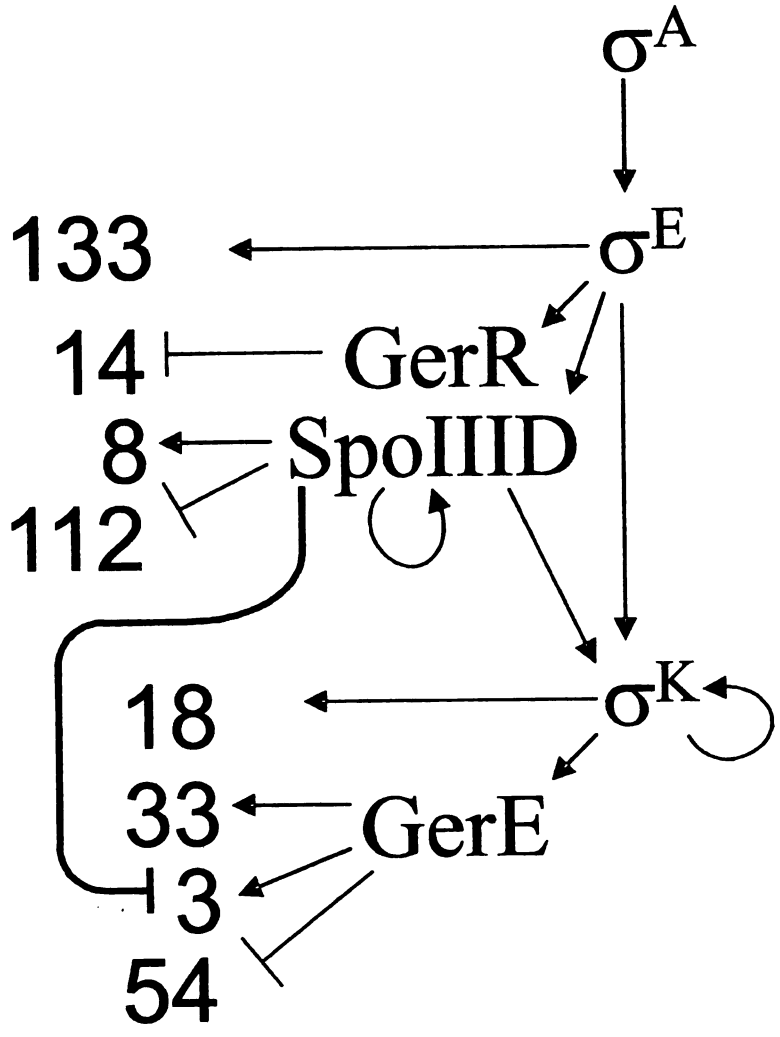

starving cell
chromosomes
polar septum
mother cell
forespore
σA
σH
σE
σF
Stage 0
Stage I
Stage II
coat
cortex
engulfment
σK
σG
Stage V
Stage IV
Stage III

poles of the cell. Near one of the poles, usually the older of the two, the cell divides to form two progeny of unequal volumes. The completion of the septum dividing the two cells is the hallmark of stage II. At the beginning of stage II, two-thirds of the chromosome belonging to the smaller progeny, or forespore, remains in the larger cell, or mother cell, and is actively pumped into the forespore. The septum then bulges toward the interior of the mother cell, and the ends pinch off such that the mother cell engulfs the forespore. Completion of engulfment is designated stage III. During stage IV, two layers of peptidoglycan form between the cell membranes, making the cell wall and cortex. A proteinaceous coat forms around the forespore in Stage V. During stage VI, the forespore matures to gain its full resistance properties. Finally, the mother cell lyses to release the fully resistant spore in Stage VII.

In order for formation of functional spores, each of the above steps has to occur in the proper order. *B. subtilis* sporulation is regulated by a complex network of interconnected regulatory pathways between the mother cell and forespore consisting of sequentially activated sigma factors and transcription factors (Figures 1.1 and 1.2) (Kroos, 2007). During stages 0 and I, σ^A and σ^H are the main sigma factors regulating sporulation-specific transcription, along with the sporulation master regulator Spo0A. After the asymmetric septation that starts stage II, σ^F becomes active in the forespore and causes the activation of σ^E in the mother cell. At this stage, transcription is further modulated by the transcription factors SpoIIID and GerR in the mother cell (Eichenberger *et al.*, 2004), and in the forespore by RsfA (Juan Wu & Errington, 2000). Engulfment in stage

Figure 1.2. Regulation of transcription in the mother cell. The number of genes positively and negatively regulated by each of the sigma factors and transcription factors in the mother cell regulatory pathway is depicted. The numbers indicate the number of genes regulated by each protein. Most of the numbers of genes listed are the results of microarray experiments (Eichenberger et al., 2004) and, thus, may not be an indication of direct activation or repression of transcription by each transcription factor. The design of the experiments did not allow for examination of the effects of SpoIIID on the σ^K regulon, so the data describing those effects comes *in vivo* and *in vitro* transcription assays of individual promoters and may not contain a complete set (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Kroos *et al.*, 1989). Adapted from (Kroos, 2007).

Fig. 1.2



III leads to the activation of σ^G in the forespore, which, in turn, signals for the activation of σ^K in the mother cell. The latter two sigma factors remain active throughout the rest of sporulation in their respective cells with the activity of σ^K further modified by SpoIIID and GerE (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Eichenberger et al., 2004) while SpoVT further regulates genes in the σ^G regulon (Wang *et al.*, 2006). Small, non-coding RNAs have been found to be regulated by Spo0A, σ^G, and σ^K, suggesting that there is a further level of sporulation control yet to be elucidated (Silvaggi *et al.*, 2006).

The sporulation gene regulatory network and the resulting morphological changes are described in more detail in the following sections.

Regulation of the Initiation of Sporulation – Spo0A

B. subtilis uses a series of kinases, KinA, KinB, KinC, KinD, and KinE (Trach & Hoch, 1993, LeDeaux & Grossman, 1995, Jiang *et al.*, 2000b), which, in response to undetermined stimuli, autophosphorylate and initiate a phosphorylation cascade by transferring their phosphates to Spo0F. Spo0F, in turn, transfers its phosphate to Spo0B. Spo0B transfers the phosphate group to Spo0A (Burbulys *et al.*, 1991). Activation (via phosphorylation) of Spo0A is the tipping point in the bistable switch leading to sporulation as active Spo0A leads to the repression of a repressor of σ^H, which stimulates the transcription of KinA, Spo0F and Spo0A, creating a positive feedback loop (Dubnau & Losick, 2006).

The level of phosphorylation of Spo0A is used to integrate a number of signals and modulate its activity and, thereby, entry into sporulation. Sporulation cannot occur until the cells reach a high density. RapA (Perego, 1997), RapB, and RapE (Jiang *et al.*, 2000a), dephosphorylate Spo0F-PO_4 at low cell densities, but are inhibited by small peptide signaling molecules when cell density increases. Nutrient limitation also has an input on levels of Spo0A-PO_4. GTP levels, serving as a stand in for intracellular energy levels, influence the ability of CodY, a repressor of the genes encoding KinB and several of the quorum sensing peptides in the Rap system (Molle *et al.*, 2003b), to bind to DNA (Ratnayake-Lecamwasam *et al.*, 2001).

As well as integrating cues of nutrient availability and cell density, the decision to initiate sporulation is influenced by the state of DNA in the cell. Blocked DNA replication leads to the production of Sda, which inhibits Spo0A activity by blocking autophosphorylation of KinA (Burkholder *et al.*, 2001). Proper chromosomal partitioning is sensed by the Soj protein. In the absence of proper chromosomal partitioning, Soj leads to the loss of Spo0A-PO_4 through Sda-mediated inactivation of KinA (Murray & Errington, 2008).

Spo0A has been shown to bind to and directly regulate 121 genes (Molle *et al.*, 2003a). Two thirds of these genes are negatively regulated. Many of these encode proteins important for vegetative growth. The remaining third that are positively regulated include (in addition to those involved in the positive feedback loop) the operons encoding the first sigma factors specific to the forespore and mother cell (σ^F and σ^E,

respectively) and their regulatory machinery; the gene encoding the protein important for axial filament formation, RacA (Ben-Yehuda *et al.*, 2003); and a protein important for asymmetric septum formation, SpoIIE (Carniol *et al.*, 2005). Spo0A also regulates mother cell-specific genes as it has been shown to be active specifically in the mother cell after asymmetric septation (Fujita & Losick, 2003). Two distinct classes of genes regulated by Spo0A have been identified (Fujita *et al.*, 2005). Genes requiring a low threshold of Spo0A for activation or repression are acted on early because they have high-affinity binding sites for Spo0A. For example, genes involved in cannibalism that delay entry into sporulation fall into this class. Genes acted on later in sporulation by Spo0A have lower affinity binding sites and are affected by Spo0A-PO_4 only when levels rise as cells commit to enter sporulation.

Stage I – Formation of the Axial Filament

Upon the initiation of sporulation, the two newly replicated chromosomes condense and extend across the length of the cell forming a long filament. Formation of this filament is affected by SMC, which is important for chromosome compaction (Britton *et al.*, 1998, Lindow *et al.*, 2002). The chromosomes attach to the opposite poles of the cells via binding by RacA near their respective origins of replication (Ben-Yehuda et al., 2003). RacA is hypothesized to interact with DivIVA, which tethers the cell division repressors MinCD near the poles (Edwards & Errington, 1997). The RacA-DivIVA interaction causes release of MinCD and allows polar septum formation (Ben-Yehuda et al., 2003). Also involved in formation of this axial filament is the DNA binding protein Soj and its partner SpoJ (Wu & Errington, 2003).

Stage II – Asymmetric Cell Division

Asymmetric cell division makes use of much the same machinery, namely FtsZ, FtsA, and MinCD, as is used during vegetative cell division (Hilbert & Piggot, 2004). As with vegetative cell division, an FtsZ ring forms at the center of the cell, but in sporulation then moves in a helical manner toward the poles, assisted by SpoIIE (Ben-Yehuda & Losick, 2002). The disruption of MinCD at the poles, mediated by the RacA-DivIVA interaction, allows polar formation of the FtsZ ring and septum formation. Though FtsZ moves toward both poles, polar septation only occurs at one, usually the older (Hitchins, 1975). Once the septum is formed at one pole, three mother cell proteins (SpoIID, SpoIIM, and SpoIIP) under the regulation of σ^E prevent the formation of a second septum at the other pole (Eichenberger *et al.*, 2001). Because polar septation occurs while the chromosomes are in an axial filament across the length of the cell, only about one third of one chromosome is in the forespore. The remaining two thirds is rapidly pumped across the septum by the SpoIIIE DNA translocase (Wu & Errington, 1998), but, in the 15-20 minutes required to transfer the DNA into the forespore, the transient genetic asymmetry of the two compartments allows for spatial regulation of gene expression (Dworkin & Losick, 2001, Frandsen *et al.*, 1999).

Regulation of early transcription in the forespore – σ^F

The first compartment-specific sigma factors involved in sporulation, σ^F and σ^E, are both made prior to septum formation but are only active in their proper compartments following completion of asymmetric cell division (Hofmeister, 1998, Lewis *et al.*, 1996).

The first cell-specific sigma factor to be activated is σ^F in the forespore. The gene encoding σ^F is the third gene in an operon that also contains the genes for SpoIIAA and SpoIIAB. SpoIIAB functions as an anti-sigma factor, binding to σ^F and keeping it inactive (Decatur & Losick, 1996). Sequestration of σ^F by SpoIIAB is relieved by the anti-anti-sigma factor, SpoIIAA (Schmidt *et al.*, 1990), which binds SpoIIAB in a phosphorylation dependent manner, making it unavailable to bind σ^F (Min *et al.*, 1993).

SpoIIAB, in addition to being able to sequester σ^F, acts as a kinase to phosphorylate and inactivate SpoIIAA. SpoIIE, which localizes to the polar septum, is a phosphatase that acts antagonistically to SpoIIAB by dephosphorylating SpoIIAA (Arigoni *et al.*, 1996, Duncan *et al.*, 1995). The ability of SpoIIE to dephosphorylate SpoIIAA is dependent on completion of the polar septum (King *et al.*, 1999). Because SpoIIE is exclusively localized to the septum and the forepore is at least four times smaller than the mother cell, the effective SpoIIE concentration in the forespore is higher than in the mother cell, which, coupled with the slow kinetics of the reaction catalyzed by SpoIIE, ensures that SpoIIE-dependent dephosphorylation of SpoIIAA is exclusive to the forespore (Iber *et al.*, 2006). In addition to sequestration by SpoIIAA, the ability of SpoIIAB to hold σ^F inactive is further modified by targeted degradation of SpoIIAB by the ClipCP protease (Pan *et al.*, 2001). Because the SpoIIA operon containing σ^F and its cognate anti- and anti-anti-sigma factors is located distal to the origin of replication, initially upon polar septation, these genes are not present in the forespore (Frandsen et al.,

1999). Degradation of SpoIIAB and the temporary inability to replenish it in the forespore, coupled with SpoIIE-dependent activation of SpoIIAA to sequester SpoIIAB allows for the activation of σ^F exclusively in the forespore. A threshold level of unphosphorylated SpoIIAA needs to be achieved before it can relieve SpoIIAB-mediated inactivation of σ^F, and a population of SpoIIAA-SpoIIAB complexes is proposed to act as a reservoir for active SpoIIAA until free SpoIIAB is degraded and the stochiometric balance is shifted in the forespore (Carniol *et al.*, 2004). Because SpoIIAB exists in two conformations (one with high affinity for SpoIIAA and low affinity for σ^F, and one with low affinity for SpoIIAA and high affinity for σ^F), as free SpoIIAA levels increase relative to the concentration of SpoIIAB, the rate of conversion of SpoIIAB to a conformation with high affinity for σ^F decreases due to increased time spent in a complex with SpoIIAA (Iber et al., 2006). While these mechanisms for removing SpoIIAB-mediated inactivation of σ^F are somewhat redundant and a mutation to any one of them has a modest effect on sporulation, a combination of eliminating genetic asymmetry (by moving the *spoIIA* locus near the origin) and eliminating septal localization of SpoIIE causes a drastic reduction in spore formation (Dworkin & Losick, 2001).

The σ^F regulon contains 55 genes that can be divided into two temporally divided classes (Steil *et al.*, 2005). While the protein RsfA is a regulator of genes under the control of σ^F (Juan Wu & Errington, 2000), it does not appear to be the primary agent of this temporal division (Steil et al., 2005), and its major role appears to be regulation of timing of the expression of the signaling protein SpoIIR (Wang et al., 2006). Included in

the genes regulated by σ^F are those encoding SpoIIR, the signal for activation of σ^E in the mother cell (Hofmeister *et al.*, 1995, Karow *et al.*, 1995); σ^G, the sigma factor controlling late events in the forespore (Steil et al., 2005); SpoIIQ, which is important for engulfment and essential for transcription of the gene encoding σ^G (Broder & Pogliano, 2006, Sun *et al.*, 2000); SpoIVB, a protease that leads to activation of σ^K in the mother cell during late sporulation (Zhou & Kroos, 2005); and BofC, a negative regulator of SpoIVB (Wakeley *et al.*, 2000a).

Regulation of early transcription in the mother cell – σ^E

As with σ^F, σ^E is synthesized before the formation of the polar septum, and is maintained in an inactive state until stage II, in this case being held inactive in a pro form and only becoming active upon removal of the N-terminal 27 residues by proteolytic cleavage (LaBell *et al.*, 1987). Pro-σ^E is initially associated with the membrane, but is processed to active σ^E, which is free in the cytoplasm, by SpoIIGA, the first gene in a two gene operon that also contains the gene for pro-σ^E (Jonas *et al.*, 1988, Stragier *et al.*, 1988, Hofmeister, 1998). SpoIIR, transcribed in the forespore under the control of σ^F, is secreted to the inter-septal space where it directly interacts with the membrane-associated N-terminal domain of a SpoIIGA dimer and is proposed to cause a change in conformation, activating the C-terminal protease domain (Imamura *et al.*, 2008). SpoIIR is one of the early class of genes regulated by σ^F (Steil et al., 2005), allowing activation

of σ^E in the mother cell to rapidly follow activation of σ^F in the forespore. Following septum formation, inactive pro-σ^E is present in both compartments, but it is selectively degraded in the forespore (Fujita & Losick, 2002). Because Spo0A is active only in the mother cell after cell division, transcription of the *spoIIG* operon increases in the mother cell (Fujita & Losick, 2003), and active σ^E only accumulates in the mother cell.

The σ^E regulon consists of 171 distinct transcription units (individual genes and operons) that can be organized into three temporally-regulated classes (Eichenberger et al., 2004, Steil et al., 2005). The earliest class of genes includes the genes encoding SpoIID, SpoIIM, and SpoIIP, which comprise the mother cell machinery essential for engulfment (Abanes-De Mello *et al.*, 2002); the SpoIIIA operon, which is essential for the activation of σ^G in the forespore (Illing & Errington, 1991); and GerR (formerly YlbO), which is a repressor of genes in the σ^E regulon (Eichenberger et al., 2004). Of the members of this class, at least SpoIIIA and SpoIID are repressed by the transcription factor SpoIIID, which, along with SpoIVB, which encodes the N-terminus of σ^K (Stragier *et al.*, 1989), is a member of the second class of the σ^E regulon (Steil et al., 2005). The class of genes within the σ^E regulon expressed latest in sporulation includes such genes as *cotE*, which encodes a spore coat protein (Zheng *et al.*, 1988) and is transcribed from multiple promoters (Zheng & Losick, 1990), and *spoVJ*, which has been shown to be activated by RNA polymerase containing both σ^E and σ^K (Foulger &

Errington, 1991), and this class appears to be a convergence of regulons of the two mother cell-specific sigmas (Steil et al., 2005).

Stage III – Engulfment of the Forespore

As described above, the mother cell proteins SpoIID, SpoIIM, and SpoIIP and the forespore protein SpoIIQ are involved in engulfment of the forespore by the mother cell. SpoIID has a peptidoglycan hydrolyase activity and has been proposed to cause the lysis of mother cell wall material starting at the septum, which leads to the forespore bulging into the mother cell at the septum (Abanes-De Mello et al., 2002). All three mother cell proteins localize to the leading edge of cell wall hydrolysis, and continued outward expansion of cell wall hydrolysis appears to drive membrane migration and engulfment. An extracellular domain of SpoIIQ interacts with a membrane-bound mother cell protein, SpoIIIAH, (Blaylock *et al.*, 2004), and this interaction appears to act as a ratchet to prevent the membrane migration from reversing direction (Broder & Pogliano, 2006). In addition, the SpoIIQ/SpoIIIAH interaction has been shown to be important for proper localization of the machinery that regulates the activity of the late mother cell sigma, σ^K (Doan *et al.*, 2005, Jiang *et al.*, 2005). The DNA translocase SpoIIIE is required for the membrane fusion that results in the final free forespore protoplast surrounded by a double membrane inside the mother cell, though this requires a domain of SpoIIIE that is distinct from that required for DNA translocation (Sharp & Pogliano, 1999, Sharp & Pogliano, 2002).

Regulation of late transcription in the forespore – σ^G

Just as σ^F and σ^E are synthesized but held inactive before asymmetric cell division, the late sigma factors, σ^G and σ^K, exist in an inactive form prior to completion of engulfment (Hilbert & Piggot, 2004). While the gene encoding σ^G is transcribed prior to engulfment, it requires an, as yet unidentified, σ^E-dependent signal from the mother cell (Partridge & Errington, 1993). Upon engulfment, a second signal from the mother cell is required to activate σ^G (Errington *et al.*, 1992). The σ^F anti-sigma factor, SpoIIAB, and a σ^G-specific anti-sigma factor, Gin, are able to bind and inactivate σ^G (Karmazyn-Campelli *et al.*, 2008, Kirchman *et al.*, 1993), but these are not sufficient to explain the inactivity of σ^G prior to signaling from the mother cell in the absence of both anti-sigma factors (Camp & Losick, 2008). The negative effects of SpoIIAB on σ^G activation appear to be mainly important for preventing aberrant activation of σ^G in the mother cell (Serrano *et al.*, 2004).

Activation of σ^G upon engulfment requires the vegetative protein SpoIIIJ (Partridge & Errington, 1993), the products of the SpoIIIA operon in the mother cell (Illing & Errington, 1991), and the forespore protein SpoIIQ (Sun et al., 2000). SpoIIIJ is a membrane protein translocase that may be involved in the insertion into the membrane of SpoIIIAE, one of the 8 genes in the SpoIIIA operon (Serrano *et al.*, 2008). SpoIIIJ produced in either the mother cell or the forespore appears to be competent to fulfill this

function (Serrano *et al.*, 2003). At least 6 members of the SpoIIIA operon associate with SpoIIQ from the forespore to form a multimeric complex spanning both membranes that is essential for maintenance of the forespore (Doan *et al.*, 2009). The proteins SpoIIIAH and SpoIIQ appear to be the minimal components of this complex required for signaling from the mother cell (in addition to their role as the ratchet in engulfment discussed above) and form a channel connecting the two cells that may be similar to those in type III secretion (Camp & Losick, 2008, Meisner *et al.*, 2008).

It has been proposed that this channel allows the transport of small molecules essential for transcription and/or translation from the mother cell into the forespore to allow expression of the σ^G operon, as presence of the channel allowed transcription by several different RNA polymerases, including σ^F-directed RNA polymerase and heterologous T7 RNA polymerase, which are unable to function at that time in the absence of the channel (Camp & Losick, 2009). In this model, σ^F would direct transcription from the gene encoding σ^G only at a low level because at the time of transcription the forespore would be depleted of at least one essential nutrient and/or σ^F would bind weakly to the σ^G promoter, and any free σ^G would be bound by the σ^G-specific anti-sigma factor Gin. Upon formation of the SpoIIIA-SpoIIQ channel complex, around the time of engulfment, the limiting nutrient(s) would enter the forespore and allow resumption of transcription. This would allow the concentration of σ^G to exceed that of Gin, and the small amount of free σ^G would direct transcription from the gene encoding σ^G, which is autoregulated (Sun *et al.*, 1991), resulting in a positive feedback loop and a rapid increase in σ^G levels.

Alternatively, presence of a gate on the mother cell face of the channel coupled with the observation that one of the other members of the SpoIIIA operon resembles an ATPase (Meisner et al., 2008), may indicate that a specific factor is actively transported across the channel to relieve a general block to transcription and/or translation in the forespore.

Like the σ^F and σ^E regulons, the σ^G regulon can be divided into early and late classes of genes (Steil et al., 2005). The early class contains genes encoding the transcription regulatory proteins SpoVT and σ^G, in addition to operons such as *spoVA*, which is involved in forespore dipicolinic acid uptake (Tovar-Rojo *et al.*, 2002), and *gerA*, *gerB*, and *gerK*, which encode the receptors leading to germination in response to various stimuli (Setlow, 2003). Many of the small acid-soluble spore proteins (SASPs) which comprise as much as 20% of the spore protein and provide much of the resistance in the spore (Driks, 2002), are encoded by genes in the late class (Steil et al., 2005). In addition to being in the late class of genes in the σ^F regulon as described above, *spoIVB*, which encodes a signal leading to activation of mother cell σ^K (Zhou & Kroos, 2005, Campo & Rudner, 2007, Dong & Cutting, 2003), is also a member of the class genes in the σ^G regulon whose activity peaks late under the influence of the transcription factor SpoVT (Bagyan *et al.*, 1996, Steil et al., 2005). SpoVT is essential for the transcription of 9 genes within the σ^G regulon, and it represses the transcription of 27 others (Wang et al., 2006).

Regulation of late transcription in the mother cell – σ^K

The regulation of σ^K begins with a non-heritable chromosomal recombination event to excise the 48 kb *skin* (sigma K intervening) element to fuse the *spoIVCB* and *spoIIIC* genes, which encode the N- and C-terminal portions of σ^K respectively (Stragier et al., 1989), by the site-specific recombinase SpoIVCA, to form the *sigK* gene (Kunkel *et al.*, 1990), which is transcribed under the direction of σ^E and SpoIIID during the second of the three temporal classes of σ^E-regulated genes (Steil et al., 2005). The newly formed *sigK* gene is transcribed by σ^E- and σ^K-directed RNA polymerase and this also requires SpoIIID (Halberg & Kroos, 1994, Kroos et al., 1989)

Like σ^E, σ^K is initially present as an inactive, membrane-associated, pro-protein (Zhang *et al.*, 1998). Cytoplasmic, active σ^K, which has been truncated by removal of 20 N-terminal amino acids as the end product of a proteolytic cascade (Kroos et al., 1989, Stragier et al., 1989, Zhou & Kroos, 2005, Campo & Rudner, 2007, Dong & Cutting, 2003). Pro-σ^K initially accumulates in the membrane in a complex with SpoIVFB, the intramembrane zinc metalloprotease responsible for cleaving it to the active form (Yu & Kroos, 2000, Rudner *et al.*, 1999); BofA, the inhibitor of SpoIVFB (Zhou & Kroos, 2004); and SpoIVFA, a scaffolding protein for the complex (Rudner & Losick, 2002). This complex has been shown to associate with the SpoIIIAH/SpoIIQ channel that is required for activation of σ^G in the forespore (Doan et al., 2005, Jiang et al., 2005).

SpoIVB, a serine protease transcribed under the direction of both σ^F and σ^G (Wang et al., 2006), serves as the forespore signal to activate σ^K (Gomez *et al.*, 1995, Zhou & Kroos, 2005, Campo & Rudner, 2007, Dong & Cutting, 2003). Activity of SpoIVB in turn is regulated by BofC which delays the autoproteolytic activation of SpoIVB (Wakeley et al., 2000a). Upon secretion into the intermembrane space and subsequent autoproteolysis to form an active state (Wakeley *et al.*, 2000b), SpoIVB cleaves the scaffold protein SpoIVFA in the mother cell membrane (Zhou & Kroos, 2005, Campo & Rudner, 2007, Dong & Cutting, 2003). Cleavage of SpoIVFA exposes BofA to cleavage by CtpB, a serine protease, of which, while produced in both the forespore and the mother cell, the forespore-derived portion is required for proper processing (Campo & Rudner, 2007, Zhou & Kroos, 2005). Cleavage of the inhibitor, BofA, allows SpoIVFB, to process pro-σ^K to active σ^K (Zhou & Kroos, 2004), which is then competent to direct transcription of the genes essential for sporulation late in the mother cell.

In addition to the *sigK* gene itself, the σ^K regulon contains at least one gene essential for cortex formation (Piggot & Coote, 1976), genes encoding spore coat proteins (Zhang *et al.*, 1994), and genes encoding proteins responsible for coat maturation (Kobayashi *et al.*, 1998) and germination (Behravan *et al.*, 2000). Temporally, genes transcribed by σ^K-containing RNA polymerase can be divided into three classes (Steil et al., 2005). The first contains those genes whose levels peak early, many of which are later repressed by the transcription factor GerE including *gerP*, an operon whose products are important for the ability of spore receptors to sense nutrients and trigger germination (Behravan et al.,

2000). For a second subset of the σ^K regulon, including the gene encoding GerE as well as the *cotA* and *cotB* genes, which encode spore coat proteins and are repressed and activated by GerE respectively (Eichenberger et al., 2004), transcript levels continue to rise throughout sporulation (Steil et al., 2005). A final class of genes regulated by σ^K is only transcribed later and includes genes shown to be dependent on GerE. These genes include those encoding the spore coat proteins *cotG* and *cotX* (Sacco *et al.*, 1995, Zhang et al., 1994). In addition to activating σ^K-directed transcription of *sigK*, SpoIIID has been shown to have regulatory effects on the σ^K regulon by negatively regulating a subset of the genes activated by GerE (Ichikawa & Kroos, 2000).

Stage IV – Cortex Formation

The cortex is composed of two distinct layers. The inner layer is similar to that found in vegetative cells and has been proposed to by synthesized from the forespore (Foster & Popham, 2002). The outer layer is synthesized by the mother cell proteins SpoVD and SpoVE (Daniel *et al.*, 1994, Henriques *et al.*, 1992). The outer cortex peptidoglycan contains a modified backbone and is significantly less cross-linked than vegetative peptidoglycan (Atrih *et al.*, 1996). This reduced cross-linking appears to be important for spore heat resistance and germination (Popham *et al.*, 1999).

Stage V – Formation of the Spore Coat

Following synthesis of the cortex, the next stage in sporulation of *B. subtilis* is formation of the proteinaceous coat. The coat has two distinct layers and consists of over 40 different peptides, comprising approximately 30% of total spore protein (Henriques &

Moran, 2007, Lai *et al.*, 2003). The SpoVM protein is crucial for coat assembly and has been shown to help target another protein essential for sporulation SpoIVA to the outer forespore membrane (Levin *et al.*, 1993, van Ooij & Losick, 2003, Price & Losick, 1999, Piggot & Coote, 1976). SpoVM has recently been shown to be targeted specifically to the forespore membrane due to the convex surface of forespore, which is not a type of membrane surface seen in the interior of bacteria (Ramamurthi *et al.*, 2009). SpoIVA recruits CotE and SpoVID, and together the 3 form a basal layer upon which the coat is built (Driks *et al.*, 1994, Ozin *et al.*, 2001). This basal layer matures through an undetermined mechanism to form the inner coat, and CotE then recruits a large number of proteins to form the outer coat (Henriques & Moran, 2007, Little & Driks, 2001, Kim *et al.*, 2006).

Stages VI-VII – Spore Maturation and Mother Cell Lysis

A number of the proteins in the outer layer of the coat have large numbers of cysteine residues, and it is predicted that they are highly crosslinked (Zhang *et al.*, 1993). In addition, a number of irreversible covalent crosslinks between coat proteins are formed, some of which are formed by the transglutaminase Tgl (Suzuki *et al.*, 2000). Currently, GerQ is the only Tgl substrate to be identified (Ragkousi & Setlow, 2004). In addition to its function as a structural member of the spore coat, GerQ is required to recruit CwlJ, which is required for cortex lysis upon germination, to the spore coat (Paidhungat *et al.*, 2001, Ragkousi *et al.*, 2003). Full protein crosslinking by Tgl does not take place until after lysis of the mother cell has occurred (Zilhao *et al.*, 2005). Lysis of the mother cell wall to release the forespore has been shown to be a function of the partially redundant peptidoglycan hydrolyases CwlB, CwlC, and CwlH (Kuroda & Sekiguchi, 1991,

Nugroho *et al.*, 1999, Smith & Foster, 1995). Following mother cell lysis, its DNA is degraded by the NucB nuclease (Hosoya *et al.*, 2007).

Germination

At least 3 spore receptors, GerA, GerB, and GerH, sense and respond to the presence of nutrients in the environment (Setlow, 2003). Activation of the nutrient receptors leads to the release of the spore's store of dipicolinic acid, comprising ~10% of the spore dry weight (Setlow *et al.*, 2008). The release of dipicolinic acid leads to the activation of SleB and CwlJ, cell cortex hydrolase enzymes which specifically degrade the outer cortex but not the inner spore cell wall (Ishikawa *et al.*, 1998, Magge *et al.*, 2008, Boland *et al.*, 2000). Hydrolysis of the cortex allows rehydration of the spore, cell swelling, and resumption of vegetative metabolism (Setlow, 2003).

The Role of SpoIIID in Sporulation

Along with GerR and GerE, SpoIIID is one of three transcription factors that modulates the activity of the mother cell sigma factors during sporulation (Figure 1.2) (Eichenberger et al., 2004). While the activity of the other two is restricted to either the σ^E (GerR) or the σ^K (GerE) regulon, SpoIIID regulates, both positively and negatively, transcription directed by both sigma factors (Halberg & Kroos, 1994). GerR, which acts concurrently with SpoIIID, is a recently identified repressor of 10 transcription units in the σ^E regulon, about which little is known (Eichenberger et al., 2004). In contrast, GerE has been well studied. The structure of GerE, which activates 27 transcription units in the σ^K regulon while repressing 36 (Eichenberger et al., 2004), has been solved to reveal

that it acts as a dimer and binds DNA through the use of a helix-turn-helix motif (Ducros *et al.*, 2001, Crater & Moran, 2001). Two distinct regions of GerE have been identified that are important for the activation of transcription of different genes (*cotC* and *cotX*) (Crater & Moran, 2002, Crater *et al.*, 2002). A corresponding region in σ^K required for the activation of one of those genes (*cotX*) has been identified as well (Wade *et al.*, 1999).

The SpoIIID locus was originally identified in UV-mutagenesis screens as a site of mutations that caused a block at stage III of sporulation (Ionesco *et al.*, 1970, Rouyard *et al.*, 1967). The activity of SpoIIID, which is a 10.8 kDa protein of 93 amino acids with a predicted helix-turn-helix motif (Kunkel *et al.*, 1989), has been shown to be exclusively required in the mother cell (de Lencastre & Piggot, 1979). In addition to activating the transcription of the *sigK* gene by both σ^E and σ^K (Halberg & Kroos, 1994, Kroos et al., 1989), SpoIIID has been shown to be essential for the transcription of 8 genes or operons in the σ^E regulon, while repressing 62 (Eichenberger et al., 2004). SpoIIID might repress more genes in the σ^E regulon, based on microarray results, but the observed negative regulation might instead be due to indirect effects. SpoIIID has also been shown to repress the transcription of at least 3 genes in the σ^K regulon (Halberg *et al.*, 1995, Ichikawa & Kroos, 2000, Kroos et al., 1989). Additionally, SpoIIID has been shown to regulate the transcription of its own gene, with expression from the *spoIIID* promoter being decreased 3- to 7-fold in a strain lacking SpoIIID (Kunkel et al., 1989, Stevens & Errington, 1990).

SpoIIID is co-transcribed with a small upstream gene, *usd*, whose translation is required for the synthesis of SpoIIID, as translation of *usd* is believed to make the Shine-Delgarno sequence of SpoIIID, which is otherwise unavailable in a stem-loop structure, accessible to binding by the ribosome (Decatur *et al.*, 1997). SpoIIID begins to accumulate about 3 hours into sporulation and peaks after about 5 hours, after which point levels decrease (Halberg & Kroos, 1992). The decrease of SpoIIID late in sporulation has been shown to be mediated by negative regulation of σ^E levels by σ^K through two distinct mechanisms (Zhang & Kroos, 1997, Zhang *et al.*, 1999). The reduction of SpoIIID during late sporulation is essential for spores to achieve full heat and lysozyme resistance, though the effects of allowing SpoIIID levels to remain high throughout sporulation can be abrogated by engineering σ^K, whose levels decrease late in sporulation in a GerE-dependent manner, to persist throughout sporulation as well (Wang *et al.*, 2007b, Wang *et al.*, 2007a).

Regulation of transcription of many genes by SpoIIID has been shown to be mediated by direct binding of SpoIIID to DNA (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang *et al.*, 1997, Eichenberger et al., 2004). Multiple binding events by SpoIIID to DNA fragments less than 300 bp in length have been shown for a number of the genes regulated by SpoIIID (Halberg & Kroos, 1994, Eichenberger et al., 2004). Analysis of several sites bound by SpoIIID led to the identification of a WWRRACARNY consensus sequence (where W is A or T, R is purine, Y is pyrimidine, and N is any nucleotide) (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang et al., 1997). While perfect matches to this sequence occur 4840 times in the *B. subtilis* genome, which is 4.2 Mb and

is predicted to contain 4,100 genes (Kunst *et al.*, 1997), at least one sequence known to be bound by SpoIIID is not included. Analysis of a larger set of promoters known and predicted to be bound by SpoIIID was used to create a consensus of which the most common bases at each position are GGACAAG (Eichenberger et al., 2004), a sequence which occurs 582 times in the genome, but is not present in a number of known SpoIIID-bound sequences. The WWRRACARNY consensus sequence is present multiple times in several sequences strongly bound by SpoIIID in DNAse I protection assays, leading to the idea that SpoIIID might bind DNA as a multimer at those sequences (Halberg & Kroos, 1994).

In addition to being a key regulator of development, because SpoIIID is found only in *Bacillus* and *Clostridium* species and their relatives, several of which are pathogens, the study of SpoIIID is interesting as it could lead to the development of specifically targeted inhibitors. Its small size, coupled with the well-characterized genetics of *Bacillus subtilis*, makes SpoIIID an attractive model to increase our knowledge of DNA binding and transcriptional regulation by helix-turn-helix proteins.

In the Chapter 2, I demonstrate that SpoIIID binds to a single match to its binding site consensus as a monomer with high affinity. I also show that binding is mediated through two distinct regions in SpoIIID. In addition to the predicted N-terminal helix-turn-helix motif, a C-terminal region is essential for interaction with DNA. This mechanism of DNA binding appears to be novel among bacterial helix-turn-helix proteins. I also examine the effects of a large number of mutations in *spoIIID* on transcription *in vivo*, as a start toward identifying residues of SpoIIID that are important for transcriptional activation.

In addition to identifying the requirements for the interaction between SpoIIID and DNA, I prepared isotopically labeled SpoIIID, both bound to DNA and free in solution, that was then used to collect NMR data for structural analysis, as described in Appendix I. In Chapter 3, the results of structure determination using these data described along with their implications for a number of the results discussed in Chapter 2. In addition, I will discuss future directions this project could take.

Chapter II: Two regions of transcription factor SpoIIID allow a monomer to bind DNA*

*This chapter is being submitted for publication with Steve J. McBryant of the Department of Biochemistry and Molecular Biology at Colorado State University as an additional author because he performed the analytical ultracentrifugation assay and was instrumental in writing the section regarding the results of that experiment.

Abstract

Nutrient limitation causes *Bacillus subtilis* to develop into two different cell types, a mother cell and a spore. SpoIIID is a key regulator of transcription in the mother cell. It positively or negatively regulates more than 100 genes, in many cases by binding to the promoter region. SpoIIID was predicted to have a helix-turn-helix motif for sequence-specific DNA binding, and a 10-bp consensus sequence was recognized in binding sites, but some strong binding sites were observed to contain more than one match to the consensus sequence, suggesting that SpoIIID might bind as a dimer, or cooperatively as monomers. Here, we show that SpoIIID binds with high affinity as a monomer to a single copy of its recognition sequence, and with little cooperativity to DNA containing more than one match to its binding site consensus sequence. Using charge reversal substitutions of residues likely to be exposed on the surface of SpoIIID, and assays for transcriptional activation *in vivo* and for DNA binding *in vitro*, we identify two regions essential for DNA binding, the putative recognition helix of the predicted helix-turn-helix motif and a basic region near the C-terminus. SpoIIID appears to be unique among prokaryotic DNA-binding proteins with a single helix-turn-helix motif in its ability to bind DNA monomerically with high affinity. We propose that the C-terminal basic region of SpoIIID makes additional contacts with DNA, analogous to the N-terminal arm of eukaryotic homeodomain proteins and the "wings" of winged-helix proteins.

Introduction

In response to nutrient limitation, the Gram-positive bacterium *Bacillus subtilis* undergoes a process of endospore formation to produce a progeny cell that can survive until conditions favorable for growth return (Kroos, 2007). The sporulation process involves creation of two distinct cell types by polar septation; a larger mother cell and a smaller forespore (see Chapter 1, Fig. 1.1). The mother cell engulfs the forespore, so it is surrounded by two membranes. A peptidoglycan cortex forms between the two membranes and a protein coat assembles on the surface of the nascent spore, which upon maturation can withstand a variety of environmental stresses. The process is completed when the mother cell lyses to release the mature spore, which can survive in a metabolically inert state for many years. *Bacilli* and *Clostridia* related to *B. subtilis* also form endospores, which can pose a threat to human health.

For a complex process like sporulation to be successful, a tightly controlled system of temporal and spatial gene regulation is essential. In *B. subtilis* this is achieved in part by a cascade of cell-type-specific σ subunits of RNA polymerase (RNAP). σ^F becomes active first, in the forespore, followed by σ^E in the mother cell, then σ^G in the forespore, and finally σ^K in the mother cell (Losick & Stragier, 1992, Kroos, 2007) (Fig. 1.1). Activation of the forespore σ factors is coupled to morphogenesis, with polar septation leading to activation of σ^F and with completion of engulfment signaling activation of σ^G. The mother cell σ factors are activated in response to signals from the forespore, with σ^F and σ^G activity required to activate σ^E and σ^K, respectively. In the mother cell, temporal regulation of transcription is further controlled by three transcription factors: GerR,

SpoIIID, and GerE (see Chapter 1, Fig. 1.2). GerR appears to act only as a transcriptional repressor of genes in the σ^E regulon early during sporulation (Eichenberger et al., 2004). SpoIIID has been shown to activate and repress transcription of genes in both the σ^E and σ^K regulons, although most of its effects are on genes in the σ^E regulon (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang et al., 1997, Eichenberger et al., 2004). Genome-wide transcriptional profiling of a *spoIIID* mutant in combination with approaches to detect and predict SpoIIID-binding sites in the genome suggests that nearly half of the 272 genes in the σ^E regulon are up- or down-regulated by SpoIIID, in many cases by binding of SpoIIID to the promoter region (Eichenberger et al., 2004). The appearance and disappearance of SpoIIID during sporulation is tightly regulated (Halberg & Kroos, 1992, Zhang & Kroos, 1997, Zhang et al., 1999), and circumventing this regulation in a way that causes the SpoIIID level to remain high late into sporulation results in spore defects (Wang et al., 2007a). GerE has been shown to activate or repress transcription of genes in the σ^K regulon later during sporulation (Eichenberger et al., 2004, Ichikawa *et al.*, 1999, Ichikawa & Kroos, 2000, Zhang et al., 1994, Zheng *et al.*, 1992. GerE crystallized as a dimer and its strongest binding sites in DNA contain inverted repeats matching the consensus sequence RWWTRGGYNNYY (R means A or G, W means A or T, Y means C or T, and N means A, C, G, or T) {Ducros, 2001 #179). Each monomer of the GerE dimer has a helix-turn-helix (HTH) motif whose recognition helix is predicted to make contacts in the major groove of DNA. Thus, GerE appears to achieve high-affinity binding to DNA by a mechanism similar to that of many other HTH-containing DNA-binding proteins (Pabo & Sauer, 1992).

How does SpoIIID achieve high-affinity binding to DNA? SpoIIID has a putative HTH motif (Kunkel et al., 1989) but appeared to be monomeric when purified from sporulating *B. subtilis* (B. Zhang and L.K., unpublished data). Because some of the strongest binding sites for SpoIIID in DNA contain inverted repeats matching the consensus sequence WWRRACARNY, it was suggested that monomers might bind cooperatively at these sites (Halberg & Kroos, 1994). However, here we show that SpoIIID binds with little cooperativity to a site containing inverted repeats and that a SpoIIID monomer is capable of binding to a single copy of the consensus sequence with high affinity using two protein domains. One region is in the predicted HTH motif near the N-terminus of the protein, but the other region is located near the C-terminus. We conclude that SpoIIID, a transcription factor crucial for *B. subtilis* sporulation and highly conserved in other *Bacilli* and *Clostridia* that form endospores, achieves high-affinity DNA-binding by an unusual mechanism.

Experimental Procedures

Plasmids - The plasmids used in this study are described in Table 2.1 and the oligonucleotides used are listed in Table 2.2. Mutations were introduced into the *spoIIID* gene using the QuikChange site-directed mutagenesis kit (Stratagene). All cloned PCR products and all mutant *spoIIID* genes were sequenced at the Michigan State University Genomics Technology Support Facility to confirm that no undesired mutations were present.

Overexpression of SpoIIID – Plasmids designed to overexpress wild-type or altered SpoIIID from a T7 RNAP promoter were transformed into *Escherichia coli* BL21 (DE3)

Table 2.1. Descriptions of certain plasmids used in this study			
Plasmid	Description	Construction	Reference or Source
pMLK83	Ap^r Nm^r; vector to make *gusA* fusions for integration at *amyE*		(Karow et al., 1995)
pAK3	Ap^r Sp^r; vector for integration at *thrC*	a spectinomycin resistance gene was inserted into the XcmI site in the *erm* gene of pDG795 (Guerout-Fleury *et al.*, 1996)	V. Chary and P. Piggot
pET-21b	Ap^r; T7		Novagen
pPH1	Ap^r Sp^r; *spoIIID*	the *spoIIID* promoter (-230 to +303 with respect to translational start) and coding regions were amplified by PCR using LK421 and LK422 as primers and PY79 chromosomal DNA as template, then the EcoRI-digested PCR product was inserted into the EcoRI site of pAK3	This work
pPH4	Ap^r; T7-*spoIIID*	the *SpoIIID* coding region (+2 to +303 with respect to translational start) was amplified by PCR using LK566 and LK567 as primers and PY79 chromosomal DNA as template, then the NdeI and BamHI digested PCR product was inserted between the NdeI and BamHI sites of pET-21b	This work
pPH7	Ap^r Nm^r; $P_{spoIVCA}$-*gusA* transcriptional fusion	the *spoIVCA* promoter region from -137 to +35 (with respect to the start of transcription) was amplified by PCR using LK616 and LK617 as primers and PY79 chromosomal DNA as template, then the SalI-BamHI-digested PCR product was inserted between the SalI and BamHI sites upstream of the *gusA* gene pMLK83	This work

Abbreviations: Ap^r, ampicillin-resistant; Nm^r, neomycin-resistant; Sp^r, spectinomycin-resistant; T7, T7 RNA polymerase promoter and a translation initiation sequence.

Table 2.2. Oligonucleotides used in this study	
Primer	Sequence[a]
LK1579 (H2E)[b] LK1580	5'-gggaggtcgagtggtgtg**ga**ggattacatcaaagagcg-3' 5'-cgctctttgatgtaatc**ct**ccacaccactcgacctccc-3'
LK1581 (D3K) LK1582	5'-ggaggtcgagtggtgtgcac**aaa**tacatcaaagagcgaac-3' 5'-gttcgctctttgatgta**ttt**gtgcacaccactcgacctcc-3'
LK1583 (K6E) LK1584	5'-gtgcacgattacatc**g**aagagcgaacaatcaag-3' 5'-cttgattgttcgctctt**c**gatgtaatcgtgcac-3'
LK1585 (E7K) LK1586	5'-gtgcacgattacatcaaa**a**agcgaacaatcaagatagg-3' 5'-cctatcttgattgttcgc**ttttt**gatgtaatcgtgcac-3'
LK1587 (R8E) LK1588	5'-gtgcacgattacatcaaagag**ga**aacaatcaagataggg-3' 5'-ccctatcttgattgttt**cc**tctttgatgtaatcgtgcac-3'
LK1591 (K14E) LK1592	5'-gaacaatcaagatagggg**g**agtacatcgtggagac-3' 5'-gtctcc**a**cgatgtact**c**cccctatcttgattgttc-3'
LK1593 (E18K) LK1594	5'-gggaagtacatcgtg**a**agacaaagaaaaccgttc-3' 5'-gaacggttttctttgtct**t**cacgatgtacttccc-3'
LK1595 (K20E) LK1596	5'-gtacatcgtggagaca**g**agaaaaccgttcgtg-3' 5'-cacgaacggttttct**c**tgtctccacgatgtac-3'
LK1599 (V23K) LK1600	5'-ggagacaaagaaaaccg**aa**cgtgtcattgcgaagg-3' 5'-ccttcgcaatgacacg**tt**cggttttctttgtctcc-3'
LK1001 (R24A) LK1002	5'-gtggagacaaagaaaaccgtt**gc**tgtcattgcgaaggaatttggtg-3' 5'-caccaaattccttcgcaatgac**agc**aacggttttctttgtctcaac-3'
LK824 (R24E) LK825	5'-gtggagacaaagaaaaccgtt**gaa**gtcattgcgaaggaatttggtg-3' 5'-caccaaattccttcggaatgac**ttc**aacggatttctttgactccac-3'
LK963 (V25E) LK964	5'-gagacaaagaaaaccgttcgtg**aa**attgcgaaggaatttggtg-3' 5'-caccaaattccttcgcaat**tt**cacgaacggttttctttgtctc-3'
LK965 (V25K) LK966	5'-ggagacaaagaaaaccgttcgt**aaa**attgcgaaggaatttggtgtttcc-3' 5'-ggaaacaccaaattccttcgcaat**ttt**acgaacggttttctttgtctcc-3'
LK1603 (I26E) LK1604	5'-gaaaaccgttcgtgtc**gag**gcgaaggaatttggtg-3' 5'-caccaaattccttcgc**ct**cgacacgaacggttttc-3'
LK967 (A27E) LK968	5'-gaaaaccgttcgtgtcattg**a**gaaggaatttggtgtttcc-3' 5'-ggaaacaccaaattccttct**c**aatgacacgaacggttttc-3'
LK826 (K28E) LK827	5'-cgttcgtgtcattgcggaggaatttggtgtttcc-3' 5'-ggaaacaccaaattcct**c**cgcaatgacacgaacg-3'
LK828 (E29K) LK829	5'-cgttcgtgtcattgcgaag**a**aatttggtgtttccaaaagtac-3' 5'-gtacttttggaaacaccaaattt**c**ttcgcaatgacacgaacg-3'

Table 2.2 (cont'd).	
LK 969 (F30A) LK970	5'-gttcgtgtcattgcgaaggaa**gct**ggtgtttccaaaagtacagtac-3' 5'-gtactgtacttttggaaacacc**agc**ttccttcgcaatgacacgaac-3'
LK1683 (G31E) LK1684	5'-gtcattgcgaaggaatttg**agg**tttccaaaagtacagtac-3' 5'-gtactgtacttttggaaac**ctc**aaattccttcgcaatgac-3'
LK1685 (V32E) LK1686	5'-gcgaaggaatttggtg**agt**ccaaaagtacagtacacaag-3' 5'-cttgtgtactgtacttttgga**ctc**accaaattccttcgc-3'
LK830 (S33A) LK831	5'-cgaaggaatttggtgtt**gcc**aaaagtacagtacacaagg-3' 5'-ccttgtgtactgtacttttgg**cgc**aacaccaaattccttcg-3'
LK971 (S33R) LK972	5'-ggttcattgcgaaggaatttggtgtt**cgt**aaaagtacagtacacaaggatttaacag-3' 5'-ctgttaaatccttgtgtactgtacttttt**acg**aacaccaaattccttcgcaatgacac-3'
LK973 (KST-343536AAA) LK974	5'-gtcattgcgaaggaatttggtgtttcc**gctgctgc**agtacacaaggatttaacagagcgtctg-3' 5'-cagacgctctgttaaatccttgtgtact**gcagcagc**ggaaacaccaaattccttcgcaatgac-3'
LK1605 (K34E) LK1606	5'-ggaatttggtgtttcc**g**aaagtacagtacacaagg-3' 5'-ccttgtgtactgtactttc**g**gaaacaccaaattcc-3'
LK1687 (S35E) LK1688	5'-cgaaggaatttggtgtttccaaa**gag**acagtacacaaggatttaac-3' 5'-gttaaatccttgtgtactgt**ctc**tttggaaacaccaaattccttcg-3'
LK1689 (T36E) LK1690	5'-ggaatttggtgtttccaaaagt**ga**agtacacaaggatttaacag-3' 5'-ctgttaaatccttgtgtact**tc**acttttggaaacaccaaattcc-3'
LK975 (VHK-373839AAA) LK976	5'-cgaaggaatttggtgtttccaaaagtacag**ctgctgc**ggatttaacagagcgtctgcctg-3' 5'-caggcagacgctctgttaaatcc**gcagcag**ctgtacttttggaaacaccaaattccttcg-3'
LK1691 (V37E) LK1692	5'-ggtgtttccaaaagtacag**a**acacaaggatttaacag-3' 5'-ctgttaaatccttgtg**tt**ctgtacttttggaaacacc-3'
LK1607 (H38E) LK1608	5'-cgtgtttccaaaagtacagta**gag**aaggatttaacagagcgtc-3' 5'-gacgctctgttaaatcctt**ctc**tactgtacttttggaaacacg-3'
LK1609 (K39E) LK1610	5'-caaaagtacagtacac**g**aggatttaacagagcgtc-3' 5'-gtcgctctgttaaatcct**c**gtgtactgtacttttg-3'
LK977 (DLT-404142AAA) LK978	5'-gtttccaaaagtaagtacagtacacaagg**ctgctg**cagagcgtctgcctgaaattaacccc-3' 5'-ggggttaatttcaggcagacgctctg**cagcag**ccttgtgtactgtacttttggaaac-3'
LK1611 (D40K) LK1612	5'-ccaaaagtacagtacacaag**aaa**ttaacagagcgtctgcctg-3' 5'-caggcagacgctctgttaa**ttt**cttgtgtactgtacttttgg-3'
LK1693 (L41K) LK1694	5'-gtacagtacacaaggat**aa**aacagagcgtctgcctg-3' 5'-caggcagacgctctgtt**tt**atccttgtgtactgtac-3'
LK1695 (T42K) LK1696	5'-cagtacacaaggatttaa**a**agagcgtctgcctg-3' 5'-caggcagacgctct**t**taaatccttgtgtactg-3'

Table 2.2 (cont'd).	
LK1613 (E43K)	5'-gtacacaaggatttaacaaagcgtctgcctgaaattaac-3'
LK1614	5'-gttaatttcaggcagacgctttgttaaatccttgtgtac -3'
LK1615 (R44E)	5'-cacaaggatttaacagaggaactgcctgaaattaaccccg-3'
LK1616	5'-cggggttaatttcaggcagttcctctgttaaatccttgtg-3'
LK1617 (E47K)	5'-cagagcgtctgcctaaaattaaccccgacttg-3'
LK1618	5'-caagtcggggttaatttttaggcagacgctctg-3'
LK1619 (D51K)	5'-ctgcctgaaattaaccccaagttggcaaacgaagtgaaag-3'
LK1620	5'-ctttcacttcgtttgccaacttggggttaatttcaggcag-3'
LK1621 (E55K)	5'-cccgacttggcaaacaaagtgaaagaaatactc-3'
LK1622	5'-gagtatttctttcacttttgtttgccaagtcccc-3'
LK1623 (K57E)	5'-cttggcaaacgaagtggaagaaatactcgattatc-3'
LK1624	5'-gataatcgagtatttcttccacttcgtttgccaag-3'
LK1625 (E58K)	5'-ggcaaacgaagtgaaaaaaatactcgattatcataaatcc-3'
LK1626	5'-ggatttatgataatccagtatttttttcacttcgtttgcc-3'
LK1627 (D61K)	5'-ggcaaacgaagtgaaagaaatactcaaaatatcataaatccatcaggc-3'
LK1628	5'-gcctgatggatttatgatatttgagtatttctttcacttcgtttgcc-3'
LK1629 (H63E)	5'-gtgaaagaaatactcgattatgagaaatccatcaggcatttaagagg-3'
LK1630	5'-cctcttaaatgcctgatggatttctcataatcgagtatttctttcac-3'
LK1631 (K64E)	5'-gaaatactcgattatcatgaatccatcaggcatttaagag-3'
LK1632	5'-ctcttaaatgcctgatggattcatgataatcgagtatttc-3'
LK1635 (H68E)	5'-cataaatccatcagggagttaagaggaggagaagcgac-3'
LK1636	5'-gtcgcttctcctcctcttaactccctgatggatttatg-3'
LK1637 (R70E)	5'-cataaatccatcaggcatttagaaggaggagaagcgac-3'
LK1638	5'-gtcgcttctcctcctctaaatgcctgatggatttatg-3'
LK1639 (E73K)	5'-ggcatttaagaggaggaaaagcgacaaagctc-3'
LK1640	5'-gagctttgtcgcttttcctcctcttaaatgcc-3'
LK1641 (K76E)	5'-ggaggagaagcgacagagctcaaatataaaaaag-3'
LK1642	5'-ctttttatatttgagctctgtcgcttctcctcc-3'
LK979 (K76Stop)	5'-caggcatttaagaggaggagaagcgacataactcaaatataaaaaagatgaaattctcg-3'
LK980	5'-cgagaatttcatctttttatatttgagttatgtcgcttctcctcctcttaaatgcctg-3'
LK1643 (K78E)	5'-gaggagaagcgacaaagctcgaatataaaaaagatg-3'
LK1644	5'-catctttttatattcgagctttgtcgcttctcctc-3'
LK1645 (K80E)	5'-gcgacaaagctcaaatatgaaaaagatgaaattctcgaag-3'
LK1646	5'-cttcgagaatttcatctttttcatatttgagctttgtcgc-3'

Table 2.2 (cont'd).	
LK1647 (K81E) LK1648	5'-gacaaagctcaaatataaa**g**aagatgaaattctcgaagg-3' 5'-ccttcgagaatttcatctt**c**tttatatttgagctttgtc-3'
LK1649 (D82K) LK1650	5'-gacaaagctcaaatataaaaaaa**a**ggaaattctcgaaggagagcctg-3' 5'-caggctctccttcgagaatttc**ctt**ttttttatatttgagctttgtc-3'
LK1651 (D82Stop) LK1652	5'-cgacaaagctcaaatataaaaaat**ag**gaaattctcgaaggagagcctg-3' 5'-caggctctccttcgagaatttc**cta**tttttatatttgagctttgtcg-3'
LK1655 (E86K) LK1656	5'-gatgaaattctc**a**aaggagagcctgttcagc-3' 5'-gctgaacaggctctcctt**t**gagaatttcatc-3'
LK1659 (E88K) LK1660	5'-gaaattctcgaagga**a**agcctgttcagcaatcg-3' 5'-cgattgctgaacaggc**t**ttccttcgagaatttc-3'
LK421 (SpoIIID -230 to +303)[c] LK422	5'-cggaattcagttcgtttcacccttgtc-3' 5'-cgaattccaagaaggcaatgccagg-3'
LK566 (SpoIIID +2 to +303)[c] LK567	5'-gggaattccatatgcacgattacatcaaagagcgaac-3' 5'-cgggatcccaagaaggcaatgccagg-3'
LK616 (SpoIVCA -248 to -59)[c] LK617	5'-acgcgtcgacgaggacctaaaggagtgtag-3' 5'-cgggatccgtcagcttgcggtccctcg-3'

[a] Boldface type indicates location of mutation in sequence.

[b] For each primer pair used in site-directed mutagenesis of pPH4, the substitution in SpoIIID is indicated in parentheses after the name of the primer, the top primer indicates the coding strand, the bottom primer indicates the template strand.

[c] For each primer pair used to clone sequences out of the *Bacillus subtilis* chromosome, the boundaries of the region (in nucleotides, with respect to the start of translation) are indicated in parentheses after the name of the primer, the top primer is the primer complementary to the upstream end of the sequence cloned, and the bottom primer is complementary to the downstream end. Sequences not part of the region cloned are underlined.

(Novagen), a strain that can be induced to synthesize T7 RNAP by the addition of isopropyl β-D-thiogalactopyranoside (IPTG). Transformants were grown on Luria-Bertani (LB) agar (Sambrook *et al.*, 1989) containing 100 μg/ml ampicillin for 10-12 h at 30°C. Small, isolated colonies (as SpoIIID overexpression causes a growth defect in *E. coli*) were selected to inoculate 0.5-1 l of LB liquid medium containing 150 μg/ml ampicillin and the culture was incubated at 37°C with vigorous shaking until it reached an OD_{600} of approximately 1, then IPTG (0.5 mM) was added and the incubation was continued for 2.5 h. Cells were then harvested by centrifugation (7000 x g for 15 min at 4°C) and stored at -70°C.

Purification of SpoIIID - Cells were resuspended in 20 ml of lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 4 mM EDTA) to which 1 Complete Mini EDTA-free protease inhibitor tablet (Roche) had been added and lysed by passage through a French pressure cell at 14,000 psi four times. The lysate was cleared by low-speed centrifugation (7000 x g for 10 min at 4°C) followed by high-speed centrifugation (approximately 175,000 x g for 1 h at 4°C). The supernatant was passed over a 1 ml SP Sepharose column (GE Healthcare) that had been equilibrated with SP column buffer (25 mM Tris-HCl pH 8.0, 0.1% Triton X-100, to which 1 Complete Mini EDTA-free protease inhibitor tablet per 10 ml had been added). The column was washed with 5 ml SP column buffer and SpoIIID was eluted with SP column buffer supplemented with a 0.2 to 0.6 M NaCl gradient. Wild-type SpoIIID and the G87Stop form eluted at ~0.5 M NaCl, while other mutant forms of SpoIIID eluted at different NaCl concentrations (K34E, K39E, and R44E at ~0.3 M; H38E and K76Stop at ~0.4M; S33R, E43K, and D82Stop at ~0.6 M). Identification of fractions containing SpoIIID and purity was

determined by separation on SDS-14% Proseive polyacrylamide gels (Lonza) with Tris-Tricine electrode buffer (0.1 M Tris, 0.1 M Tricine, 0.1% SDS) followed by staining with Coomassie solution (0.1% Coomassie brilliant blue R-250, 10% acetic acid, 45% ethanol) and destaining with a 10% acetic acid, 45% ethanol solution. Fractions containing SpoIIID were pooled, dispensed into 60 µl aliquots, and stored at -70°C for later use in electrophoretic mobility shift assays (EMSAs). Alternatively, for analytical ultracentrifugation, SpoIIID was purified as above except that SP Column buffer did not contain 0.1% Triton X-100, and fractions containing SpoIIID were pooled following purification over the SP Sepharose column and passed over a 1-ml HiTrap™ Heparin HP column (GE Healthcare) that had been equilibrated with 10 mM potassium phosphate buffer pH 7.0 (buffer 1). SpoIIID was eluted with successive washes of buffer 1 supplemented with 0.6 M, 0.8 M, 1 M, 1.2 M, and 1.4 M NaCl. The SpoIIID used for analytical ultracentrifugation eluted at 1.0 M NaCl. The concentration of SpoIIID purified by either method was determined by measuring the absorbance at 280 nm.

EMSAs - Oligonucleotides corresponding to the probes listed in Table 2.4 were synthesized at the Michigan State University Genomics Technology Support Facility, labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (New England Biolabs), and annealed by allowing to cool to room temperature after incubating in a boiling water bath for 10 min. Each annealed probe was visualized by autoradiography after electrophoresis through a 15% polyacrylamide gel, the band corresponding to the appropriate size was excised, and the labeled probe DNA was eluted by incubation at 37°C overnight in 150 µl TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

Labeled probe DNA (~13 nM) was incubated at 30°C for 30 min in buffer [10 mM Tris-HCl pH 7.5, 50 mM NaCl, 10 mM EDTA, 5% glycerol, 0.1 mM double-stranded poly dI·dC (Roche)] with SpoIIID at the concentrations indicated. Reaction mixtures were electrophoresed on an 8% polyacrylamide gel using 0.5x TBE (45 mM Tris-borate, 1 mM EDTA) and dried. Bands were visualized using a Storm 820 PhosphorImager (Molecular Dynamics) and quantified with ImageQuant software (GE Healthcare). The apparent dissociation constant (K_d) was determined by plotting the linear range of the log of the ratio of bound to free probe DNA versus the log [SpoIIID], observing the [SpoIIID] at which the line intersected the *x*-axis (*i.e.* [bound DNA] = [free DNA]), and subtracting 6.5 nM (half the concentration of the DNA probe), since this is the [SpoIIID] bound to DNA when [bound DNA] = [free DNA].

Preparation of Samples for Analytical Ultracentrifugaion. Oligonucleotides identical in sequence to probe 11 (Table 2.4) were obtained from the Michigan State University Genomics Technology Support Facility and were resuspended in water. Equimolar concentrations were combined with buffer 1 supplemented with 50 mM NaCl, placed in a boiling water bath for 10 min, and allowed to cool to room temperature. SpoIIID purified as described above using a Heparin HP column was diluted tenfold in buffer 1 to a 0.1 M NaCl final concentration. Three different amounts of diluted SpoIIID (10.8, 12, and 13.8 nmol) were added to a constant amount of probe 11 DNA (12 nmol) prepared as described above and adjusted to 0.1 M NaCl in buffer 1, resulting in 3 mixtures with slightly different protein to DNA ratios. To each mixture, 1 Complete Mini EDTA-free protease inhibitor tablet was added and the mixtures were incubated at 4°C with rotation for 1 h. SpoIIID/DNA complexes were concentrated using Amicon Ultra 4 (5K MWCO)

(Millipore) filtration devices to a final volume of 1 ml, then shipped on ice to the Colorado State University Specialized Facility for Protein Characterization, where they were dialyzed extensively versus buffer 1 supplemented with 0.1 M NaCl prior to analytical ultracentrifugation.

Analytical Ultracentrifugation - All experiments were performed in a Beckman XL-I using the absorbance optical system and a 4-hole, AN60-Ti rotor. Sedimentation velocity (SV) was performed in a 1.2-cm, 2-sector EPON centerpiece, while sedimentation equilibrium (EQ) was performed in a 1.2-cm, 6-sector EPON centerpiece. For SV, 400 µl of sample at $A_{260} = 0.5$ was sedimented at 55,000 rpm (244,000 x g) for 4 h at 22 °C, with a radial step size of 0.002 cm in the continuous scanning mode. A total of 53 scans were analyzed using the method of Demeler and van Holde (Demeler & van Holde, 2004) to yield the diffusion-corrected, integral distribution of S over the boundary [*G(s)*] within Ultrascan (version 9.4 for Linux). Sedimentation coefficients (*s*) were corrected to that in water at 20°C ($s_{20,w}$). *C(s)* fitting to determine the hydrodynamic properties and, ultimately, model the molecular mass of the complex, was performed within Ultrascan. The solvent density and viscosity were calculated within Ultrascan (1.0033 g/ml and 1.0095 cP, respectively). The partial specific volumes (v-bar) for the SpoIIID protein (0.747 ml/g), DNA (0.55 ml/g), and the SpoIIID/DNA complex (0.68 ml/g) were also estimated using Ultrascan.

Based on the hydrodynamic properties returned from the SV experiments, the conditions for performing EQ (i.e., the appropriate rotor speeds for achieving equilibrium at $\sigma = 1\text{-}4$ and the approximate time to equilibrium) were modeled using Ultrascan. For EQ, 100 µl samples with 0.2, 0.5, and 0.7 absorbance units at 230, 260, and 280 nm,

respectively, were loaded into the centerpiece. The samples were centrifuged to equilibrium (as evidenced by overlaying absorbance scans collected 4-8 hours apart) at 28,000, 34,600, 41,300, and 48,000 rpm (63,000, 96,500, 137,500, and 185,800 x g). Ten scans at 0.001 cm radial step size were averaged for each data set. A total of 72 data sets were available for global fitting. The data were globally fit to various models within Ultrascan (molecular weight distribution model, single, ideal species model, and two component, non-interacting model) as described in the Results section. The fit of the data to the models was judged by the randomness of the residuals and the variance (with respect to the degrees of freedom that each model poses).

Construction of B. subtilis *Strains that Express SpoIIID from the* thrC *Locus* – Competent *B. subtilis* cells were prepared by the Gronigen method as described previously (Harwood & Cutting, 1990). Strains BK395 (containing the *spoIIID83* mutation) (Kunkel *et al.*, 1988)and wild-type PY79 (as a control) (Youngman *et al.*, 1984) were transformed with pPH7 and transformants were selected on LB agar containing 5 μg/ml neomycin to create PH1001 and PH1003, respectively, with the $P_{spoIVCA}$-*gusA* fusion inserted by double crossover into the *amyE* gene as determined by loss of amylase activity (Harwood & Cutting, 1990). Competent PH1001 cells were further transformed with pPH1 (creating strain PH2001) or derivatives in which the *spoIIID* allele had been modified by site-directed mutagenesis (Stratagene) (see Table 2.3 for strain designations), or with pAK3 as a control (creating strain PH2000). Transformants were selected on LB agar containing 100 μg/ml spectinomycin and 5 μg/ml neomycin, chromosomal DNA was prepared as described previously

Table 2.3. Mutations in *spoIIID* and corresponding plasmid and *B. subtilis* strain designations

spoIIID mutation	pAK3 derivative	*B. subtilis* strain	pET-21b derivative
N/A	pAK3	PH2000	
w.t.	pPH1	PH2001	pPH4
H2E	pPH101	PH2101	
D3K	pPH102	PH2102	
K6E	pPH103	PH2103	
E7K	pPH104	PH2104	
R8E	pPH105	PH2105	
K11E	pPH106	PH2106	
K14E	pPH107	PH2107	
E18K	pPH108	PH2108	
K20E	pPH109	PH2109	
K21E	pPH110	PH2110	
V23K	pPH111	PH2111	
R24A	pPH22	PH2022	
R24E	pPH9	PH2009	
V25A	pPH112	PH2112	
V25E	pPH13	PH2013	
V25K	pPH14	PH2014	
I26E	pPH113	PH2113	
A27E	pPH15	PH2015	
K28E	pPH10	PH2010	
E29K	pPH11	PH2011	

Table 2.3 (cont'd).			
F30A	pPH16	PH2016	
G31E	pPH142	PH2142	
V32E	pPH143	PH2143	
S33A	pPH12	PH2012	
S33R	pPH17	PH2017	pPH75
KST343536AAA	pPH18	PH2018	
K34E	pPH114	PH2114	pPH214
S35E	pPH144	PH2144	
T36E	pPH145	PH2145	
VHK373839AAA	pPH19	PH2019	
V37E	pPH146	PH2146	
H38E	pPH115	PH2115	pPH215
K39E	pPH116	PH2116	pPH216
DLT404142AAA	pPH20	PH2020	
D40K	pPH117	PH2117	pPH217
L41K	pPH147	PH2147	
T42K	pPH148	PH2148	
E43K	pPH118	PH2118	pPH218
R44E	pPH119	PH2119	pPH219
E47K	pPH120	PH2120	
D51K	pPH121	PH2121	
E55K	pPH122	PH2122	
K57E	pPH123	PH2123	

Table 2.3 (cont'd).			
E58K	pPH124	PH2124	
D61K	pPH125	PH2125	
H63E	pPH126	PH2126	
K64E	pPH127	PH2127	
R67E	pPH128	PH2128	
H68E	pPH129	PH2129	
R70E	pPH130	PH2130	
E73K	pPH131	PH2131	
K76E	pPH132	PH2132	
K76Stop	pPH21	PH2021	pPH56
K78E	pPH133	PH2133	
K80E	pPH134	PH2134	
K81E	pPH135	PH2135	
D82K	pPH136	PH2136	
D82Stop	pPH137	PH2137	pPH237
E83K	pPH138	PH2138	
E86K	pPH139	PH2139	
G87Stop	pPH140	PH2140	pPH240
E88K	pPH141	PH2141	

(Harwood & Cutting, 1990), and PCRs with primers LK2189 and LK2190 (bordering *thrC*) and with primers LK2234 and LK2235 (bordering the mutant *spoIIID83* allele at the native locus and not complementary to sequences present in pPH1 and its derivatives) were used to identify strains with *spoIIID* from pPH1 or its derivatives inserted by double crossover into the *thrC* gene and with no insertion in *spoIIID83* at the native locus.

Western Blot Analysis – B. subtilis strains were induced to sporulate by nutrient exhaustion as described previously (Harwood & Cutting, 1990), and 1-ml samples were collected at hourly intervals by centrifugation (14,000 x g for 1 min), the supernatant was decanted, the cell pellet was rinsed with 50 mM Tris-HCl pH 8.0, centrifugation and decanting were repeated, and the cell pellet was stored at -20°C. Whole-cell extracts were prepared as described previously (Healy *et al.*, 1991) except the lysis buffer did not contain PMSF or DNase I. After adding 1 vol of 2X sample buffer [50 mM Tris-HCl pH 6.8, 4% SDS, 20% (vol/vol) glycerol, 200 mM DTT, 0.03% bromophenol blue] and boiling for 3 min, samples were subjected to Western blot analysis as described previously (Kroos *et al.*, 2002). Anti-SpoIIID antiserum (Halberg & Kroos, 1992) was used at 1:10,000 dilution. Signals were detected using an LAS-3000 imager (Fujifilm) and analyzed with Multigauge version 3.0 software (Fujifilm).

Measurement of spoIVCA-gusA *Reporter Expression – B. subtilis* strains were induced to sporulate and samples were collected as described above. Cells were resuspended, treated with lysozyme, permeabilized with toluene, and assayed for enzyme activity as described previously (Miller, 1972) except 4-nitrophenyl-β-D-glucuronide (NPG) served as substrate. One unit of enzyme hydrolyzes 1 μmol of substrate/min per A_{595} of original cell density. The background activity (as determined by the level of β-

glucuronidase activity in a strain lacking *spoIIID* assayed in the same experiment) was subtracted from each sample and the values were determined as a percentage of the value for a strain bearing wild type *spoIIID* in the same experiment. The average of 3 biological replicates was determined for each mutation.

Results

Binding of SpoIIID to DNA Containing Three Matches to the SpoIIID Binding Site Consensus Sequence - To examine the binding of SpoIIID to DNA, recombinant SpoIIID was overproduced in *E. coli* and purified (Fig. 2.1A). The purified SpoIIID was then tested for its ability to bind to DNA containing three matches to the SpoIIID binding site consensus sequence (Probe 1, Table 2.4). This sequence was chosen because it resembles a sequence in the *B. subtilis* chromosome (68 to 93 bp downstream of the *sigK* transcriptional start site) that was shown previously to be protected from DNase I digestion in footprinting assays with SpoIIID at a low concentration, indicating the presence of one or more high-affinity binding sites (Halberg & Kroos, 1994). The Probe 1 sequence differs from that of the *B. subtilis* chromosome at position 12 from its 5' end, where a G to C change in the second match to the SpoIIID binding site consensus sequence eliminated its only mismatch. Using EMSAs, SpoIIID and Probe 1 formed a single discrete complex (Fig. 2.1B) with an apparent K_d of 13 nM (Table 2.4), and a Hill coefficient of 1.4 (Fig. 2.1C). The formation of a single complex with little cooperativity of binding suggests that a single molecule of SpoIIID binds to a single site in Probe 1. Because SpoIIID purified from sporulating *B. subtilis* is primarily in a monomeric state (B. Zhang and L.K., unpublished data), it is likely that recombinant SpoIIID is

Figure 2.1. Purification of SpoIIID and binding to different DNA probes. A) Purified SpoIIID (8 μg) was subjected to SDS-PAGE followed by Coomassie Blue staining. Size (in kDa) of molecular weight markers is shown. Arrowhead indicates SpoIIID. B) SpoIIID at 0, 140, 70, 35, 18, or 9 nM (lanes 1-6) was incubated with DNA probe (6 nM) containing three matches to the SpoIIID binding site consensus sequence (Probe 1, Table 1) or a single match to the consensus sequence (Probes 2-4, Table 1) and subjected to EMSAs. Bands were quantified with a phosphorimager. For each probe, all lanes are from the same gel, but intervening lanes were removed for clarity. Filled arrowheads indicate bound DNA and unfilled arrowheads indicate free DNA. C) Hill plots of data obtained in B for Probes 1 (◆), 2 (■), and 3 (▲) with best-fit lines.

Fig. 2.1

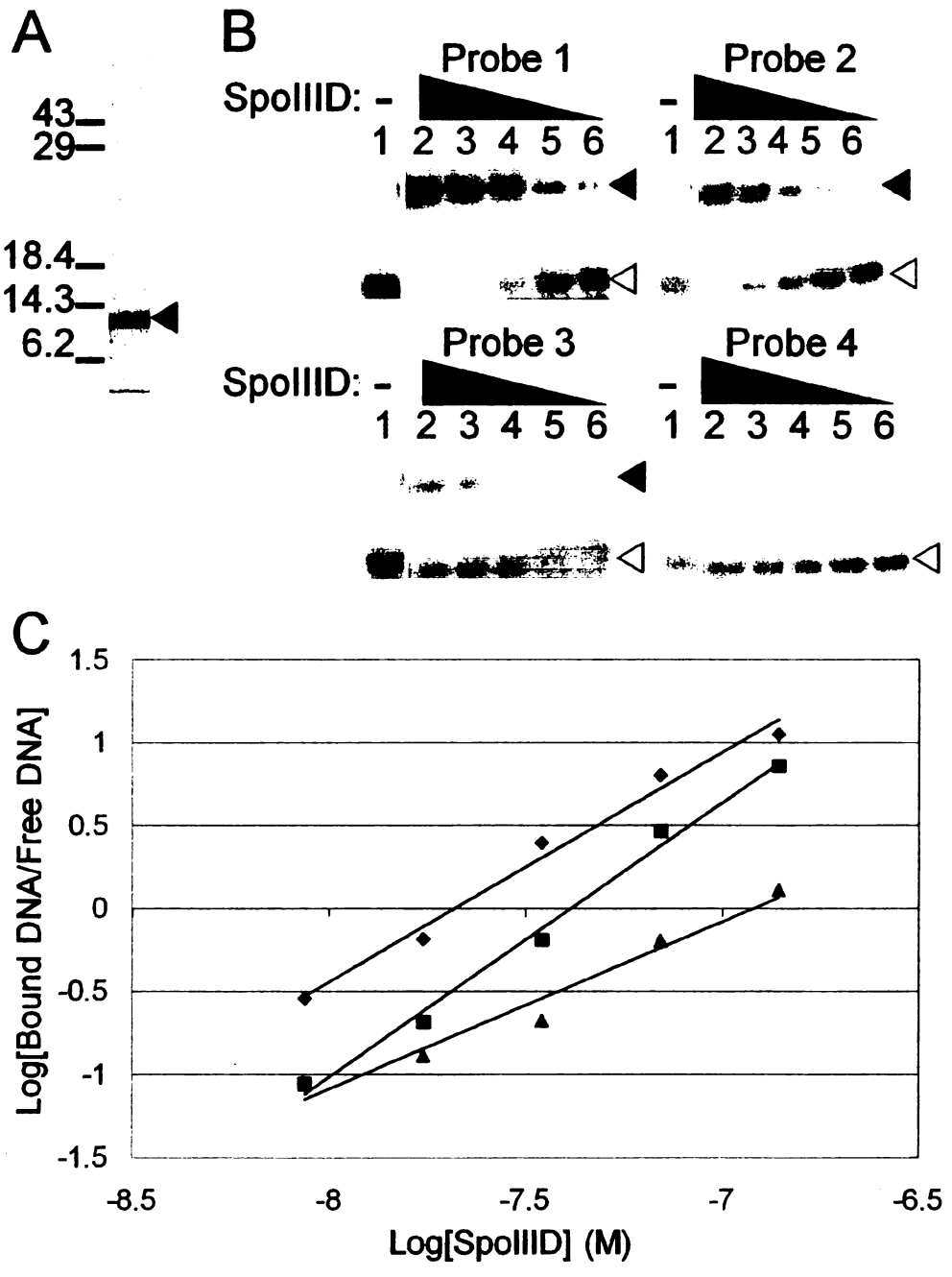

A
43
29
18.4
14.3
6.2
B
Probe 1
Probe 2
SpoIIID: -
1 2 3 4 5 6
Probe 3
Probe 4
C
Log[Bound DNA/Free DNA]
1.5
1
0.5
0
-0.5
-1
-1.5
-8.5
-8
-7.5
-7
-6.5
Log[SpoIIID] (M)

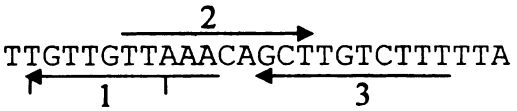
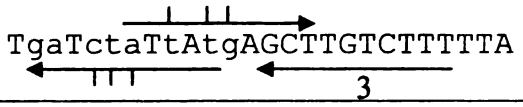
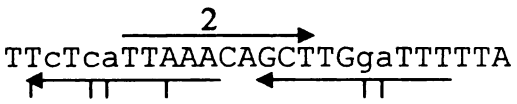
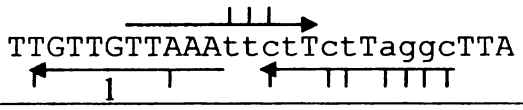
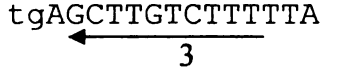
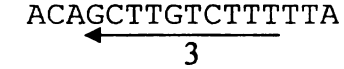
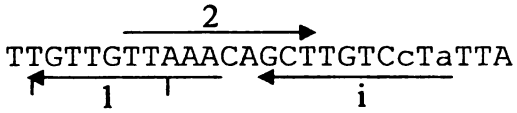
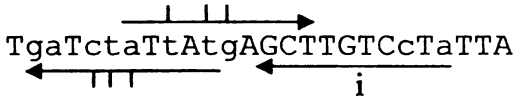
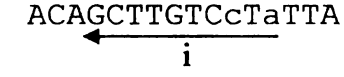
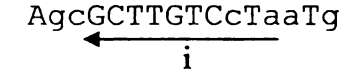
Table 2.4. Sequences of probes used and dissociation constants for SpoIIID-binding		
Probe	Sequence[a]	Average K_d (nM)[b]
1		10 ± 1 (5)
2		28 ± 1 (3)
3		120 ± 35 (5)
4		>15000 (3)
5		28 ± 7 (3)
6		24 ± 5 (4)
7		10 ± 2 (4)
8		21 ± 5 (5)
9		19 ± 3 (4)
10		8 ± 2 (10)

Table 2.4 (cont'd).		
11	tgcGCTTGTCcTaaT cgCGAACAGgAttaa i	8 ± 5 (3)
12	gcGCTTGTCcTaaT i	9 ± 3 (3)
13	AgcGCTTGTCTTTaTg i	16 ± 8 (4)

[a] Matches to the SpoIIID binding site consensus sequence are marked by arrows indicating their orientation. Vertical bars indicate mismatches to the 5'-WWRRACARNY-3' consensus sequence. Upper case letters indicate matches to the sequence of Probe 1, and lower case letters indicate differences from Probe 1. For clarity, the sequence for a single strand of DNA is shown for each probe except Probe 11 for which both strands are shown to demonstrate the possibility of multimerization due to interactions between the unpaired complementary nucleotide on the 5' end of each strand. The idealized (i) consensus sequence is 5'-TAGGACAAGC-3'.

[b] K_d is determined as in Figure 1 and the number of determinations is indicated in parentheses.

monomeric and binds primarily to one of the three matches to the SpoIIID binding site consensus sequence in Probe 1.

To further characterize binding of SpoIIID to DNA, the contribution of each match to the SpoIIID binding site consensus sequence in Probe 1 was tested individually by altering the sequence of the other two matches (Table 2.4, Probes 2-4). Binding of SpoIIID to each of these probes was assayed using EMSAs. Probe 2, with only the third match to the consensus sequence intact, allowed the highest affinity binding, exhibiting only a 2.4-fold increase in apparent K_d compared with Probe 1 (Table 2.4 and Fig. 2.1). Probe 3, with only the second match to the consensus intact, exhibited considerably lower binding affinity (higher apparent K_d) than Probes 1 and 2 (Table 2.4 and Fig. 2.1). Probe 4, with only the first match (containing 2 mismatches to the consensus), did not exhibit binding even at the highest concentration of SpoIIID tested (Table 2.4 and Fig. 2.1B). These results suggest that the third match to the consensus sequence is the major determinant of SpoIIID binding to Probe 1. The second match to the consensus sequence also seems to play a role in binding, while the first match appears to be dispensable.

Since the second and third matches to the consensus sequence appeared to contribute weakly and strongly, respectively, to the binding of SpoIIID to Probe 1, we considered the possibility that more molecules of SpoIIID bind to Probe 1 than to Probes 2 or 3, although this did not appear to be the case based on comparison of the migration of shifted complexes on separate gels (Fig. 2.1B). When compared on the same gel, Probes 1, 2, and 3 formed complexes with SpoIIID that co-migrated (Fig 2.2A). Taken together, the results of EMSAs with Probes 1-4 suggest that a single molecule of SpoIIID binds

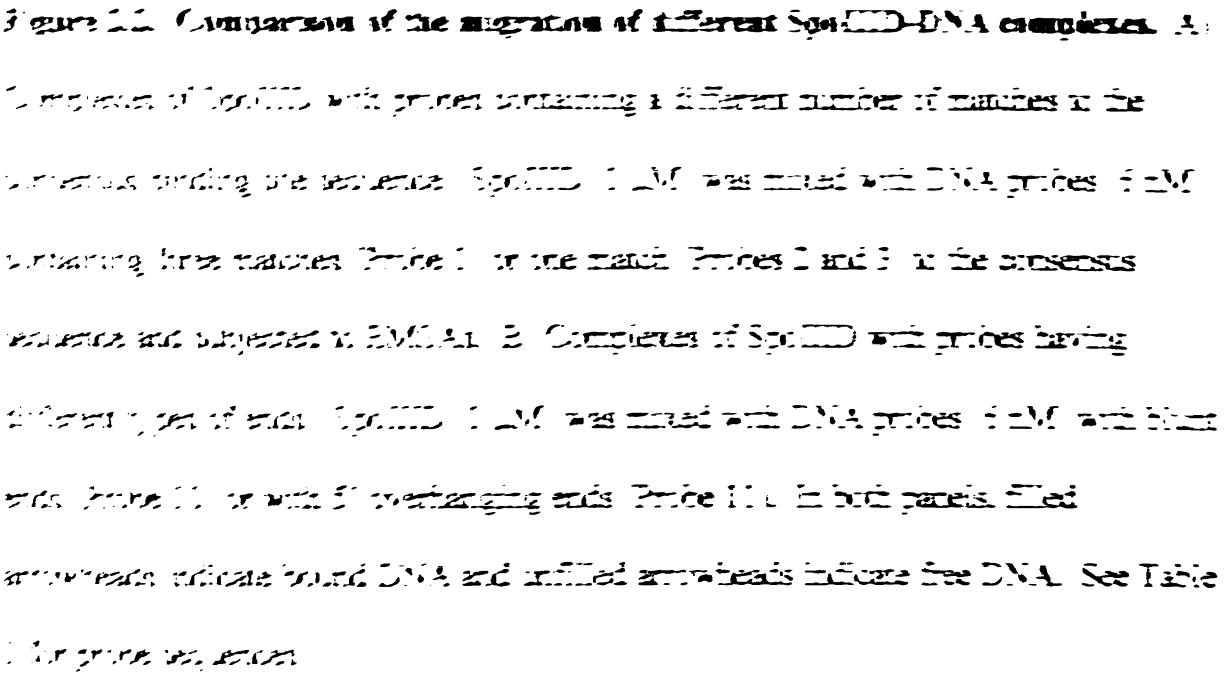
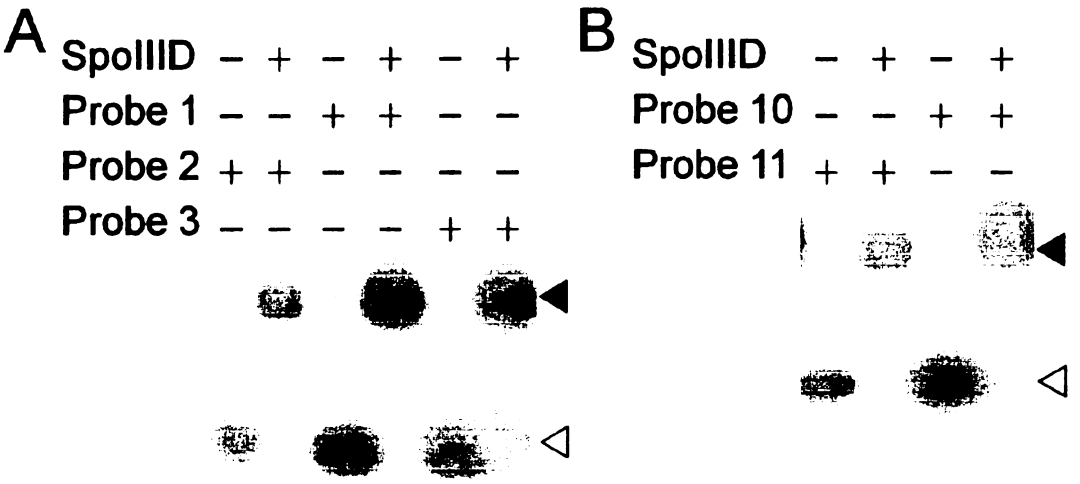


Fig. 2.2



with high affinity to the third match to the consensus sequence or with lower affinity to the second match.

Binding of SpoIIID to DNA Containing a Single Match to the SpoIIID Binding Site Consensus Sequence – If SpoIIID binds with high affinity to the third match to the consensus sequence in Probes 1 and 2 as suggested by the results presented above, it might be possible to shorten those probes and still observe high-affinity binding (i.e., comparable to that observed for Probe 2 in which the first and second matches to the consensus were altered). Probes 1 and 2 were shortened to include only three base pairs on each side of the third match to the consensus sequence. No matter whether the first three base pairs were identical to those in Probe 2 (Probe 5, Table 2.4) or Probe 1 (Probe 6, Table 2.4), the shortened probes exhibited apparent K_d values similar to Probe 2 (Table 2.4). These results demonstrate that a single match to the consensus sequence is sufficient for high-affinity binding of SpoIIID to DNA.

An examination of the sequences in 20 SpoIIID binding sites mapped by DNase I footprinting in several studies revealed the most common nucleotide for each position of the binding site consensus sequence (Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang et al., 1997, Eichenberger et al., 2004). We call this sequence, 5'-TAGGACAAGC-3', the idealized consensus sequence. When this sequence was substituted for the third match to the consensus sequence in Probe 1 (Probe 7, Table 2.4), Probe 2 (Probe 8, Table 2.4), or Probe 6 (Probe 9, Table 2.4), no significant change in the apparent K_d for SpoIIID binding was observed (Table 2.4). In these contexts, the two sequences exhibited similar affinity for SpoIIID.

We serendipitously discovered that the sequence context of the match to the consensus sequence can affect SpoIIID binding affinity. When two of the three base pairs on each side of the idealized consensus were changed in Probe 9, the binding affinity increased twofold (Probe 10, Table 2.4). The rationale for this change came from studies with the idealized consensus sequence in the context of shorter ends. Probe 11 (Table 2.4) was created to allow potential base-pair interactions at the ends to form repeats of the DNA sequence. Probe 11 was bound with high affinity (Table 2.4), but this did not require the single nucleotide overhangs at the ends (Probe 12, Table 2.4). Probe 10 was designed to have the same two base pairs at each end of the idealized consensus sequence as Probes 11 and 12, and all three probes bound SpoIIID with about twofold higher affinity than Probe 9 (Table 2.4). Likewise, changing the context of the third match to the consensus sequence in this way (Probe 13, Table 2.4) lowered the average apparent K_d (compare Probe 6, Table 2.4), although the effect was less than twofold. We conclude that the sequence surrounding the consensus sequence can have a small effect on the affinity of SpoIIID for short DNA probes.

While Probe 11, with single base overhangs at its 5' ends, was designed to potentially allow a DNA repeat with bound SpoIIID to form, the migration of the complex that formed was indistinguishable from the complex formed with Probe 10, which has blunt ends (Fig. 2.2B). Moreover, the Hill coeffients were 1.4, 1.4, and 1.3 for Probes 10, 11, and 12, respectively, so binding did not appear to be cooperative. Rather, it appears that a single molecule of SpoIIID (likely a monomer) can bind to a single match to the consensus sequence with high affinity.

Analytical Ultracentrifugation of SpoIIID and DNA – To determine the composition of the complex formed between SpoIIID and DNA containing a single copy of the idealized consensus sequence, we used analytical ultracentrifugation. SpoIIID was mixed with Probe 11 (Table 2.4) at slightly different ratios and the homogeneity of the complex was assessed using absorbance-detected sedimentation velocity experiments. We monitored the sedimentation of the complex by detecting both the DNA component (at 260 nm) and the protein component (at 230 nm). The data were analyzed using the boundary analysis method of Demeler and van Holde (Demeler & van Holde, 2004), which results in the diffusion-corrected, integral distribution of S over the entire boundary. A homogenous sample will have a vertical or near vertical *G(s)* plot, while a heterogeneous sample will have a *G(s)* plot with one or more positive slopes. As shown in Figure 2.3A, the plots that result from analysis of the data from detecting at 260 nm were nearly vertical, and >80% of each sample sedimented at ~2.5S. Less than 20% of a smaller, slower sedimenting material was also detected. Analysis of the data from detecting the sedimentation of the protein component of the complex at 230 nm yielded similar plots, and also detected a small fraction of slower sedimenting material (~1S, data not shown). Further analysis of the data from the 1:1 mixture of Probe 11 with SpoIIID, using *C(s)* analysis (Schuck & Rossmanith, 2000), resulted in the observation of two components: ~12% of a 1.4S and 9000 Da component, and ~83% of a 2.45S and 21,000 Da component. This suggested that the majority of the sample exhibited sedimentation properties consistent with a complex made up of one duplex of Probe 11 DNA and one monomer of SpoIIID (calculated mol. wt. of complex = 19,946 Da). However, while the *C(s)* analysis is a good first step toward analyzing a macromolecular complex in solution,

Figure 2.3. Analytical ultracentrifugation of the SpoIIID-DNA complex. A) Sedimentation velocity experiments. The resulting G(s) distributions for the 1.1:1 (●), 1:1 (○) and 0.9:1 (▲) complexes of DNA:SpoIIID are shown. Sedimentation velocity data were analyzed as described in Experimental Procedures. Sedimentation coefficients (*s*) were corrected to that in water at 20°C ($s_{20,w}$). B) Sedimentation equilibrium of the 1:1 complex of DNA:SpoIIID. Samples containing 1.1-6.3 μM of the complex were sedimented to equilibrium at 28,000, 34,600, 41,300, and 48,000 rpm at 5°C. Two equilibrium scans were collected at each speed 4-8 hours apart to ensure the sample was at equilibrium. The resulting 72 data sets were globally fit to numerous models within Ultrascan. Plot overlays (lower part) for the fit of the sedimentation equilibrium data to a two-component, non-interacting model are shown. The model (solid lines) closely overlays the data (open circles). The residuals for the fit are shown in the upper part. The variance for the fit was 2.93×10^{-5}.

Fig. 2.3

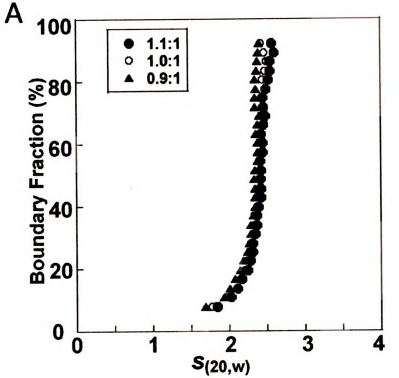
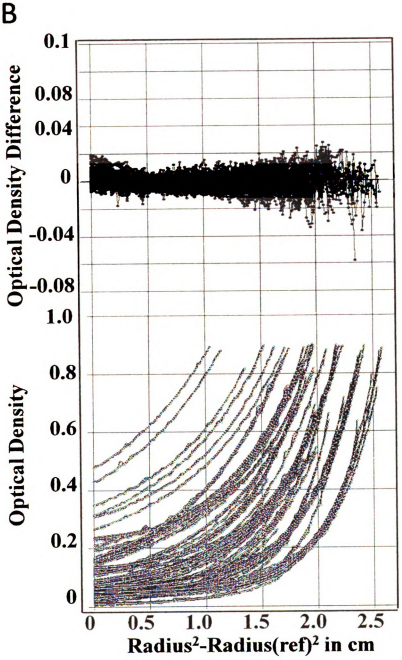


Fig. 2.3 (cont'd.)


B
Optical Density Difference
0.1
0.08
0.04
0
-0.04
-0.08
Optical Density
1.0
0.8
0.6
0.4
0.2
0
0
0.5
1.0
1.5
2.0
2.5
Radius²-Radius(ref)² in cm

it assumes that all components share a similar partial specific volume, so the mass values returned are only accurate for a truly homogenous sample. Nonetheless, despite the small fraction of slower sedimenting material in the sample, *G(s)* and *C(s)* analyses showed the 1:1 mixture of Probe 11 with SpoIIID to be an excellent candidate for sedimentation equilibrium analysis to rigorously determine the molecular mass of the complex.

For sedimentation equilibrium analysis, a total of 72 data sets were collected using 4 different rotor speeds and 9 sample concentrations (from 1.1 to 6.3 μM complex) with detection at 230, 260, and 280 nm. The data were best fit (variance = 2.93×10^{-5}) using a 2-component, non-interacting species model (Fig 2.3B), as indicated by the sedimentation velocity analysis presented above. The two fitted components exhibited molecular masses of 9,579 Da and 20,310 Da (partial specific volumes of 0.55 ml/g and 0.68 ml/g, respectively), the latter of which is in excellent agreement (within 1.8%) with the calculated mol. wt. of a 1:1 complex between Probe 11 DNA and a SpoIIID monomer. The smaller component most closely approximated the mass of the unliganded DNA (within 4.5%), though this analysis cannot preclude the possibility that it represents the unliganded SpoIIID monomer (within 12%). Taken together, the analytical ultracentrifugation analyses indicate that a SpoIIID monomer forms a 1:1 complex with Probe 11 DNA, which contains a single copy of the idealized consensus sequence.

Structure Prediction for SpoIIID – SpoIIID has been predicted to contain an HTH DNA-binding motif (residues 23 to 42) (Kunkel et al., 1989). In agreement with this prediction, modeling of SpoIIID based primarily on similarity of its predicted secondary structure with structures from the Protein Data Bank resulted in tertiary structure

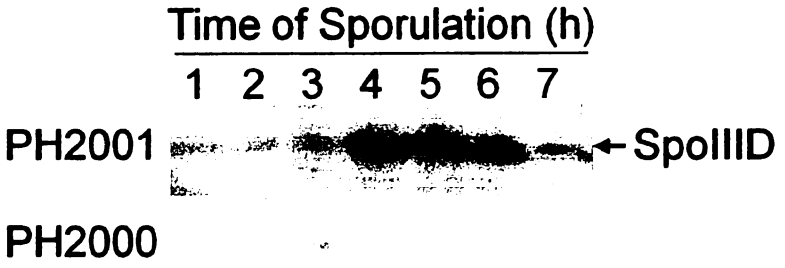
predictions for SpoIIID based on the structures of two HTH DNA-binding proteins, cI repressor protein of bacteriophage λ (W. Wedemeyer, unpublished data) and SinR of *B. subtilis* (data not shown). Both predictions depict a protein with an N-terminal core comprised of four α-helices, including an HTH motif (helices 2 and 3). The predictions differ at the C-terminus where the structure based on cI predicts a fifth α-helix (residues 76-85) extending away from the rest of the protein, while the SinR-based structure predicts the region is disordered. The SinR and cI DNA-binding domains exhibit a high degree of structural similarity and both proteins form oligomers (in contrast to SpoIIID).

Mutational Analysis of spoIIID – To begin testing the validity of the structure prediction for SpoIIID and to identify residues important for the function of SpoIIID, we established a convenient system for analysis of *spoIIID* mutations. To measure the ability of SpoIIID to bind DNA and activate transcription, we constructed a *spoIVCA-gusA* transcriptional fusion reporter. Transcription from the *spoIVCA* promoter by σ^E RNA polymerase was shown previously to be activated by SpoIIID *in vitro* (Halberg & Kroos, 1994), and *spoIVCA* failed to be expressed in *spoIIID* mutant cells during sporulation (Kunkel et al., 1988, Sato *et al.*, 1994). To allow expression of mutant *spoIIID* alleles, we used a vector plasmid designed to allow genes to be recombined into the *B. subtilis* chromosome at an ectopic site (*thrC*), which plays no known role in sporulation gene expression. A DNA fragment encompassing *spoIIID* and its promoter region was cloned into the vector plasmid. The resulting plasmid (pPH1) was transformed into *B. subtilis* strain PH1001, which contains the *spoIVCA-gusA* reporter integrated at the *amyE* locus in a *spoIIID* mutant background. Recombination between the plasmid and the chromosome resulted in replacement of the *thrC* gene with a copy

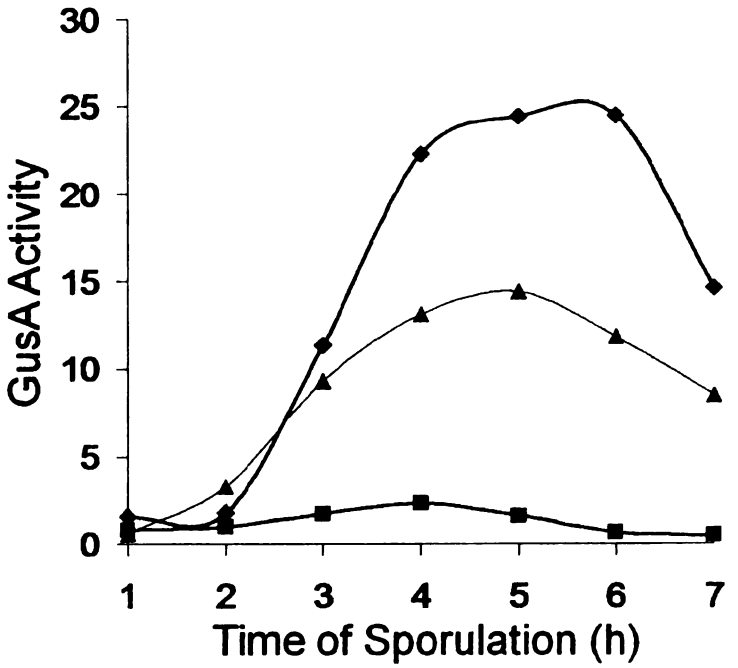
Figure 2.4. System for mutational analysis of SpoIIID. A) Western blot of ectopic SpoIIID expression during sporulation. *B. subtilis* strains with (PH2001) or without (PH2000) *spoIIID* at the *thrC* locus, and bearing a mutation that prevents *spoIIID* expression from the native locus, were induced to sporulate by nutrient exhaustion and samples collected at hourly intervals were subjected to Western blot analysis with antibodies against SpoIIID. B) Expression of a *spoIVCA-gusA* reporter during sporulation. *B. subtilis* strains PH2001 (◆), PH2000 (■), and PH1003 (▲) containing the *spoIVCA-gusA* reporter at the *amyE* locus were induced to sporulate by nutrient exhaustion and samples collected at hourly intervals were assayed for GusA activity. PH1003 expresses *spoIIID* from the native locus. Points are the average of 3 determinations.

Fig. 2.4

A


Time of Sporulation (h)
1 2 3 4 5 6 7
PH2001
SpoIIID
PH2000

B


GusA Activity
30
25
20
15
10
5
0
1 2 3 4 5 6 7
Time of Sporulation (h)

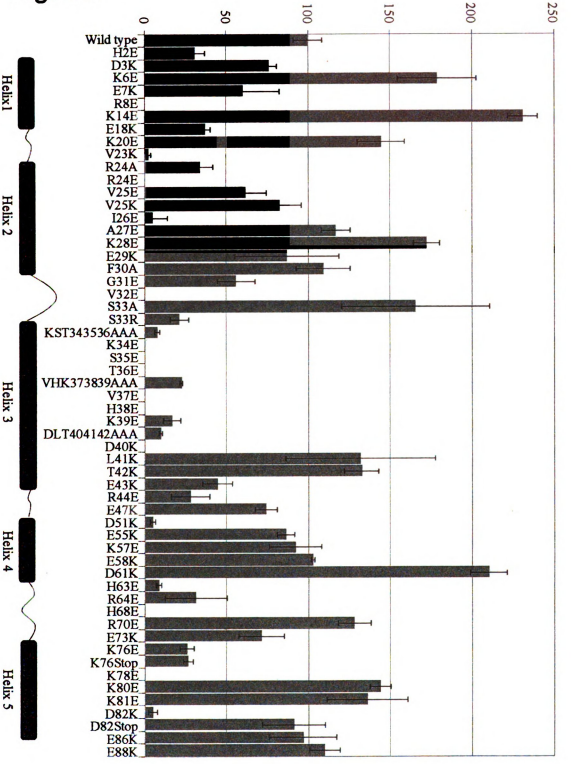
interrupted by *spoIIID* and a gene coding for neomycin resistance (used for selection), creating strain PH2001. As expected, the ectopic copy of *spoIIID* was expressed normally during sporulation (Fig 2.4A). SpoIIID began to accumulate by 3 h into sporulation and its concentration rose by 4 h then began to decline by 6 h, as observed previously for wild-type *B. subtilis* expressing SpoIIID from the native *spoIIID* locus (Halberg & Kroos, 1992). As a negative control, the vector plasmid (with no *spoIIID* gene) was transformed into *B. subtilis* strain PH1001, creating strain PH2000 with only the neomycin resistance gene at the ectopic site. This strain failed to accumulate SpoIIID during sporulation, although a weak signal presumably due to a protein that co-migrated with SpoIIID and cross-reacted with antibodies against SpoIIID was observed (Fig 2.4A).

The two *B. subtilis* strains described above that were subjected to Western blot analysis (Fig 2.4A) were assayed for GusA activity from the *spoIVCA-gusA* reporter during sporulation (Fig 2.4B). For comparison, activity from the *spoIVCA-gusA* reporter in an otherwise wild-type background (*B. subtilis* strain PH1003; expressing SpoIIID from the native *spoIIID* locus) was measured. GusA activity increased similarly in *B. subtilis* strain PH2001, which expresses *spoIIID* ectopically, and in the wild-type strain PH1003. The increase in GusA activity (Fig 2.4B) correlated with accumulation of SpoIIID (Fig 2.4A) for PH2001. In the negative control strain lacking *spoIIID* at the ectopic site, PH2000, *spoIVCA-gusA* failed to be expressed (Fig 2.4B). These results demonstrate that *spoIVCA-gusA* expression provides an assay for functional SpoIIID, which can be synthesized ectopically, establishing a system for mutational analysis *spoIIID*.

To identify residues important for the function of SpoIIID, we subjected the *spoIIID* gene to site-directed mutagenesis and used the system described above to express the mutant allele and measure the effect on *spoIVCA-gusA* expression during sporulation. Because SpoIIID has many charged residues and these are likely to be surface-exposed, our mutagenesis strategy was to create charge reversals throughout SpoIIID. In addition, we mutated every residue in the predicted HTH DNA-binding motif (residues 23-42). Each mutant allele was placed at the ectopic *thrC* site (as verified by diagnostic PCR) in the chromosome of the *B. subtilis* strain (PH1001) that contains the *spoIVCA-gusA* reporter integrated at the *amyE* locus in a *spoIIID* mutant background. Expression of *spoIVCA-gusA* was measured at 5 h into sporulation, which was the time when expression in a strain with the wild-type *spoIIID* gene at the ectopic site reached a plateau (Fig 2.4B). As shown in Figure 2.5, charge reversals in the first 21 residues of SpoIIID had modest effects (less than threefold change) on expression of the reporter, with the exception of the R8E substitution. Substitutions in the predicted first helix (residues 23-30) of the HTH changed reporter expression less than twofold, with the exceptions of V23K, R24A, R24E, and I26E. Some substitutions in the predicted turn (residues 31-33) of the HTH reduced reporter expression more than fourfold (V32E and S33R) but others had less than a twofold effect (G31E and S33A). Strikingly, most substitutions in the predicted second helix (residues 34-42) of the HTH, which is the predicted "recognition helix" that would interact directly with DNA in the major groove, dramatically reduced or abolished expression of the *spoIVCA-gusA* reporter. Charge reversals in the C-terminal half of SpoIIID had modest effects (less than threefold change) on expression of the reporter, with the exceptions of the R44E, D51K, H63E, H68E, K76E, K78E, and

Figure 2.5. Quantification of the effects of mutations in *spoIIID* on transcription *in vivo*. β-glucuronidase activity as a percentage of wild type levels was determined as described in Experimental Procedures. Bars indicate averages of 3 biological replicates and error bars indicate one standard deviation. Solid bars below the *x*-axis indicate the positions of the helices in the predicted structure of SpoIIID based on the cI repressor protein of bacteriophage λ.

Fig. 2.5


0
50
100
150
200
250
Wild type
H2E
D3K
K6E
E7K
R8E
K14E
E18K
K20E
V23K
R24A
R24E
V25E
V25K
I26E
A27E
K28E
E29K
F30A
G31E
V32E
S33A
S33R
KST343536AAA
K34E
S35E
T36E
VHK373839AAA
V37E
H38E
K39E
DLT404142AAA
D40K
L41K
T42K
E43K
R44E
E47K
D51K
E55K
K57E
E58K
D61K
H63E
R64E
H68E
R70E
E73K
K76E
K76Stop
K78E
K80E
K81E
D82K
D82Stop
E86K
E88K
Helix1
Helix 2
Helix 3
Helix 4
Helix 5

D82K substitutions. We also made C-terminal truncations of SpoIIID by substituting a stop codon for codons that normally specify residues 76 or 82 of the 93-residue protein. Elimination of 12 (D82Stop) residues from the C-terminal end of SpoIIID had little effect on reporter expression, but elimination of 18 (K76Stop) residues reduced reporter expression about fourfold. In summary, our mutational analysis revealed primarily two regions that are important for producing functional SpoIIID; the predicted DNA recognition helix (residues 34-42) and a basic region near the C-terminus (residues 63-81). Charge reversals at a few other positions (R8, R24, R44, and D51) dramatically reduced or abolished expression of the SpoIIID-dependent reporter.

Expression of the *spoIVCA-gusA* reporter presumably requires that SpoIIID accumulate, bind to the *spoIVCA* promoter region, and activate transcription by σ^E RNA polymerase (Halberg & Kroos, 1994, Kunkel et al., 1988, Sato et al., 1994). To distinguish between mutations that prevent SpoIIID accumulation versus those that prevent DNA binding and/or transcriptional activation, we performed Western blot analysis with antibodies against SpoIIID. For several strains with reduced or abolished reporter expression, we could not detect SpoIIID reliably at 5 h into sporulation (data not shown). Because SpoIIID positively autoregulates its own expression (Kunkel et al., 1989), mutations in *spoIIID* that impair expression of the *spoIVCA-gusA* reporter were expected to reduce accumulation of SpoIIID. Presumably, the mutations that allow some reporter expression do allow the altered SpoIIID protein to accumulate at a low concentration that we could not detect reliably above the background signal in the negative control strain (PH2000), which lacks *spoIIID* at the ectopic site (Fig 2.4A).

Binding of Altered SpoIIID Proteins to DNA - To further characterize altered SpoIIID proteins, we engineered expression of several in *E. coli* (Table 2.5). All the altered proteins we tested accumulated normally in soluble form; however, the D40K substitution caused SpoIIID to be unstable during purification. The D40K substitution might alter the structure of SpoIIID in a way that makes it more susceptible to *E. coli* proteases. The other altered proteins were stable during purification, suggesting that they are not unfolded.

As a step toward distinguishing between altered SpoIIID proteins that are defective for DNA binding and those specifically impaired in transcriptional activation, their ability to bind to DNA containing the idealized consensus sequence (Probe 10, Table 2.4) was measured using EMSAs. Figure 2.6 shows representative results and Table 2.5 lists the average apparent K_d of altered SpoIIID proteins. The S33R substitution in the predicted turn of the HTH increased the apparent K_d more than tenfold (Fig 2.6B and Table 2.5), indicating that the protein is impaired for DNA binding, and providing an explanation for the reduced *spoIVFCA-gusA* expression observed when the S33R substitution was tested *in vivo* as described above (Fig 2.5). K34E, H38E, and K39E substitutions in the predicted DNA recognition helix abolished detectable binding (Fig 2.6C and Table 2.5), supporting the structure prediction for this region. These *in vitro* results are in qualitative agreement with our *in vivo* results as the K34E and H38E substitutions abolished *spoIVCA-gusA* expression, and the K39E substitution reduced expression about fivefold (Fig 2.5). The ability to detect expression of the reporter in the case of the K39E substitution suggests that this assay provides a more sensitive measure for functional SpoIIID than does the *in vitro* DNA-binding assay. The E43K substitution in the residue

Figure 2.6. Binding of altered SpoIIID proteins to DNA. Mutant alleles of *spoIIID* were expressed in *E. coli* and the proteins were purified over an SP Sepharose column. No protein or decreasing concentrations of SpoIIID (numbers indicate nM concentration of SpoIIID) were mixed with a DNA probe (13 nM) containing a single idealized consensus sequence for SpoIIID binding (Probe 10) and analyzed for binding as in Figure 1. Wild-type (WT) (A) and representative mutant proteins showing 15-fold reduced binding (B); no binding (C); and 3-fold reduced binding (D). Intervening lanes were removed for clarity in (A), (B), and (D). Filled arrowheads indicate bound DNA and unfilled arrowheads indicate free DNA.

Fig. 2.6

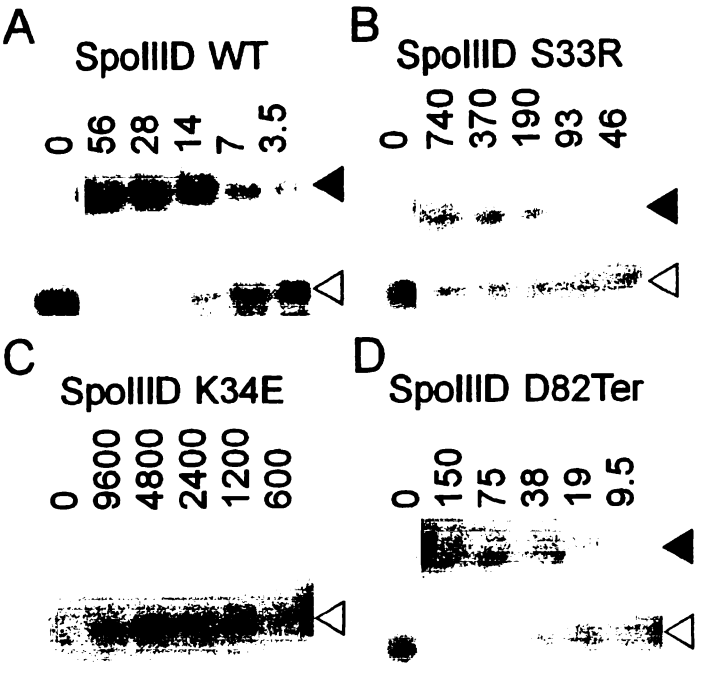

A
SpoIIID WT
0
56
28
14
7
3.5
B
SpoIIID S33R
0
740
370
190
93
46
C
SpoIIID K34E
0
9600
4800
2400
1200
600
D
SpoIIID D82Ter
0
150
75
38
19
9.5

Table 2.5. Binding of altered SpoIIID proteins to DNA	
SpoIIID Protein[a]	Average K_d(nM)[b,c]
Wild Type	8 ± 2 (10)
S33R	122 ± 23 (3)
K34E	>9600 (3)
H38E	>1500 (4)
K39E	>1100 (3)
D40K	Not Determined[d]
E43K	74 ± 20 (3)
R44E	>5000 (3)
K76Stop	>4100 (3)
D82Stop	23 ± 7 (3)

[a] Mutant alleles of *spoIIID* were expressed in *E. coli* and purified over an SP Sepharose column.

[b] K_d was determined as in Figure 1 using Probe 10 which contains one copy of the idealized SpoIIID binding site consensus sequence.

[c] The number of determinations is indicated in parentheses.

[d] D40K was unstable in repeated purification attempts.

following the predicted recognition helix reduced reporter expression about twofold (Fig 2.5) and increased the apparent K_d about sevenfold (Table 2.5), so a considerable defect in DNA binding as measured *in vitro* nevertheless allowed substantial reporter expression *in vivo*. Likewise, the R44E substitution and the K76Stop truncation reduced reporter expression only about fourfold (Fig 2.5), but the purified proteins failed to bind detectably to Probe 10 (Table 2.5). The results with the K76Stop truncation indicate that a region near the C-terminus of SpoIIID is important for DNA binding. On the other hand, the D82Stop truncation had little effect on reporter expression (Fig 2.5) or DNA binding (Fig 2.6D and Table 2.5). We conclude that in addition to the predicted recognition helix of the HTH, one or more residues in the C-terminal region of SpoIIID spanning from K76 to K81 are critical for DNA binding.

Discussion

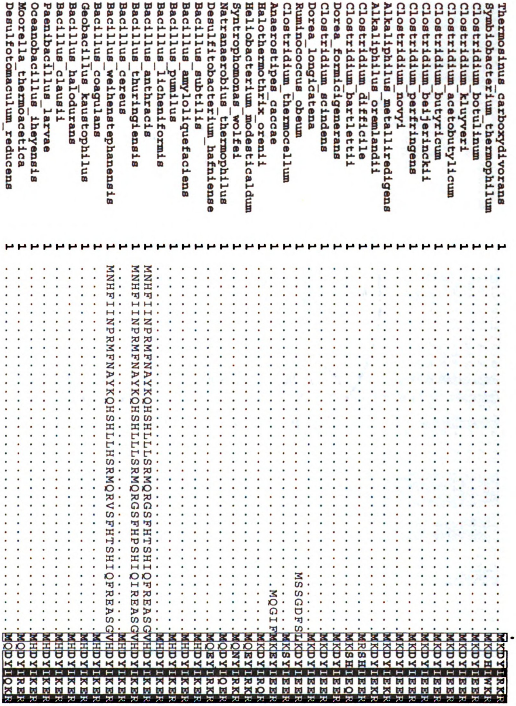
We have discovered that a highly conserved, key regulator of gene expression during sporulation binds to specific sequences in DNA as a monomer with high affinity by using at least two regions. Dwindling resources in the mother cell during sporulation might have favored evolution of a small (*B. subtilis* SpoIIID is 93 residues) transcription factor capable of high-affinity binding to specific sites in the chromosome as a monomer. If our structure prediction for SpoIIID is correct, as suggested by preliminary NMR data of SpoIIID in complex with Probe 11 DNA (B. Chen, P. H., A. Liu, H. Yan, and L. K., unpublished data), SpoIIID achieves high-affinity binding by a novel mechanism involving an HTH motif (whose recognition helix presumably contacts the major groove of DNA) followed by an additional α-helix (the C-terminal basic region) that makes contacts with DNA. Below, we discuss our findings in the context of this emerging

model and we note additional studies that will be required to understand not only how SpoIIID binds DNA but how it functions as an activator and repressor of transcription. SpoIIID is highly conserved among *Bacilli* and *Clostridia* related to *B. subtilis* that also form endospores (Fig 2.7). Non-spore-forming genera like *Listeria* and *Staphylococcus*, although closer phylogenetically to *Bacilli* than are *Clostridia*, do not harbor SpoIIID orthologs in their genomes, as is true for many other sporulation-specific genes (Eichenberger et al., 2004, Eichenberger *et al.*, 2003). To our knowledge, the role of SpoIIID has not been investigated in any organism other than *B. subtilis*. However, a recent study showed that expression of *spoIIID* in *Clostridium perfringens* is not under control of σ^E RNAP (Harry *et al.*, 2009), as it is in *B. subtilis* (Kunkel et al., 1989). Strikingly, SpoIIID orthologs exhibit highest identity in their putative recognition helix (residues 34-42) and in their C-terminal basic region (residues 63-81) (Fig 2.7). Conservation in the putative recognition helix was noted previously and it was inferred that SpoIIID orthologs would bind to similar DNA sequences (Eichenberger et al., 2004). This inference is strengthened by our finding that the C-terminal basic region is critical for DNA binding and the observation that this region is highly conserved. The high degree of conservation in regions shown in this study to be important for DNA binding suggests that SpoIIID orthologs likely are key regulators of gene expression during endospore formation by many different bacteria, including several human pathogens.

Endospore formation is triggered by nutrient limitation, and a recent study suggests that the mother cell shares its resources with the developing forespore (Camp & Losick, 2009), so the need to conserve biosynthetic capacity might have favored evolution of a highly efficient mode of DNA binding by SpoIIID. Proteins encoded in the *spoIIIA*

Figure 2.7. Alignment of SpoIIID orthologs. The highest scoring ortholog of SpoIIID, as identified by a blastp search (Altschul *et al.*, 1990, Gish & States, 1993), from each species containing one was aligned using ClustalW (Larkin *et al.*, 2007). The results were then visualized using ESPript (Gouet *et al.*, 1999). White letters with a black background indicate complete conservation.

Fig. 2.7



Thermosinus_carboxydivorans
Symbiobacterium_thermophilum
Clostridium_botulinum
Clostridium_kluyveri
Clostridium_acetobutylicum
Clostridium_butyricum
Clostridium_beijerinckii
Clostridium_perfringens
Clostridium_novyi
Alkaliphilus_metalliredigens
Alkaliphilus_oremlandii
Clostridium_difficile
Clostridium_bartlettii
Dorea_formicigenerans
Clostridium_scindens
Dorea_longicatena
Ruminococcus_obeum
Clostridium_thermocellum
Anaerostipes_caccae
Halothermothrix_orenii
Heliobacterium_modesticaldum
Syntrophomonas_wolfei
Natranaerobius_thermophilus
Desulfitobacterium_hafniense
Bacillus_subtilis
Bacillus_amyloliquefaciens
Bacillus_pumilus
Bacillus_licheniformis
Bacillus_anthracis
Bacillus_thuringiensis
Bacillus_cereus
Bacillus_weihenstephanensis
Bacillus_coagulans
Geobacillus_kaustophilus
Bacillus_halodurans
Bacillus_clausii
Paenibacillus_larvae
Oceanobacillus_iheyensis
Moorella_thermoacetica
Desulfotomaculum_reducens

Fig. 2.7 (cont'd.)

Fig. 2.7 (cont'd.)

operon are co-expressed with SpoIIID under σ^E RNAP control in the mother cell, and SpoIIIA proteins are involved in forming channels that span the intermembrane space between the mother cell and the forespore by interacting with the extracellular domain of SpoIIQ, which is made in the forespore under σ^F RNAP control and is inserted into the inner membrane surrounding the forespore (Camp & Losick, 2008, Meisner et al., 2008). The channels are necessary to permit transcription and/or translation in the forespore around the time engulfment of the forespore is completed, suggesting that the channels are feeding tubes through which the mother cell supplies small molecules (e.g., ATP, nucleotides, amino acids) to the forespore (Camp & Losick, 2009). If this model is correct, the ability to produce a small protein like SpoIIID that can bind to specific sites with high affinity as a monomer and regulate many genes might have provided a selective advantage evolutionarily.

Most bacterial transcription factors are homodimers that bind to palindromic sites in DNA (Huffman & Brennan, 2002). HTH-containing transcription factors are generally no exception to this rule, so it was somewhat surprising that SpoIIID behaved as a monomer during gel filtration chromatography after purification from sporulating *B. subtilis* (B. Zhang and L.K., unpublished data). Members of the AraC/XylS family of transcription factors bind to DNA as monomers, but the monomer contains two HTH motifs (Gallegos *et al.*, 1997), and SpoIIID is predicted to contain only one (Kunkel et al., 1989). Bacteriophage λ excisionase (Xis) is involved in site-specific recombination of DNA, rather than transcription, but Xis has a single, winged HTH and is a small protein that is monomeric in solution and can bind to a single site with moderate affinity, and binds cooperatively to DNA containing more than one copy of its recognition

sequence (Abbani *et al.*, 2007, Bushman *et al.*, 1984, Sam *et al.*, 2004). We thought that SpoIIID might likewise achieve high-affinity binding *via* cooperative interactions. Instead, we found that SpoIIID bound to a high-affinity site containing three matches to its binding site consensus sequence with little cooperativity, forming a single shifted complex in EMSAs, primarily by binding to the best match to an idealized consensus sequence (5'-TAGGACAAGC-3') deduced from 17 known SpoIIID binding sites (Figs. 1 and 2, and Table 2.4). The best match in Probe 1 to the idealized consensus sequence is the third match (5'-AAAGACAAGC-3'; Table 2.4, see the arrow labeled 3 and note that this is the sequence of the other strand), which has just two mismatches, whereas the second match (5'-TTAAACAGCT-3'; Table 2.4, see the arrow labeled 2) has six mismatches. Non-conformity of the second match to the idealized consensus sequence might account for poor binding of SpoIIID to Probe 3, but further studies are needed with shorter probes containing only the second match. Although SpoIIID appears to bind with lower affinity to the second match than to the third match, simultaneous binding to both matches in Probe 1 would have been expected to produce a second shifted complex of lower mobility. However, a second complex was not observed even at high (> 1 μM) SpoIIID concentrations (data not shown), suggesting that binding to the two sites is mutually exclusive. The first match in Probe 1 (5'-TTTAACAACA-3'; Table 2.4, see the arrow labeled 1 and note that this is the sequence of the other strand) has five mismatches to the idealized consensus, but two of these (at the third and tenth positions) are also mismatches to the more general WWRRACARNY consensus, perhaps explaining our inability to detect SpoIIID binding to Probe 4.

The lack of cooperative binding of SpoIIID to DNA containing three matches to its binding site consensus sequence motivated us to explore further whether SpoIIID achieves high-affinity DNA-binding by an unusual mechanism. We found that SpoIIID can bind with high affinity to a 14-bp duplex containing one copy of the idealized consensus sequence with an additional 2 bp on each end (Probe 12, Table 2.4). Whether this constitutes a minimal high-affinity site remains to be tested with shorter duplexes. We used analytical ultracentrifugation analyses to show that SpoIIID binds as a monomer to Probe 11, which contains a single copy of the idealized consensus sequence (Table 2.4 and Fig 2.3). Eukaryotic homeodomain proteins can bind as monomers to short DNA segments because in addition to their HTH they have an N-terminal arm that interacts with the adjacent minor groove (Tullius, 1995). As is typical for homeodomain proteins and other HTH proteins (Aravind *et al.*, 2005), our structural modeling of SpoIIID predicts an α-helix preceding the HTH, but unlike homeodomain proteins, this helix is predicted to extend to the N-terminus of SpoIIID, leaving no room for an N-terminal arm. Moreover, our mutational analysis indicated that most charge reversals in predicted helix 1 of SpoIIID had little or no effect on its function, as determined by ability to activate expression of *spoIVCA-gusA* reporter during sporulation (Fig 2.5). In contrast, several charge reversals in the highly-conserved C-terminal basic region (residues 63-81) (Fig 2.7) of SpoIIID greatly reduced or abolished reporter expression, as did elimination of the 18 C-terminal residues (K76Stop) (Fig 2.5). The truncated K76Stop SpoIIID protein did not bind to DNA *in vitro* (Table 2.5). We propose that the C-terminal basic region of SpoIIID makes additional contacts with DNA, analogous to the N-terminal arm of eukaryotic homeodomain proteins, allowing a monomer to bind with high affinity.

Another strategy that DNA-binding proteins containing a single HTH use to increase affinity of their interaction with DNA is the winged-helix motif (Brennan, 1993). Hepatocyte nuclear factor 3γ provided the first glimpse of this motif bound to DNA and it was recognized that histone H5 has a similar structure (Clark *et al.*, 1993). These monomeric, αβ proteins have "wings" (loops) that together with the recognition helix of the HTH interact with DNA, resembling a butterfly perched on a rod. Two examples from bacteriophages, Xis (mentioned above) and MuR (a transcriptional repressor of phage Mu), have a single "wing" following their HTH that interacts with the minor groove, and these monomeric, αβ proteins increase their affinity for DNA by cooperative binding (Abbani et al., 2007, Bushman et al., 1984, Sam et al., 2004, Wojciak *et al.*, 2001). Although our results show that SpoIIID does not require cooperativity to achieve high-affinity binding, the possibility that it contains two "wings" following its HTH that allow it to bind DNA with high affinity as a monomer should be considered. Recently, the C-terminal winged-helix domain of the bacterial transcription factor, OmpR, has been shown to bind to DNA as a monomer, although the affinity of the interaction ($K_d \geq 1.5$ μM) was not as high as that measured for SpoIIID (Table 2.4). However, neither our structural modeling nor preliminary NMR data of SpoIIID in complex with Probe 11 DNA (B. Chen, P. H., A. Liu, H. Yan, and L. K., unpublished data) support the idea that SpoIIID is an αβ protein with a winged helix. Rather, both the modeling and the data suggest that the HTH of SpoIIID is followed by an α-helix.

We propose that the C-terminal basic region of SpoIIID is an α-helix that makes additional contacts with DNA, analogous to the N-terminal arm of eukaryotic homeodomain proteins and the "wings" of winged-helix proteins, but structurally distinct.

The structure could be similar to that of the *E. coli* PurR repressor, whose HTH is followed by a "hinge" helix that becomes ordered upon ligand binding and interacts with the minor groove of DNA (Schumacher *et al.*, 1994). PurR is dimeric and other characterized members of the LacI family are dimeric or tetrameric. SpoIIID appears to be unique among prokaryotic DNA-binding proteins with a single HTH in its ability to bind DNA monomerically with high affinity. Efforts to determine the structure of SpoIIID in complex with Probe 11 DNA (B. Chen, P. H., A. Liu, H. Yan, and L. K., unpublished data) promise to reveal the role of the C-terminal basic region and the predicted HTH in DNA binding.

While SpoIIID has been implicated in up- or down-regulation of 122 genes in the σ^E regulon by transcriptome analysis (Eichenberger et al., 2004) and in the regulation of a few genes in the σ^K regulon by biochemical studies (Ichikawa & Kroos, 2000), only 20 SpoIIID binding sites have been mapped by DNase I footprinting (Eichenberger et al., 2004, Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang et al., 1997). The positions of the binding sites suggest that SpoIIID can repress transcription by interfering with RNAP (σ^E or σ^K) or activator (GerE) binding. The two sites from which SpoIIID activates transcription overlap the *spoIVCA* and *sigK* -35 promoter regions, suggesting that SpoIIID contacts RNAP, most likely the sigma factor (i.e., σ^E at the *spoIVCA* promoter, and both σ^E and σ^K at the *sigK* promoter) (Halberg & Kroos, 1994). Part of our motivation for making charge reversal substitutions throughout SpoIIID was to identify residues that might contact RNAP. Such "positive control mutants" would have mutations in *spoIIID* that reduce or eliminate transcriptional activation without impairing

DNA binding. Our mutational analysis identified several charged residues that are likely surface-exposed and are candidates for making contact with RNAP because charge reversal at these positions reduced *spoIVCA-gusA* expression more than threefold: R8, D51, H63, K64, H68, K76, and K78 (Fig 2.5). Altered SpoIIID proteins with charge reversals at these positions are candidates for overexpression, purification, and EMSA studies, as shown here for K34E, H38E, K39E, E43K, and R44E, with the results showing involvement of these residues in DNA binding (Table 2.5). K76 and/or K78 might also be important for DNA binding, because the K76Stop protein failed to bind DNA detectably and the K82Stop protein bound DNA almost normally (Table 2.5). Although R24A and R24E substitutions reduced and eliminated *spoIVCA-gusA* expression, respectively (Fig 2.5), the purified R24A and R24E proteins showed reduced and undetectable binding to Probe 1 DNA, respectively (data not shown), so R24 is important for DNA binding and is therefore not included in the list above of candidate residues for making contact with RNAP. D82 is not a candidate residue for contacting RNAP because D82Stop (lacking D82) activated *spoIVCA-gusA* expression (Fig 2.5) and bound to DNA (Table 2.5), but the D82K substitution is interesting because it eliminates reporter expression (Fig 2.5). We speculate that D82K might prevent K76 and/or K78 from contacting RNAP. Our findings provide a foundation for further studies aimed at elucidating the mechanism of transcriptional activation by SpoIIID.

Chapter III: Summary and Perspectives*

*The figures presented in this chapter are adapted from those generated by Bin Chen for a paper based on the structure of SpoIIID upon which I will be second author (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data).

Summary

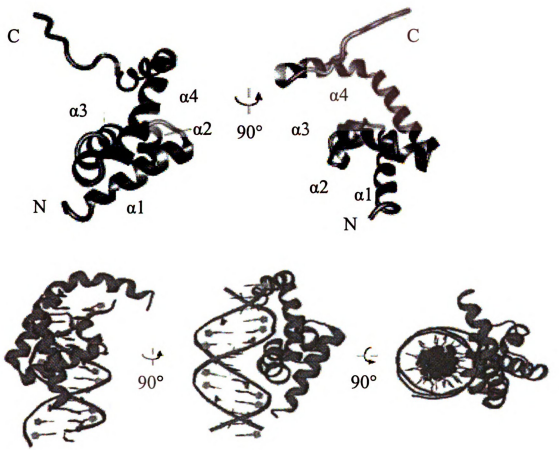
The goal of this project was to determine the requirements, both in SpoIIID and its DNA substrate, for DNA binding by SpoIIID. To that end, we showed that a SpoIIID monomer is able to bind with high affinity to a single match to the SpoIIID-binding consensus sequence on a segment of DNA with as few as 2 base pairs on each side of the consensus. Further, we demonstrated that there are two distinct regions of SpoIIID required for binding to DNA. In addition to an N-terminal region which we have predicted to be the third helix (*i.e.*, the recognition helix) of a traditional HTH motif, we have also identified a C-terminal DNA-binding determinant that appears to relieve the need for the dimerization that is required by many helix-turn-helix proteins (Huffman & Brennan, 2002).

The NMR Solution Structure of SpoIIID

Structural Features of SpoIIID – In conjunction with our studies of DNA binding by SpoIIID, the NMR solution structure of SpoIIID in complex with DNA was determined (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). In agreement with the structural predictions described in Chapter II, the initial three helices form the three-helical bundle that includes an N-terminal α-helix followed by an HTH found in many DNA-binding proteins (Aravind et al., 2005). While, in the predicted structures, a fourth long helix completes the core of the protein, in the NMR structure this sequence is separated into a short helical region (residues 44-48) that extends helix 3, the recognition helix, and is connected to it by a kink, and, after a sharp turn, continues as an α-helix (residues 51-65) that remains within the core of the protein (Fig 3.1) (B. Chen,

Figure 3.1. The solution structure of SpoIIID. The 20 lowest-energy structural models of SpoIIID as determined by NMR were averaged and depicted in a ribbon representation in two different angles with the alpha helices labeled (A). Modeled B-form DNA was docked to SpoIIID (residues 1-81), from which the C-terminal unstructured region had been removed, using docking software and 3 different angles are shown (B). Adapted from figures prepared by Bin Chen (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data).

Fig. 3.1



P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). After a kink at residue 66, the NMR structure shows that helix 4 extends out from the core (residues 67-81) in a manner that bears some similarities to helix 5 (residues 76-85) in the predicted structure based on the cI repressor of bacteriophage λ. Residues 81-93 appear to be disordered in SpoIIID, based on the NMR structure.

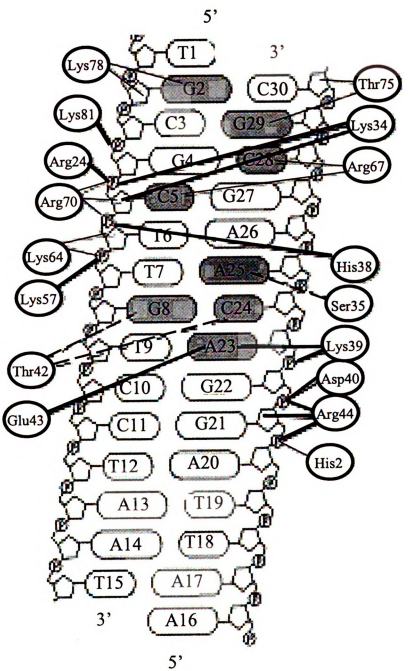
DNA-binding by SpoIIID – Analysis of the structural data provided confirmation of the observations that SpoIIID contains two distinct regions that interact with DNA (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). While the data used for structure determination were collected when SpoIIID was in the presence of probe 11 DNA (see Table 2.4 of Chapter II), the DNA used was not isotopically labeled, so the residues important for protein/DNA interactions must be inferred by two methods. In the first, which identifies residues whose NMR signal is modified by proximity to DNA, two distinct regions of SpoIIID were shown to be significantly near DNA. The first region, encompassing helix 3 (S35, T36, H38, K39, E43, and R44), matches with the data in Chapter II (Table 2.5) describing the effects of mutations to the putative recognition helix on DNA binding *in vitro*. The C-terminal set of residues that this method determines to be near DNA was somewhat N-terminal (K57, K64, R67, G71, A74, T75, and K76) to the region identified by the *in vitro* assay of DNA-binding. Data showing a mutation deleting residues 76-93 (K76Stop) results in a complete loss in DNA binding by SpoIIID, while a mutation truncating residues 81-93 (D82Stop) has little effect, may indicate the importance of residue 76, or it may show that the K76Stop deletion alters the C-terminal structure of SpoIIID such that this region is no longer able to effectively interact with DNA.

Figure 3.2. Interactions between SpoIIID and DNA. Docking and energy minimization software was used to model the interactions between SpoIIID and DNA, depicted in ribbon and space-filling representations respectively, and the side chains important residues in the interaction are represented in ball and stick form (A). The important interactions between SpoIIID and DNA were summarized in schematic form (B). Thick and thin lines between residues and DNA indicate hydrogen bonding with the major and minor grooves, respectively. Dashed lines indicate hydrogen bond interactions mediated by a water molecule. Adapted from figures prepared by Bin Chen (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data).

Fig. 3.2

A

Fig. 3.2 (cont'd.)


B
5'
3'
T1
Lys78
G2
C30
Thr75
Lys81
C3
G29
Lys34
Arg24
G4
C28
Arg67
Arg70
C5
G27
T6
A26
Lys64
T7
A25
His38
Lys57
G8
C24
Ser35
T9
A23
Lys39
Thr42
C10
G22
Asp40
Glu43
C11
G21
Arg44
T12
A20
His2
A13
T19
A14
T18
T15
A17
3'
A16
5'

In addition to directly examining the effects of DNA on the NMR signals of residues of SpoIIID, docking and energy minimization software was used to model SpoIIID onto a model of DNA containing the sequence used in structure determination (Fig 3.2A) (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). Results of this modeling indicate that DNA interacts with SpoIIID in the positively charged cleft between helices 3 and 4 (Fig. 3.2B). Four residues in the HTH motif (S35, K39, T42 and E43) and 3 C-terminal residues (R67, T75, and K78) appear to make base-specific interactions while some of these and a number of other residues from both regions interact with the sugar-phosphate backbone (R24, K34, H38, K39, D40, R44, K57, K64, R70, T75, K78, and K81).

A Structural Basis for Understanding Effects of Amino Acid Substitutions to SpoIIID – The molecular dynamics simulations used to examine SpoIIID/DNA interactions also identified two salt bridges between residues of SpoIIID that appear to stabilize the complex. The first, connecting helices 1 and 3, involves residues R8 and D40 (Fig 3.3A). The importance of these residues is supported both by the result that altered SpoIIID containing a charge reversal substitution, D40K, was unstable during multiple purification attempts and by the findings that *B. subtilis* strains designed to express the charge reversal substitutions R8E or D40K did not express SpoIIID-dependent *spoIVCA-gusA* reporter. A second salt bridge is predicted by the model to connect helices 2 and 4, through interactions between R24 and E73 (Fig 3.3B). While substitutions for R24 impaired DNA binding *in vitro* and decreased SpoIIID-dependent transcription *in vivo*,

Figure 3.3. Intramolecular interactions in SpoIIID. The 3-dimensional structure of SpoIIID identified potential salt-bridge interactions between arginine 8 and aspartate 40 (A) and arginine 24 and glutamate 73 (B). A core of hydrophobic residues between helices 1-4 was also identified (C). Adapted from figures prepared by Bin Chen (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data).

Fig. 3.3

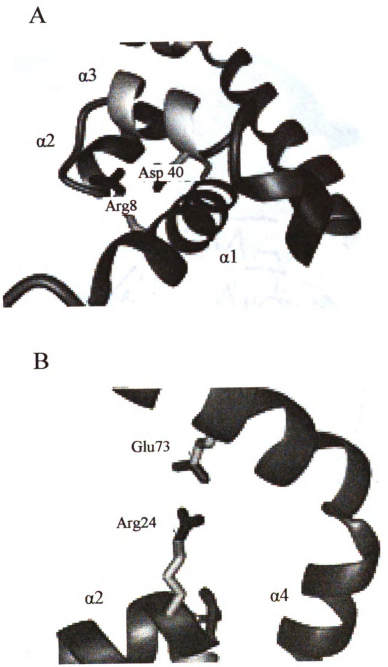
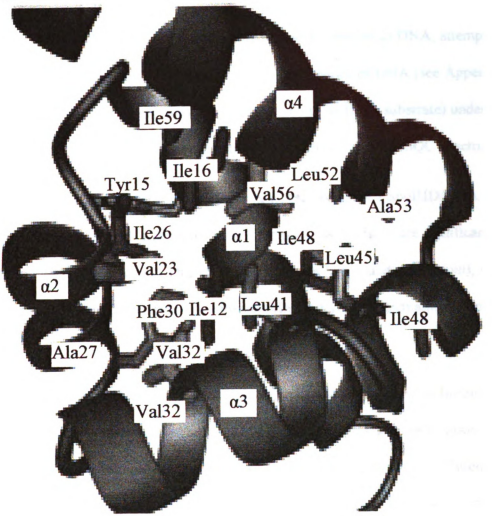


Fig. 3.3 (cont'd.)

C



the E73K substitution did not impair SpoIIID-dependent *in vivo* transcription, so the predicted salt bridge might not be crucial for the structure of SpoIIID and the requirement for R24 could be explained by its interaction with the DNA phosphate backbone.

In addition to determining the structure of SpoIIID bound to DNA, attempts were made to collect structural data for SpoIIID in the absence of DNA (see Appendix 1). Due to the low solubility of apo-SpoIIID (that is, lacking its DNA substrate) under conditions suitable for NMR, we were only able to obtain a single ^{1}H-^{15}N HSQC spectrum.

Comparison of this spectrum to a similar one obtained from a SpoIIID/DNA complex, demonstrates that, with a few exceptions, the peak distributions are significantly different (Appendix 1, Table A1.1). This result (though based on a single spectrum), seems to indicate that SpoIIID undergoes a considerable change in conformation upon binding DNA.

The structure of SpoIIID provides insights into the effects of substitutions to SpoIIID on transcription. Because SpoIIID positively regulates its own transcription 3- to 7-fold (Kunkel et al., 1989, Stevens & Errington, 1990), we were unable to differentiate between SpoIIID mutants that were defective in activation of their own promoter and mutants that were unstable *in vivo*. In an attempt to circumvent this problem, the coding region for *spoIIID* was fused to the *spoIID* promoter, which does not require SpoIIID for activation (Eichenberger et al., 2004). These fusions were unable to produce SpoIIID that accumulated *in vivo* (see Appendix 2). Because we are unable to distinguish between altered proteins that fail to accumulate and those that are unable to activate transcription, residues such as I26 and V37, which, when substituted, resulted in

complete loss of *in vivo* transcription, may be causing their effects not because they are involved directly in transcription activation or DNA-binding, but because introducing a charge at those positions disrupts the hydrophobic core indicated by the structure (Fig. 3.3C) (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data) and thereby destabilizes the protein so that it is unable to accumulate. The valine at residue 23 is positioned near this hydrophobic core and may interact with it as well (Fig. 3.3C), and the addition of a charge at this position may also affect the stability of SpoIIID. Oddly substitution of a charged residue (lysine) for L41, which also appears to contribute to the hydrophobic core, resulted in a slight increase in *spoIVCA-gusA* expression. Residues 31-33 comprise the turn of the HTH motif, so introduction of bulky charged residues at these positions could also result in a less stably folded SpoIIID structure, explaining the dramatic effects of the V32E and S33R mutations on transcription *in vivo*. Also, the purified S33R protein exhibited about 10-fold lower affinity for DNA (Chapter II, Table 2.5).

Many of the other mutations identified as resulting in a loss of transcription *in vivo*, alter residues that appear from structural data to be directly involved in interactions with, or in very close proximity to, the DNA (Fig 3.2B) (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). These residues include K34, S35, T36, H38, K39, E43, R44, K64, K76, and K78. Interestingly, charge reversal substitutions K57E and K81E did not impair transcription *in vivo*, although these two residues are predicted to interact with the phosphates in the backbone of DNA, suggesting these contacts are not essential.

The mutational analysis of SpoIIID did reveal, however, 4 mutations that effect transcription *in vivo* but are not in residues expected to interact with DNA. All but one of

these residues, D51, are adjacent to residues that are predicted to be involved in protein/DNA interactions (Fig. 3.2) (B. Chen, P. Himes, A. Liu, H. Yan, and L. Kroos, unpublished data). The aspartate at residue 51 may be on the same face of helix 4 as lysine 57, which as noted above is predicted to contact a phosphate in the DNA backbone but the K57E substitution had no effect on transcription *in vivo*. In any case, the negatively charged aspartate at position 51 is not predicted to interact with DNA. A substitution of lysine for aspartate at residue 82 resulted in a nearly 20-fold decrease in activation of transcription (Chapter II, Fig. 2.5). This result was unexpected in light of the lack of effects on transcription of the K81E substitution (despite K81 being predicted to contact a phosphate in the DNA backbone as noted above) and the D82Stop truncation. Because aspartate 82 is the first residue of the disordered region, it is possible that a negative charge at this location has an effect on the overall flexibility of the region that negatively impacts DNA binding. Alternatively, as suggested in Chapter II, the D82 substitution might prevent K76 and/or K78 from contacting RNAP, but, in the light of the structural data, K76 and/or K78 might contact DNA. The other two residues at which substitutions result in a significant loss in activation of *in vivo* transcription, histidines at positions 63 and 68, while adjacent to residues predicted to be important for DNA-binding, are on the opposite face of the helix and have long flexible side chains with positive charges, so it is attractive to hypothesize that these positive charges could serve as the basis for a charge-charge interaction involved in recruiting RNA polymerase to SpoIIID-dependent promoters. The potential importance of histidine 63 is further emphasized by the observation that a substitution with a glutamate at that residue results in a 10-fold loss of activation of transcription *in vivo*, while substitution of glutamate for

the adjacent lysine 64, which the structure-based modelling depicts as being involved in DNA-binding, only causes a 3-fold decrease. Further emphasizing the importance of a positive charge in this region, a substitution for a positive charge at a nearby residue (D61E) results in a 2-fold increase in *in vivo* transcription.

Future Directions

To further examine the role of SpoIIID in activation of transcription, it would be interesting to further examine the effects of the substitutions at positions 51, 63, 68, and 82. To confirm that these residues are not involved in DNA binding, the altered SpoIIID would be over expressed in *E. coli*, purified, and used subjected to *in vitro* DNA binding assays as in Chapter 2. Any mutant that retains the ability bind DNA with wild-type affinity would then be assayed for its ability to activate transcription *in vitro* using RNA polymerase containing σ^E. To that end, I have transformed DNA encoding a histidine-tagged β'-subunit into a strain of *B. subtilis* that contains a deletion of the gene encoding σ^G. This should simplify purification of σ^E RNAP. *In vitro* transcription assays would be performed on a SpoIIID- and σ^E-dependent promoter, presumably that of *spoIVCA*, though the *sigK* promoter may be interesting to study as well because it is subject to more complex regulation. Finally, should one or several of the altered SpoIIID proteins be demonstrated to be competent for DNA binding but not transcriptional activation, a comparison of their ability with that of wild-type SpoIIID to cause a supershift in *in vitro* DNA-binding assays, when mixed with a radio-labeled SpoIIID-dependent promoter and σ^E-dependent RNA polymerase, could be used to determine whether the altered proteins

have a defect in the ability to recruit RNA polymerase, or whether the defect occurs at some later stage like open complex formation.

Because the RNAP used in previous studies to demonstrate that SpoIIID activates transcription by σ^E RNAP *in vitro* was only partially purified (Halberg & Kroos, 1994), it is possible that *in vitro* transcription experiments using wild-type SpoIIID and the histidine-tagged σ^E RNAP described above would be unsuccessful due to the requirement for a previously-unidentified cofactor in SpoIIID/RNA polymerase interactions. Should this be the case, a *B. subtilis* strain expressing altered SpoIIID that is competent to bind DNA *in vitro*, but unable to support transcription *in vivo*, could be subjected to mutagenesis to identify a mutation that allows SpoIIID-dependent transcription. If wild-type SpoIIID does activate transcription by highly purified σ^E RNAP *in vitro*, but one or more of the altered SpoIIID proteins does not, it would still be interesting to screen for suppressor mutations, though in a more directed manner. The *sigE* gene would be the most likely location of such mutations, but, because SpoIIID also activates σ^K-dependent transcription of *sigK* (Kroos et al., 1989), they may be located in the *sigK* gene or the gene for a subunit of the RNAP core, most likely for the α-subunit based on paradigms from other transcriptional activators.

It would also be interesting to characterize the conformational shift in SpoIIID that my preliminary results indicate occurs upon DNA binding. While under the conditions used, apo-SpoIIID did not appear to be stable in high enough concentrations for NMR-based structure determination, other conditions may be found that allow better protein stability

without interfering with NMR data collection. Alternatively, though I was unable to get crystals of SpoIIID to form in the presence or absence of DNA, truncating the C-terminal 12 residues of SpoIIID which are not essential for DNA binding and are disordered when SpoIIID is bound to DNA may allow for the formation of crystals. Analysis of the apparent conformational shift in SpoIIID may provide new insights into how SpoIIID interacts with DNA, and, coupled with identification of the residues involved in transcriptional activation, allow a more complete picture of the mechanisms underlying transcriptional activation in bacteria.

Appendices

Appendix 1: Preparation of SpoIIID for NMR-based structural analysis

Abstract

While a number of predictions have been made with regard to the structural properties of SpoIIID and the manner in which it contacts DNA, a better understanding of the mechanisms by which SpoIIID functions would be provided by the determination of the structure of SpoIIID. To that end, ^{15}N-singly labeled and ^{15}N-, ^{13}C-doubly labeled SpoIIID was prepared and purified. SpoIIID, either mixed with DNA or free, was analyzed via NMR. Although SpoIIID in the absence of DNA appears to be have been unstable under the conditions tested, precluding a definitive result, less than 25% of the peaks in the spectrum obtained could be aligned with the spectra of SpoIIID in the presence of DNA, suggesting that the majority of the amino groups resided in different chemical environments depending on whether SpoIIID is bound to DNA or free.

Introduction

SpoIIID is a small (10.8 kDa, 93 amino acids) protein that binds DNA in a sequence-specific manner to both activate and repress transcription (Eichenberger et al., 2004, Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Kunkel et al., 1989, Zhang et al., 1997). SpoIIID has previously been predicted to make use of a helix-turn-helix motif to interact with DNA (Kunkel et al., 1989). In Chapter II, we describe more detailed predictions of the 3-dimensional structure of SpoIIID that suggest that it is comprised of a 4 α-helix bundle potentially followed by a fifth α-helix that extends away from the protein. We also identify two distinct regions of SpoIIID that are required for interactions between a monomer of SpoIIID and DNA. While DNA-binding studies suggest that when SpoIIID functions as a transcriptional repressor it binds at or near the

start site of transcription or the binding site of a transcriptional activator and may competitively inhibit interactions between the activator or RNA polymerase with DNA, when SpoIIID activates transcription it binds near the -35 region where it is ideally positioned to recruit RNAP to promoters to which RNAP has low affinity (Eichenberger et al., 2004, Halberg & Kroos, 1994, Ichikawa & Kroos, 2000, Zhang et al., 1997). In an attempt to identify residues important for the recruitment of RNAP to SpoIIID-dependent promoters, in Chapter II we performed a screen of a large number of substitutions in SpoIIID that either reversed the charges at many positions or introduced charges at positions where there had previously been an uncharged residue. Mutations for that screen were chosen based on their charge, under the assumption that charged residues would be solvent-exposed, or on their location in the 3-dimensional models of SpoIIID. To determine the structure of SpoIIID, which would provide insights into which of the residues that we identified in Chapter II as being deficient in transcriptional activation *in vivo* were not involved in DNA-binding as well as provide other insights into the function of SpoIIID, we purified isotopically-labeled SpoIIID protein and subjected it to NMR analysis both in the presence and absence of DNA. The resulting spectra have few peaks in common, suggesting that SpoIIID undergoes a substantial change in conformation upon DNA-binding.

Experimental Procedures

Overexpression and Purification of SpoIIID – A small colony of *Escherichia coli* BL21 (DE3) (Novagen) strain containing the plasmid pPH7 (see Chapter II, Table 2.2) was selected as described in Chapter II to inoculate 2 l of M9 medium (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 18 mM ^{15}N-NH_4Cl, 0.1 mM $CaCl_2$, 0.2%

(w/v) dextrose, supplemented with 2 ml Advanced Formula Multivitamin (Meijer) extract [crushed tablet was mixed with 3 ml 1 N HCl, and solids were cleared]) containing 150 μg/ml ampicillin. When SpoIIID was ^{15}N, ^{13}C-doubly labeled, 0.2% (w/v) ^{13}C-glucose was substituted for 0.2% (w/v) dextrose. The culture was grown and expression of SpoIIID was induced as described in Chapter II. SpoIIID was purified over a SP Sepharose column (GE Healthcare) as described in Chapter II for SpoIIID prepared for use in analytical ultracentrifugation, but the column was 7.5 ml rather than the 1 ml column used in Chapter II. The protein was further purified over the 1-ml HiTrap™ Heparin HP column (GE Healthcare) as described in Chapter II for SpoIIID used for analytical ultracentrifugation.

Preparation of Samples for NMR Analysis – Oligonucleotides containing the sequences 5'-attaggacaagcgc-3' and 5'tgcgcttgtcctaa-3' were obtained from W.M. Keck Oligonucleotide Synthesis Facility, Yale University, resuspeded in 10 mM potassium phosphate buffer pH 7 supplemented with 50 mM NaCl, and equimolar concentrations were boiled for 10 minutes and, then, allowed to cool to room temperature. SpoIIID purified as described above was diluted tenfold in 10 mM potassium phosphate buffer pH 7 and 500 nmol SpoIIID was mixed with 600 nmol annealed DNA inverted several times to mix and stored on ice for 1 hour. Alternatively, for SpoIIID in the absence of DNA, SpoIIID purified as described above was diluted in 10 mM potassium phosphate buffer pH 7 to a final NaCl concentration of 0.6 M. The SpoIIID or SpoIIID/DNA mixture was concentrated to a final volume of 0.5 ml using Amicon Ultra 4 (5K MWCO) (Millipore) centrifugal filter devices, then transferred to a microcentrifuge tube and centrifuged at

16,000 x g for 10 minutes at 4°C, and the supernatant was transferred to a fresh tube. Samples were supplemented with 50 μM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) and 1% (w/v) sodium azide and 10% D_20 and transferred to NMR tubes for analysis.

Results and Discussion

Stability of SpoIIID in the absence of DNA – We have previously shown that, at concentrations suitable for structural studies (e.g., 500 μM – 1mM SpoIIID), requires high concentrations of salt to remain stable in the absence of DNA (P.H and L.K, unpublished data). For example, using dialysis to reduce the concentration of NaCl from 1 M in 100 mM increments, SpoIIID spontaneously precipitated when the NaCl concentration decreased below 600 mM (or 500 mM NaCl when the buffer was supplemented with 0.1% (v/v) Triton X-100). While there was no visible precipitation following concentration or subsequent centrifugation, the NMR signal resulting from SpoIIID in the absence of DNA, was significantly smaller than expected and as a result a large number of scans had to be taken to amplify the signal. This implies that there was a substantial decrease of protein in the SpoIIID sample in the absence of DNA during or after concentration and that 600 mM NaCl was insufficient to maintain protein stability in these conditions.

Comparison of NMR Spectra of SpoIIID in the Presence or Absence of DNA – As can be seen in Table A1.1, when we compared the ^{1}H-^{15}N HSQC spectrum of SpoIIID in the presence of DNA with that from free SpoIIID only a small portion of the peaks (each peak corresponds to an amino group in the protein) aligned. While each of the spectra

had over 100 peaks only 26 peaks had coordinates in common. Although this experiment was only performed once, and on one protein sample that appeared to be unstable, this result implies that the majority of the amino groups, and therefore residues, of SpoIIID are in different chemical environments in the two conditions analyzed. This result leads us to hypothesize that the conformation of SpoIIID undergoes a significant change upon interaction with DNA.

Table A1.1 Comparison of the Coordinates of ^{1}H-^{15}N HSQC spectra of SpoIIID in the presence or absence of DNA

SpoIIID with DNA			Free SpoIIID			
Peak no.[a]	N (ppm)	HN (ppm)	Peak no.[a]	N (ppm)	HN (ppm)	Match no.[a]
			0	130.48	7.962	
			1	128.679	8.312	
			2	128.672	7.899	
0	126.936	8.356				
1	126.784	8.515				
2	126.779	8.446				
3	126.779	8.434				
4	126.784	8.429				
5	126.779	8.338	3	126.699	8.309	1
			4	126.644	7.791	
			5	125.555	8.857	
6	124.948	6.622				
7	124.805	8.41				
8	124.713	8.755				
9	124.713	8.498				
10	124.705	8.674				
11	124.645	8.406	6	124.62	8.338	2
			7	124.568	8.457	
			8	124.349	9.091	
			9	124.374	7.831	
12	124.252	9.243				
			10	124.21	7.891	
13	124.169	8.374	11	124.157	8.337	3
			12	124.153	8.233	
14	123.993	8.207				
15	123.985	7.617				
16	123.965	7.805				
17	123.896	7.382				
			13	123.861	8.175	
			14	123.747	8.551	
18	123.439	8.278				
19	123.443	8.256				
20	123.441	8.228				
21	123.43	8.341				
22	123.431	8.299				
23	123.435	8.083	15	123.406	8.061	4
			16	123.358	8.136	
24	123.404	7.653	17	123.369	7.61	5

Table A1.1 (cont'd).						
			18	123.174	7.829	
25	123.21	8.523	19	123.145	8.485	6
26	123.062	8.285	20	123.071	8.281	7
			21	123.017	9.006	
			22	122.898	7.307	
			23	122.823	8.448	
			24	122.836	6.996	
			25	122.788	8.083	
			26	122.723	8.615	
			27	122.487	8.03	
27	122.47	8.473				
			28	122.439	7.709	
28	122.404	7.523				
29	122.325	7.83				
30	122.238	8.701				
31	122.236	8.557				
32	122.253	8.493				
33	122.256	8.487				
34	122.262	8.168				
35	122.231	8.708				
36	122.231	8.573				
			29	122.23	9.073	
37	122.216	8.372				
38	122.197	10.855				
			30	122.179	8.347	
			31	122.148	8.129	
39	122.075	6.337				
40	122.022	8.35				
41	121.986	7.551	32	121.942	7.536	8
			33	121.913	7.852	
42	121.813	9.118				
43	121.789	8.325				
44	121.786	8.321	34	121.728	8.376	9
45	121.807	7.769				
46	121.768	8.188	35	121.735	8.15	10
47	121.758	7.193				
			36	121.704	7.801	
48	121.698	8.401				
			37	121.533	8.104	
			38	121.459	7.191	
49	121.456	8.31				
			39	121.445	8.674	
50	121.42	9.038				
			40	121.405	7.345	

Table A1.1 (cont'd).						
			41	121.301	8.046	
51	121.254	9.572				
52	121.267	7.322				
53	121.242	7.813				
54	121.126	8.616				
55	121.125	8.61				
56	121.112	8.501				
57	121.112	8.416				
58	121.135	8.275				
59	121.103	8.203				
60	121.072	8.338				
61	121.036	8.064				
			42	121.019	8.733	
			43	120.99	8.332	
			44	120.99	7.794	
			45	120.939	8.183	
62	120.896	8.484				
63	120.897	8.431				
64	120.898	8.418				
65	120.89	8.587				
66	120.89	8.496				
67	120.888	8.257	46	120.862	8.232	11
68	120.809	8.141				
69	120.746	8.209				
			47	120.659	10.443	
			48	120.594	8.895	
70	120.547	8.28				
			49	120.531	8.004	
			50	120.499	8.453	
			51	120.422	8.779	
			52	120.427	8.709	
			53	120.372	8.222	
			54	120.299	8.147	
71	120.202	8.487				
72	120.117	7.8	55	120.085	7.614	12
73	119.934	8.045				
			56	119.915	8.563	
74	119.863	8.908	57	119.901	8.852	13
75	119.843	7.591				
			58	119.683	9.067	
76	119.618	8.594	59	119.676	8.467	14
			60	119.615	7.333	
77	119.512	9.022				
78	119.475	8.685				

Table A1.1 (cont'd).						
79	119.354	7.859	61	119.375	7.79	15
80	119.345	7.624				
			62	119.331	8.403	
81	119.292	7.812				
			63	119.25	8.745	
82	119.172	8.379				
83	119.134	8.471	64	119.116	8.484	16
84	119.143	8.388				
			65	119.033	7.908	
			66	118.964	7.829	
			67	118.889	8.715	
			68	118.754	8.01	
85	118.645	8.247				
			69	118.403	7.469	
86	118.347	7.85				
87	117.855	8.061	70	117.913	8.04	17
88	117.868	7.975				
89	117.84	8.868				
			71	117.796	7.981	
90	117.734	7.073				
			72	117.664	7.906	
91	117.488	8.318	73	117.489	8.517	18
92	117.33	7.774				
93	116.877	8.122	74	116.826	8.279	19
			75	116.539	7.119	
			76	116.465	8.257	
			77	116.422	8.17	
94	116.329	8.337				
			78	116.294	7.45	
			79	116.033	7.265	
95	115.947	7.67				
96	115.972	7.401				
97	115.922	8.568	80	115.921	8.182	20
98	115.856	7.124				
			81	115.819	7.928	
99	115.768	7.929				
			82	115.664	8.495	
100	115.566	8.317				
101	115.46	7.411				
			83	115.425	8.649	
102	114.865	8.49				
103	114.785	8.609				
104	114.617	8.034				
			84	114.483	8.389	

Table A1.1 (cont'd).						
105	114.467	9.181				
106	114.405	8.003				
107	114.341	7.951				
			85	114.28	8.257	
			86	113.584	8.077	
108	113.171	6.851				
			87	113.214	8.294	
			88	113.153	7.573	
109	113.029	6.852				
			89	112.965	7.979	
			90	112.951	7.161	
			91	112.893	8.738	
110	112.891	7.588				
111	112.891	5.851				
112	112.754	7.549				
			92	112.808	7.783	
113	112.73	6.909	93	112.789	6.906	21
114	112.698	7.816				
115	112.638	6.909				
116	112.612	7.515				
117	112.615	6.8				
118	112.46	6.851				
			94	112.425	7.975	
119	112.218	7.586	95	112.239	7.557	22
120	112.207	6.855	96	112.223	6.845	23
121	112.101	6.956				
122	112.069	6.854				
			97	112.004	7.961	
			98	111.824	6.945	
			99	111.442	8.515	
123	111.172	8.18				
			100	111.166	7.976	
124	110.884	7.648				
125	110.805	8.608				
126	110.188	8.354	101	110.026	8.304	24
127	109.956	8.318				
128	109.866	8.064				
129	109.673	8.815				
			102	109.524	8.32	
130	109.182	6.978				
			103	109.161	8.622	
			104	108.969	8.26	
			105	107.075	7.665	
131	105.734	8.104	106	105.715	8.069	25

Table A1.1 (cont'd).						
			107	105.067	8.077	
132	104.58	7.446	108	104.432	7.453	26
133	103.742	7.819				

[a]The peaks are numbered for each spectrum separately in order of descending nitrogen signal (from 130 to 103 ppm). Peaks of similar characteristics between the two spectra are aligned and given a match number, again organized by descending nitrogen signal, identify the 26 peaks that may correspond to amino groups in the same chemical environment when SpoIIID is with or without DNA.

Appendix 2: Removing *spoIIID* from Dependence on SpoIIID for Transcriptional Activation

Abstract

Because SpoIIID has been shown to positively autoregulate transcription of *spoIIID*, amino acid substitutions in SpoIIID that cause decreased stability are indistinguishable from substitutions that impair the ability of SpoIIID to activate transcription when SpoIIID accumulation is assayed. In an attempt to circumvent the SpoIIID-dependence of *spoIIID*, the *spoIIID* coding region was fused to the promoter of a gene that is expressed at a similar time in sporulation to similar levels. When this fusion was inserted into the chromosome of a *B. subtilis* strain unable to express *spoIIID* from its native locus, the fusion was unable to support SpoIIID-dependent gene expression significantly above background levels.

Introduction

While the gene encoding the *Bacillus subtilis* transcription regulator SpoIIID is transcribed by σ^E RNA polymerase, full expression of *spoIIID* is dependent on the presence of SpoIIID (Kunkel et al., 1989, Stevens & Errington, 1990). This requirement for SpoIIID has been proposed to be the basis for the delay of peak expression of *spoIIID* after activation of σ^E (Steil et al., 2005). Surprisingly, there does not appear to be a match to the SpoIIID DNA-binding consensus sequence 5'-WWRRACARNY-3' (where W is A or T, R is purine, Y is pyrimidine, N is any nucleotide) (Halberg & Kroos, 1994) near the *spoIIID* promoter. Expression of *spoIIID* is further regulated by the presence of a small upstream cistron, *usd*, with which *spoIIID* is co-transcribed (Decatur et al., 1997). While there is no known function for the peptide encoded by *usd*, its translation has been proposed to be required to disrupt a stem-loop structure in the mRNA which otherwise

renders the *spoIIID* Shine-Dalgarno sequence inaccessible to the ribosome and this is supported by the finding that mutation disrupting the *usd* ribosome-binding sequence, causes a lack of SpoIIID accumulation (Decatur et al., 1997).

Expression of the *spoIID* gene, the product of which is an essential part of the engulfment apparatus (Abanes-De Mello et al., 2002), is also dependent on σ^E RNA polymerase (Eichenberger et al., 2004, Tatti *et al.*, 1995). Although expression of *spoIID* is repressed by SpoIIID (Eichenberger et al., 2004), even in the presence of SpoIIID, expression levels of *spoIID* are approximately threefold higher than those of *spoIIID* (Tatti et al., 1995). Here we present the results of a fusion between the wild-type or a mutant *spoIID* promoter and *spoIIID*.

Experimental Procedures

Plasmids – The plasmids and oligonucleotides used in this study are described in Tables A2.1 and A2.2 respectively.

Construction of spoIID-spoIIID *fusions* – To create the *spoIID(-31AT)-spoIIID* fusion, primers LK1061 and LK1062 were used to PCR amplify the *spoIID* promoter (-291 to -16 relative to translational start with the A to T mutation at -31 relative to transcriptional start) from pSpoIID-LacZ -31AT. The product was purified and 200 ng of the reaction was used as the upstream primer in a second PCR reaction with primer LK1063 to amplify the *spoIIID* coding region (-21 to +303 with respect to translational start) from pPH1. The product of this reaction was used with primers LK1061 and LK1063 in a third PCR reaction to amplify the product. Alternatively, for the *spoIID(-31AT)-H(6)-spoIIID* fusion, primers LK1061 and LK1207 were used to PCR amplify the *spoIID*

promoter (-291 to -16 relative to translational start) from pSpoIID-LacZ -31AT. The product was purified and 200 ng was used as the upstream primer in a second PCR reaction with primer LK1063 to amplify the *spoIIID* coding region (+1 to +363 with respect to the translational start) containing the N-terminal histidine tag from pPH26. Between the sequences contained in pSpoIID-LacZ -31AT and pPH26, LK1207 contains sequence corresponding to the 5' untranslated region (-21 to -1 with respect to translational start), including the ribosome binding site, of *spoIIID*. The product of this reaction was used with primers LK1061 and LK1063 in a third PCR reaction to amplify the product. The -31AT mutation was converted to wild-type sequence by using primers LK1345 and LK1346 and the QuikChange site-directed mutagenesis kit (Stratagene) on plasmids pPH35 and pPH36 to make plasmids pPH47 and pPH48, respectively. All products of fusions and mutagenesis were sequenced at the Michigan State University Genomics Technology Support Facility to ensure that the desired sequence was present.

Construction of B. subtilis *Strains that Express* spoIID-spoIIID *fusions from the* thrC *Locus* – Competent PH1001 cells were transformed with pPH35, pPH36, pPH47, and pPH48 to create PH2035, PH2036, PH2047, and PH2048 as described in Chapter II. Verification of double crossover insertion into the *thrC* gene and no insertion into *spoIIID* was performed as described in Chapter II.

Measurement of spoIVCA-gusA *Reporter Expression* – Sporulation and sample collection were performed as described in Chapter II. Enzymatic activity of each sample was assayed, activity of PH2000 strain (background activity) was subtracted and the average of the activity as a percentage of PH2001 (wild type) activity for 3 biological replicates was determined as described in Chapter II.

Table A2.1. Descriptions of certain plasmids used in this study			
Plasmid	Description	Construction	Reference or Source
pSpoIID-LacZ - 31AT	Ap^r Nm^r; vector to make *gusA* fusions for integration at *amyE*		(Eichenberger et al., 2004, Tatti et al., 1995)
pAK3	Ap^r Sp^r; vector for integration at *thrC*		V. Chary and P. Piggot
pET-28a	Kan^r; T7; C-terminal His(6) tag		Novagen
pPH26	Ap^r; T7-H(6)-*spoIIID*	the *spoIIID* coding region (+2 to +303 with respect to translational start) was amplified by PCR using LK566 and LK567 as primers and PY79 chromosomal DNA as template, then the NdeI and BamHI digested PCR product was inserted between the NdeI and BamHI sites of pET-28a	This work
pPH1	Ap^r Sp^r; *spoIIID*		Chapter II
pPH35	Ap^r Sp^r; $P_{spoIID(-31AT)}$ -*spoIIID* transcriptional fusion	the *spoIID(-31AT)-spoIIID* fusion was created as described in Experimental Procedures then the EcoRI-digested PCR product was inserted into the EcoRI site of pAK3	This work
pPH36	Ap^r Sp^r; $P_{spoIID(-31AT)}$ -*H(6)-spoIIID* transcriptional fusion	the *spoIID(-31AT)-H(6)-spoIIID* fusion was created as described in Experimental Procedures then the EcoRI-digested PCR product was inserted into the EcoRI site of pAK3	This work
pPH47	Ap^r Sp^r; P_{spoIID}- *spoIIID* transcriptional fusion	the -31AT mutation in pPH35 was converted to wild-type sequence with site-directed mutagenesis using primers LK1345 and LK1346	This work
pPH48	Ap^r Sp^r; $P_{spoIID)}$-*H(6)-spoIIID* transcriptional fusion	the -31AT mutation in pPH36 was converted to wild-type sequence with site-directed mutagenesis using primers LK1345 and LK1346	This work

Abbreviations: Ap^r, ampicillin-resistant; Nm^r, neomycin-resistant; Sp^r, spectinomycin-resistant; Kan^r, kanamycin-resistant; T7, T7 RNA polymerase promoter and a translation initiation sequence; H(6) N-terminal histidine tag sequence from pET-28a.

Table A2.2. Oligonucleotides used in this study	
Primer	Sequence[a]
LK1345 (-31TA)[b]	5'-ccaaaacgagagtcatattagcttgtccctgccc-3'
LK1346	5'-gggcagggacaagctaatatgactctcgttttgg-3'
LK566 (SpoIIID +2 to +303)[c]	5'-gggaattccatatgcacgattacatcaaagagcgaac-3'
LK567	5'-cgggatcccaagaaggcaatgccagg-3'
LK1061 (*spoIID* -291 to -16)[c]	5' -ccgctcgaggaattcaagcttgccgctctgggcgc-3'
LK1062 (*spoIIID* -21 to +303)[c]	5'-caccactcgacctccctaaaa*g*ctcgggattcgactctagt-3'
LK1063	5'-cgggatccgaattccaagaaggcaatgccaggg-3'
LK1061 (*spoIID* -291 to -16)[c]	5' -ccgctcgaggaattcaagcttgccgctctgggcgc-3'
LK1207 (*H(6)-spoIIID* +1 to +363)[c]	5'-gatgatgatggctgctgcgga*t*accactcgacctccctaaaa*g*ctcgggattcgactctagtc -3'
LK1063	5'-cgggatccgaattccaagaaggcaatgccaggg-3'

[a] Boldface type indicates location of mutation in sequence.

[b] For the primer pair used in site-directed mutagenesis, the substitution is indicated in parentheses, the top primer indicates the coding strand, the bottom primer indicates the template strand.

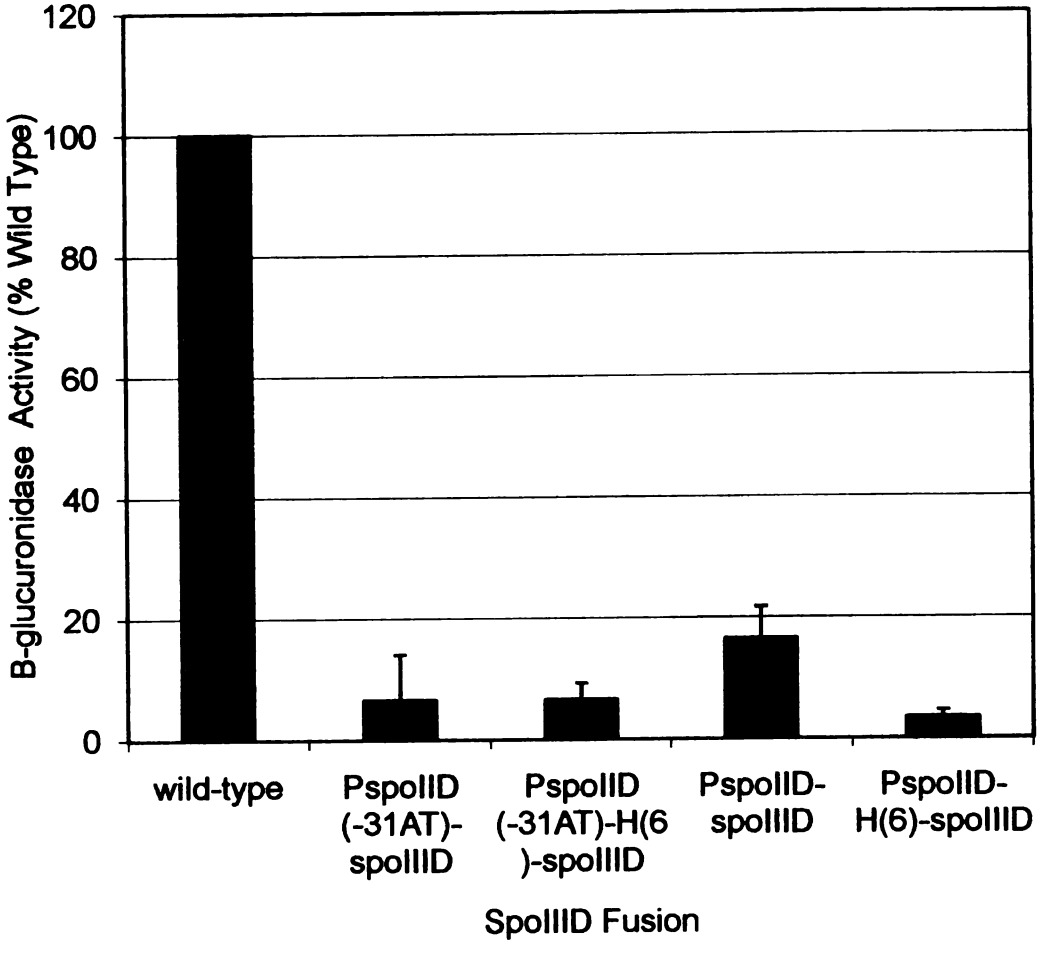
[c] For each primer set used to clone sequences, the boundaries of the region (in nucleotides, with respect to the start of translation) are indicated in parentheses after the name of the primer, the upper primer is the primer complementary to the upstream end of the sequence cloned, and the lower primer is complementary to the downstream end. Sequences not part of the region cloned are underlined. Italicized type indicates the final base of each sequence when two sequences are joined.

Results and Discussion

Expression of spoIVCA-gusA *in a* B. subtilis *strain engineered to produce SpoIIID under the control of the* spoIID *promoter* – While expression from translational fusions of the *spoIID* promoter to *lacZ* are approximately threefold higher than similar fusions of *spoIIID* to *lacZ*, a single-base pair A to T substitution at -31 in the *spoIID* promoter reduced expression to approximately threefold less than that seen using the *spoIIID* promoter (Tatti et al., 1995). Both wild-type and -31AT versions of the *spoIID* promoter were transcriptionally fused to SpoIIID. The ultimate base of the *spoIID* promoter in these fusions immediately preceded the Shine-Dalgarno sequence from *spoIID* and the next base was the first base of the *spoIIID* Shine-Dalgarno sequence to eliminate the 5' portion of the stem-loop that is believed to sequester the *spoIIID* Shine-Dalgarno sequence at its native locus (Decatur et al., 1997). Because the polyclonal α-SpoIIID antibodies we use to detect SpoIIID primarily interact with the C-terminus of SpoIIID and we planned to perform C-terminal truncations of SpoIIID, similar fusions were made containing the N-terminal polyhistidine tag from pET-28a, so that they could be used to detect SpoIIID with α-His(6) antibodies. None of these fusions, when transformed into *B. subtilis*, were competent to activate expression of the SpoIIID-dependent *spoIVCA-gusA* transcriptional fusion (Figure A2.1). Because the wild-type SpoIIID is encoded by these fusions, we are unable to determine whether there is an inherent flaw in these constructs or if SpoIIID more fully represses the *spoIID* promoter when that promoter drives SpoIIID expression than when expression of SpoIIID is driven by the *spoIIID* promoter.

Figure A2.1. Quantification of the effects of *spoIID-spoIIID* fusions on transcription *in vivo*. β-glucuronidase activity as a percentage of wild type levels was determined as described in Experimental Procedures. Bars indicate averages of 3 biological replicates and error bars indicate one standard deviation.

Fig. A2.1


120
100
80
60
40
20
0
B-glucuronidase Activity (% Wild Type)
wild-type
PspoIID (-31AT)-spoIIID
PspoIID (-31AT)-H(6)-spoIIID
PspoIID-spoIIID
PspoIID-H(6)-spoIIID
SpoIIID Fusion

References

REFERENCES

Abanes-De Mello, A., Y. L. Sun, S. Aung & K. Pogliano, (2002) A cytoskeleton-like role for the bacterial cell wall during engulfment of the Bacillus subtilis forespore. *Genes Dev* 16: 3253-3264.

Abbani, M. A., C. V. Papagiannis, M. D. Sam, D. Cascio, R. C. Johnson & R. T. Clubb, (2007) Structure of the cooperative Xis-DNA complex reveals a micronucleoprotein filament that regulates phage lambda intasome assembly. *Proc Natl Acad Sci U S A* 104: 2109-2114.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman, (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.

Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu & L. M. Iyer, (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* 29: 231-262.

Arigoni, F., L. Duncan, S. Alper, R. Losick & P. Stragier, (1996) SpoIIE governs the phosphorylation state of a protein regulating transcription factor sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* 93: 3238-3242.

Atrih, A., P. Zollner, G. Allmaier & S. J. Foster, (1996) Structural analysis of Bacillus subtilis 168 endospore peptidoglycan and its role during differentiation. *J Bacteriol* 178: 6173-6183.

Bagyan, I., J. Hobot & S. Cutting, (1996) A compartmentalized regulator of developmental gene expression in Bacillus subtilis. *J Bacteriol* 178: 4500-4507.

Behravan, J., H. Chirakkal, A. Masson & A. Moir, (2000) Mutations in the gerP locus of Bacillus subtilis and Bacillus cereus affect access of germinants to their targets in spores. *J Bacteriol* 182: 1987-1994.

Ben-Yehuda, S. & R. Losick, (2002) Asymmetric cell division in B. subtilis involves a spiral-like intermediate of the cytokinetic protein FtsZ. *Cell* 109: 257-266.

Ben-Yehuda, S., D. Z. Rudner & R. Losick, (2003) RacA, a bacterial protein that anchors chromosomes to the cell poles. *Science* 299: 532-536.

Blaylock, B., X. Jiang, A. Rubio, C. P. Moran, Jr. & K. Pogliano, (2004) Zipper-like interaction between proteins in adjacent daughter cells mediates protein localization. *Genes Dev* 18: 2916-2928.

Boland, F. M., A. Atrih, H. Chirakkal, S. J. Foster & A. Moir, (2000) Complete spore-cortex hydrolysis during germination of Bacillus subtilis 168 requires SleB and YpeB. *Microbiology* 146 (Pt 1): 57-64.

Brennan, R. G., (1993) The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell* 74: 773-776.

Britton, R. A., D. C. Lin & A. D. Grossman, (1998) Characterization of a prokaryotic SMC protein involved in chromosome partitioning. *Genes Dev* 12: 1254-1259.

Broder, D. H. & K. Pogliano, (2006) Forespore engulfment mediated by a ratchet-like mechanism. *Cell* 126: 917-928.

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

Burkholder, W. F., I. Kurtser & A. D. Grossman, (2001) Replication initiation proteins regulate a developmental checkpoint in Bacillus subtilis. *Cell* 104: 269-279.

Bushman, W., S. Yin, L. L. Thio & A. Landy, (1984) Determinants of directionality in lambda site-specific recombination. *Cell* 39: 699-706.

Camp, A. H. & R. Losick, (2008) A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol* 69: 402-417.

Camp, A. H. & R. Losick, (2009) A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes Dev* 23: 1014-1024.

Campo, N. & D. Z. Rudner, (2007) SpoIVB and CtpB are both forespore signals in the activation of the sporulation transcription factor sigmaK in Bacillus subtilis. *J Bacteriol* 189: 6021-6027.

Carniol, K., S. Ben-Yehuda, N. King & R. Losick, (2005) Genetic dissection of the sporulation protein SpoIIE and its role in asymmetric division in Bacillus subtilis. *J Bacteriol* 187: 3511-3520.

Carniol, K., P. Eichenberger & R. Losick, (2004) A threshold mechanism governing activation of the developmental regulatory protein sigma F in Bacillus subtilis. *J Biol Chem* 279: 14860-14870.

Chai, Y., F. Chu, R. Kolter & R. Losick, (2008) Bistability and biofilm formation in Bacillus subtilis. *Mol Microbiol* 67: 254-263.

Clark, K. L., E. D. Halay, E. Lai & S. K. Burley, (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* 364: 412-420.

Crater, D. L. & C. P. Moran, Jr., (2001) Identification of a DNA binding region in GerE from Bacillus subtilis. *J Bacteriol* 183: 4183-4189.

Crater, D. L. & C. P. Moran, Jr., (2002) Two regions of GerE required for promoter activation in Bacillus subtilis. *J Bacteriol* 184: 241-249.

Crater, D. L., K. H. Wade, O. Resnekov, H. T. Ichikawa, L. Kroos, J. A. Brannigan & C. P. Moran, Jr., (2002) A mutation in GerE that affects cotC promoter activation in Bacillus subtilis. *Biochim Biophys Acta* 1576: 30-38.

Daniel, R. A., S. Drake, C. E. Buchanan, R. Scholle & J. Errington, (1994) The Bacillus subtilis spoVD gene encodes a mother-cell-specific penicillin-binding protein required for spore morphogenesis. *J Mol Biol* 235: 209-220.

de Lencastre, H. & P. J. Piggot, (1979) Identification of different sites of expression for spo loci by transformation of Bacillus subtilis. *J Gen Microbiol* 114: 377-389.

Decatur, A., M. T. McMurry, B. N. Kunkel & R. Losick, (1997) Translation of the mRNA for the sporulation gene spoIIID of Bacillus subtilis is dependent upon translation of a small upstream open reading frame. *J Bacteriol* 179: 1324-1328.

Decatur, A. L. & R. Losick, (1996) Three sites of contact between the Bacillus subtilis transcription factor sigmaF and its antisigma factor SpoIIAB. *Genes Dev* 10: 2348-2358.

Demeler, B. & K. E. van Holde, (2004) Sedimentation velocity analysis of highly heterogeneous systems. *Anal Biochem* 335: 279-288.

Doan, T., K. A. Marquis & D. Z. Rudner, (2005) Subcellular localization of a sporulation membrane protein is achieved through a network of interactions along and across the septum. *Mol Microbiol* 55: 1767-1781.

Doan, T., C. Morlot, J. Meisner, M. Serrano, A. O. Henriques, C. P. Moran, Jr. & D. Z. Rudner, (2009) Novel secretion apparatus maintains spore integrity and developmental gene expression in Bacillus subtilis. *PLoS Genet* 5: e1000566.

Dong, T. C. & S. M. Cutting, (2003) SpoIVB-mediated cleavage of SpoIVFA could provide the intercellular signal to activate processing of Pro-sigmaK in Bacillus subtilis. *Mol Microbiol* 49: 1425-1434.

Driks, A., (2002) Proteins of the Spore Coat. In: *Bacillus subtilis* and Its Closest Relatives: from Genes to Cells. J. A. H. Abraham L. Sonenshein, and R. Losick (ed). Washington, D.C.: ASM Press, pp. 527-535.

Driks, A., S. Roels, B. Beall, C. P. Moran, Jr. & R. Losick, (1994) Subcellular localization of proteins involved in the assembly of the spore coat of Bacillus subtilis. *Genes Dev* 8: 234-244.

Dubnau, D. & R. Losick, (2006) Bistability in bacteria. *Mol Microbiol* 61: 564-572.

Dubnau, D. & R. Provvedi, (2000) Internalizing DNA. *Res Microbiol* 151: 475-480.

Ducros, V. M., R. J. Lewis, C. S. Verma, E. J. Dodson, G. Leonard, J. P. Turkenburg, G. N. Murshudov, A. J. Wilkinson & J. A. Brannigan, (2001) Crystal structure of GerE, the ultimate transcriptional regulator of spore formation in Bacillus subtilis. *J Mol Biol* 306: 759-771.

Duncan, L., S. Alper, F. Arigoni, R. Losick & P. Stragier, (1995) Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. *Science* 270: 641-644.

Dworkin, J. & R. Losick, (2001) Differential gene expression governed by chromosomal spatial asymmetry. *Cell* 107: 339-346.

Edwards, D. H. & J. Errington, (1997) The Bacillus subtilis DivIVA protein targets to the division septum and controls the site specificity of cell division. *Mol Microbiol* 24: 905-915.

Eichenberger, P., P. Fawcett & R. Losick, (2001) A three-protein inhibitor of polar septation during sporulation in Bacillus subtilis. *Mol Microbiol* 42: 1147-1162.

Eichenberger, P., M. Fujita, S. T. Jensen, E. M. Conlon, D. Z. Rudner, S. T. Wang, C. Ferguson, K. Haga, T. Sato, J. S. Liu & R. Losick, (2004) The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. *PLoS Biol* 2: e328.

Eichenberger, P., S. T. Jensen, E. M. Conlon, C. van Ooij, J. Silvaggi, J. E. Gonzalez-Pastor, M. Fujita, S. Ben-Yehuda, P. Stragier, J. S. Liu & R. Losick, (2003) The sigmaE regulon and the identification of additional sporulation genes in Bacillus subtilis. *J Mol Biol* 327: 945-972.

Errington, J., L. Appleby, R. A. Daniel, H. Goodfellow, S. R. Partridge & M. D. Yudkin, (1992) Structure and function of the spoIIIJ gene of Bacillus subtilis: a vegetatively expressed gene that is essential for sigma G activity at an intermediate stage of sporulation. *J Gen Microbiol* 138: 2609-2618.

Foster, S. J. & D. L. Popham, (2002) Structure and Synthesis of Cell Wall, Spore Cortex, Teichoic Acids, S-Layers, and Capsules. In: *Bacillus subtilis* and Its Closest Relatives: from Genes to Cells. J. A. H. Abraham L. Sonenshein, and R. Losick (ed). Washington, D.C.: ASM Press, pp. 21-41.

Foulger, D. & J. Errington, (1991) Sequential activation of dual promoters by different sigma factors maintains spoVJ expression during successive developmental stages of Bacillus subtilis. *Mol Microbiol* 5: 1363-1373.

Frandsen, N., I. Barak, C. Karmazyn-Campelli & P. Stragier, (1999) Transient gene asymmetry during sporulation and establishment of cell specificity in Bacillus subtilis. *Genes Dev* 13: 394-399.

Fujita, M., J. E. Gonzalez-Pastor & R. Losick, (2005) High- and low-threshold genes in the Spo0A regulon of Bacillus subtilis. *J Bacteriol* 187: 1357-1368.

Fujita, M. & R. Losick, (2002) An investigation into the compartmentalization of the sporulation transcription factor sigmaE in Bacillus subtilis. *Mol Microbiol* 43: 27-38.

Fujita, M. & R. Losick, (2003) The master regulator for entry into sporulation in Bacillus subtilis becomes a cell-specific transcription factor after asymmetric division. *Genes Dev* 17: 1166-1174.

Gallegos, M. T., R. Schleif, A. Bairoch, K. Hofmann & J. L. Ramos, (1997) Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* 61: 393-410.

Gish, W. & D. J. States, (1993) Identification of protein coding regions by database similarity search. *Nat Genet* 3: 266-272.

Gomez, M., S. Cutting & P. Stragier, (1995) Transcription of spoIVB is the only role of sigma G that is essential for pro-sigma K processing during spore formation in Bacillus subtilis. *J Bacteriol* 177: 4825-4827.

Gonzalez-Pastor, J. E., E. C. Hobbs & R. Losick, (2003) Cannibalism by sporulating bacteria. *Science* 301: 510-513.

Gouet, P., E. Courcelle, D. I. Stuart & F. Metoz, (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15: 305-308.

Guerout-Fleury, A. M., N. Frandsen & P. Stragier, (1996) Plasmids for ectopic integration in Bacillus subtilis. *Gene* 180: 57-61.

Halberg, R. & L. Kroos, (1992) Fate of the SpoIIID switch protein during Bacillus subtilis sporulation depends on the mother-cell sigma factor, sigma K. *J Mol Biol* 228: 840-849.

Halberg, R. & L. Kroos, (1994) Sporulation regulatory protein SpoIIID from Bacillus subtilis activates and represses transcription by both mother-cell-specific forms of RNA polymerase. *J Mol Biol* 243: 425-436.

Halberg, R., V. Oke & L. Kroos, (1995) Effects of Bacillus subtilis sporulation regulatory protein SpoIIID on transcription by sigma K RNA polymerase in vivo and in vitro. *J Bacteriol* 177: 1888-1891.

Harry, K. H., R. Zhou, L. Kroos & S. B. Melville, (2009) Sporulation and enterotoxin (CPE) synthesis are controlled by the sporulation-specific sigma factors SigE and SigK in Clostridium perfringens. *J Bacteriol* 191: 2728-2742.

Harwood, C. R. & S. M. Cutting, (1990) *Molecular Biological Methods for Bacillus*. John Wiley & Sons Ltd, Chichester, England.

Healy, J., J. Weir, I. Smith & R. Losick, (1991) Post-transcriptional control of a sporulation regulatory gene encoding transcription factor sigma H in Bacillus subtilis. *Mol Microbiol* 5: 477-487.

Henriques, A. O., H. de Lencastre & P. J. Piggot, (1992) A Bacillus subtilis morphogene cluster that includes spoVE is homologous to the mra region of Escherichia coli. *Biochimie* 74: 735-748.

Henriques, A. O. & C. P. Moran, Jr., (2007) Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* 61: 555-588.

Hilbert, D. W. & P. J. Piggot, (2004) Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiol Mol Biol Rev* 68: 234-262.

Hitchins, A. D., (1975) Polarized relationship of bacterial spore loci to the "old" and "new" ends of sporangia. *J Bacteriol* 121: 518-523.

Hofmeister, A., (1998) Activation of the proprotein transcription factor pro-sigmaE is associated with its progression through three patterns of subcellular localization during sporulation in Bacillus subtilis. *J Bacteriol* 180: 2426-2433.

Hofmeister, A. E., A. Londono-Vallejo, E. Harry, P. Stragier & R. Losick, (1995) Extracellular signal protein triggering the proteolytic activation of a developmental transcription factor in B. subtilis. *Cell* 83: 219-226.

Hosoya, S., Z. Lu, Y. Ozaki, M. Takeuchi & T. Sato, (2007) Cytological analysis of the mother cell death process during sporulation in Bacillus subtilis. *J Bacteriol* 189: 2561-2565.

Huffman, J. L. & R. G. Brennan, (2002) Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol* 12: 98-106.

Iber, D., J. Clarkson, M. D. Yudkin & I. D. Campbell, (2006) The mechanism of cell differentiation in Bacillus subtilis. *Nature* 441: 371-374.

Ichikawa, H., R. Halberg & L. Kroos, (1999) Negative regulation by the Bacillus subtilis GerE protein. *J Biol Chem* 274: 8322-8327.

Ichikawa, H. & L. Kroos, (2000) Combined action of two transcription factors regulates genes encoding spore coat proteins of Bacillus subtilis. *J Biol Chem* 275: 13849-13855.

Illing, N. & J. Errington, (1991) The spoIIIA operon of Bacillus subtilis defines a new temporal class of mother-cell-specific sporulation genes under the control of the sigma E form of RNA polymerase. *Mol Microbiol* 5: 1927-1940.

Imamura, D., R. Zhou, M. Feig & L. Kroos, (2008) Evidence that the Bacillus subtilis SpoIIGA protein is a novel type of signal-transducing aspartic protease. *J Biol Chem* 283: 15287-15299.

Ionesco, H., J. Michel, B. Cami & P. Schaeffer, (1970) Symposium on bacterial spores: II. Genetics of sporulation in Bacillus subtilis Marburg. *J Appl Bacteriol* 33: 13-24.

Ishikawa, S., K. Yamane & J. Sekiguchi, (1998) Regulation and characterization of a newly deduced cell wall hydrolase gene (cwlJ) which affects germination of Bacillus subtilis spores. *J Bacteriol* 180: 1375-1380.

Jiang, M., R. Grau & M. Perego, (2000a) Differential processing of propeptide inhibitors of Rap phosphatases in Bacillus subtilis. *J Bacteriol* 182: 303-310.

Jiang, M., W. Shao, M. Perego & J. A. Hoch, (2000b) Multiple histidine kinases regulate entry into stationary phase and sporulation in Bacillus subtilis. *Mol Microbiol* 38: 535-542.

Jiang, X., A. Rubio, S. Chiba & K. Pogliano, (2005) Engulfment-regulated proteolysis of SpoIIQ: evidence that dual checkpoints control sigma activity. *Mol Microbiol* 58: 102-115.

Johnsen, P. J., D. Dubnau & B. R. Levin, (2009) Episodic selection and the maintenance of competence and natural transformation in Bacillus subtilis. *Genetics* 181: 1521-1533.

Jonas, R. M., E. A. Weaver, T. J. Kenney, C. P. Moran, Jr. & W. G. Haldenwang, (1988) The Bacillus subtilis spoIIG operon encodes both sigma E and a gene necessary for sigma E activation. *J Bacteriol* 170: 507-511.

Juan Wu, L. & J. Errington, (2000) Identification and characterization of a new prespore-specific regulatory gene, rsfA, of Bacillus subtilis. *J Bacteriol* 182: 418-424.

Karmazyn-Campelli, C., L. Rhayat, R. Carballido-Lopez, S. Duperrier, N. Frandsen & P. Stragier, (2008) How the early sporulation sigma factor sigmaF delays the switch to late development in Bacillus subtilis. *Mol Microbiol* 67: 1169-1180.

Karow, M. L., P. Glaser & P. J. Piggot, (1995) Identification of a gene, spoIIR, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in Bacillus subtilis. *Proc Natl Acad Sci U S A* 92: 2012-2016.

Kearns, D. B. & R. Losick, (2005) Cell population heterogeneity during growth of Bacillus subtilis. *Genes Dev* 19: 3083-3094.

Kim, H., M. Hahn, P. Grabowski, D. C. McPherson, M. M. Otte, R. Wang, C. C. Ferguson, P. Eichenberger & A. Driks, (2006) The Bacillus subtilis spore coat protein interaction network. *Mol Microbiol* 59: 487-502.

King, N., O. Dreesen, P. Stragier, K. Pogliano & R. Losick, (1999) Septation, dephosphorylation, and the activation of sigmaF during sporulation in Bacillus subtilis. *Genes Dev* 13: 1156-1167.

Kirchman, P. A., H. DeGrazia, E. M. Kellner & C. P. Moran, Jr., (1993) Forespore-specific disappearance of the sigma-factor antagonist spoIIAB: implications for its role in determination of cell fate in Bacillus subtilis. *Mol Microbiol* 8: 663-671.

Kobayashi, K., K. Hashiguchi, K. Yokozeki & S. Yamanaka, (1998) Molecular cloning of the transglutaminase gene from Bacillus subtilis and its expression in Escherichia coli. *Biosci Biotechnol Biochem* 62: 1109-1114.

Kroos, L., (2007) The Bacillus and Myxococcus developmental networks and their transcriptional regulators. *Annu Rev Genet* 41: 13-39.

Kroos, L., B. Kunkel & R. Losick, (1989) Switch protein alters specificity of RNA polymerase containing a compartment-specific sigma factor. *Science* 243: 526-529.

Kroos, L., Y. T. Yu, D. Mills & S. Ferguson-Miller, (2002) Forespore signaling is necessary for pro-sigmaK processing during Bacillus subtilis sporulation despite the loss of SpoIVFA upon translational arrest. *J Bacteriol* 184: 5393-5401.

Kunkel, B., L. Kroos, H. Poth, P. Youngman & R. Losick, (1989) Temporal and spatial control of the mother-cell regulatory gene spoIIID of Bacillus subtilis. *Genes Dev* 3: 1735-1744.

Kunkel, B., R. Losick & P. Stragier, (1990) The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* 4: 525-535.

Kunkel, B., K. Sandman, S. Panzer, P. Youngman & R. Losick, (1988) The promoter for a sporulation gene in the spoIVC locus of Bacillus subtilis and its use in studies of temporal and spatial control of gene expression. *J Bacteriol* 170: 3513-3522.

Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin & et al., (1997) The complete genome sequence of the gram-positive bacterium Bacillus subtilis. *Nature* 390: 249-256.

Kuroda, A. & J. Sekiguchi, (1991) Molecular cloning and sequencing of a major Bacillus subtilis autolysin gene. *J Bacteriol* 173: 7304-7312.

LaBell, T. L., J. E. Trempy & W. G. Haldenwang, (1987) Sporulation-specific sigma factor sigma 29 of Bacillus subtilis is synthesized from a precursor protein, P31. *Proc Natl Acad Sci U S A* 84: 1784-1788.

Lai, E. M., N. D. Phadke, M. T. Kachman, R. Giorno, S. Vazquez, J. A. Vazquez, J. R. Maddock & A. Driks, (2003) Proteomic analysis of the spore coats of Bacillus subtilis and Bacillus anthracis. *J Bacteriol* 185: 1443-1454.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson & D. G. Higgins, (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

LeDeaux, J. R. & A. D. Grossman, (1995) Isolation and characterization of kinC, a gene that encodes a sensor kinase homologous to the sporulation sensor kinases KinA and KinB in Bacillus subtilis. *J Bacteriol* 177: 166-175.

Levin, P. A., N. Fan, E. Ricca, A. Driks, R. Losick & S. Cutting, (1993) An unusually small gene required for sporulation by Bacillus subtilis. *Mol Microbiol* 9: 761-771.

Lewis, P. J., T. Magnin & J. Errington, (1996) Compartmentalized distribution of the proteins controlling the prespore-specific transcription factor sigmaF of Bacillus subtilis. *Genes Cells* 1: 881-894.

Lindow, J. C., R. A. Britton & A. D. Grossman, (2002) Structural maintenance of chromosomes protein of Bacillus subtilis affects supercoiling in vivo. *J Bacteriol* 184: 5317-5322.

Little, S. & A. Driks, (2001) Functional analysis of the Bacillus subtilis morphogenetic spore coat protein CotE. *Mol Microbiol* 42: 1107-1120.

Lopez, D., M. A. Fischbach, F. Chu, R. Losick & R. Kolter, (2009) Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis. *Proc Natl Acad Sci U S A* 106: 280-285.

Losick, R. & P. Stragier, (1992) Crisscross regulation of cell-type-specific gene expression during development in B. subtilis. *Nature* 355: 601-604.

Magge, A., A. C. Granger, P. G. Wahome, B. Setlow, V. R. Vepachedu, C. A. Loshon, L. Peng, D. Chen, Y. Q. Li & P. Setlow, (2008) Role of dipicolinic acid in the germination, stability, and viability of spores of Bacillus subtilis. *J Bacteriol* 190: 4798-4807.

Meisner, J., X. Wang, M. Serrano, A. O. Henriques & C. P. Moran, Jr., (2008) A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc Natl Acad Sci U S A* 105: 15100-15105.

Miller, J. H., (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Min, K. T., C. M. Hilditch, B. Diederich, J. Errington & M. D. Yudkin, (1993) Sigma F, the first compartment-specific transcription factor of B. subtilis, is regulated by an anti-sigma factor that is also a protein kinase. *Cell* 74: 735-742.

Molle, V., M. Fujita, S. T. Jensen, P. Eichenberger, J. E. Gonzalez-Pastor, J. S. Liu & R. Losick, (2003a) The Spo0A regulon of Bacillus subtilis. *Mol Microbiol* 50: 1683-1701.

Molle, V., Y. Nakaura, R. P. Shivers, H. Yamaguchi, R. Losick, Y. Fujita & A. L. Sonenshein, (2003b) Additional targets of the Bacillus subtilis global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J Bacteriol* 185: 1911-1922.

Murray, H. & J. Errington, (2008) Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* 135: 74-84.

Nugroho, F. A., H. Yamamoto, Y. Kobayashi & J. Sekiguchi, (1999) Characterization of a new sigma-K-dependent peptidoglycan hydrolase gene that plays a role in Bacillus subtilis mother cell lysis. *J Bacteriol* 181: 6230-6237.

Ozin, A. J., C. S. Samford, A. O. Henriques & C. P. Moran, Jr., (2001) SpoVID guides SafA to the spore coat in Bacillus subtilis. *J Bacteriol* 183: 3041-3049.

Pabo, C. O. & R. T. Sauer, (1992) Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* 61: 1053-1095.

Paidhungat, M., K. Ragkousi & P. Setlow, (2001) Genetic requirements for induction of germination of spores of Bacillus subtilis by Ca(2+)-dipicolinate. *J Bacteriol* 183: 4886-4893.

Pan, Q., D. A. Garsin & R. Losick, (2001) Self-reinforcing activation of a cell-specific transcription factor by proteolysis of an anti-sigma factor in B. subtilis. *Mol Cell* 8: 873-883.

Partridge, S. R. & J. Errington, (1993) The importance of morphological events and intercellular interactions in the regulation of prespore-specific gene expression during sporulation in Bacillus subtilis. *Mol Microbiol* 8: 945-955.

Perego, M., (1997) A peptide export-import control circuit modulating bacterial development regulates protein phosphatases of the phosphorelay. *Proc Natl Acad Sci U S A* 94: 8612-8617.

Piggot, P. J. & J. G. Coote, (1976) Genetic aspects of bacterial endospore formation. *Bacteriol Rev* 40: 908-962.

Piggot, P. J. & D. W. Hilbert, (2004) Sporulation of Bacillus subtilis. *Curr Opin Microbiol* 7: 579-586.

Popham, D. L., M. E. Gilmore & P. Setlow, (1999) Roles of low-molecular-weight penicillin-binding proteins in Bacillus subtilis spore peptidoglycan synthesis and spore properties. *J Bacteriol* 181: 126-132.

Price, K. D. & R. Losick, (1999) A four-dimensional view of assembly of a morphogenetic protein during sporulation in Bacillus subtilis. *J Bacteriol* 181: 781-790.

Ragkousi, K., P. Eichenberger, C. van Ooij & P. Setlow, (2003) Identification of a new gene essential for germination of Bacillus subtilis spores with Ca2+-dipicolinate. *J Bacteriol* 185: 2315-2329.

Ragkousi, K. & P. Setlow, (2004) Transglutaminase-mediated cross-linking of GerQ in the coats of Bacillus subtilis spores. *J Bacteriol* 186: 5567-5575.

Ramamurthi, K. S., S. Lecuyer, H. A. Stone & R. Losick, (2009) Geometric cue for protein localization in a bacterium. *Science* 323: 1354-1357.

Ratnayake-Lecamwasam, M., P. Serror, K. W. Wong & A. L. Sonenshein, (2001) Bacillus subtilis CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev* 15: 1093-1103.

Rouyard, J. F., H. Ionesco & P. Schaeffer, (1967) [Genetic classification of certain sporulation mutants of Bacillus subtilis, Marburg]. *Ann Inst Pasteur (Paris)* 113: 675-683.

Rudner, D. Z., P. Fawcett & R. Losick, (1999) A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proc Natl Acad Sci U S A* 96: 14765-14770.

Rudner, D. Z. & R. Losick, (2002) A sporulation membrane protein tethers the pro-sigmaK processing enzyme to its inhibitor and dictates its subcellular localization. *Genes Dev* 16: 1007-1018.

Ryter, A., (1965) [Morphologic Study of the Sporulation of Bacillus Subtilis.]. *Ann Inst Pasteur (Paris)* 108: 40-60.

Sacco, M., E. Ricca, R. Losick & S. Cutting, (1995) An additional GerE-controlled gene encoding an abundant spore coat protein from Bacillus subtilis. *J Bacteriol* 177: 372-377.

Sam, M. D., D. Cascio, R. C. Johnson & R. T. Clubb, (2004) Crystal structure of the excisionase-DNA complex from bacteriophage lambda. *J Mol Biol* 338: 229-240.

Sambrook, J., E. F. Fritsch & T. Maniatis, (1989) *Molecular Cloning: A laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sato, T., K. Harada, Y. Ohta & Y. Kobayashi, (1994) Expression of the Bacillus subtilis spoIVCA gene, which encodes a site-specific recombinase, depends on the spoIIGB product. *J Bacteriol* 176: 935-937.

Schmidt, R., P. Margolis, L. Duncan, R. Coppolecchia, C. P. Moran, Jr. & R. Losick, (1990) Control of developmental transcription factor sigma F by sporulation regulatory proteins SpoIIAA and SpoIIAB in Bacillus subtilis. *Proc Natl Acad Sci U S A* 87: 9221-9225.

Schuck, P. & P. Rossmanith, (2000) Determination of the sedimentation coefficient distribution by least-squares boundary modeling. *Biopolymers* 54: 328-341.

Schumacher, M. A., K. Y. Choi, H. Zalkin & R. G. Brennan, (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* 266: 763-770.

Serrano, M., L. Corte, J. Opdyke, C. P. Moran, Jr. & A. O. Henriques, (2003) Expression of spoIIIJ in the prespore is sufficient for activation of sigma G and for sporulation in Bacillus subtilis. *J Bacteriol* 185: 3905-3917.

Serrano, M., A. Neves, C. M. Soares, C. P. Moran, Jr. & A. O. Henriques, (2004) Role of the anti-sigma factor SpoIIAB in regulation of sigmaG during Bacillus subtilis sporulation. *J Bacteriol* 186: 4000-4013.

Serrano, M., F. Vieira, C. P. Moran, Jr. & A. O. Henriques, (2008) Processing of a membrane protein required for cell-to-cell signaling during endospore formation in Bacillus subtilis. *J Bacteriol* 190: 7786-7796.

Setlow, B., P. G. Wahome & P. Setlow, (2008) Release of small molecules during germination of spores of Bacillus Species. *J Bacteriol* 190: 4759-4763.

Setlow, P., (2003) Spore germination. *Curr Opin Microbiol* 6: 550-556.

Sharp, M. D. & K. Pogliano, (1999) An in vivo membrane fusion assay implicates SpoIIIE in the final stages of engulfment during Bacillus subtilis sporulation. *Proc Natl Acad Sci U S A* 96: 14553-14558.

Sharp, M. D. & K. Pogliano, (2002) Role of cell-specific SpoIIIE assembly in polarity of DNA transfer. *Science* 295: 137-139.

Silvaggi, J. M., J. B. Perkins & R. Losick, (2006) Genes for small, noncoding RNAs under sporulation control in Bacillus subtilis. *J Bacteriol* 188: 532-541.

Smith, T. J. & S. J. Foster, (1995) Characterization of the involvement of two compensatory autolysins in mother cell lysis during sporulation of Bacillus subtilis 168. *J Bacteriol* 177: 3855-3862.

Steil, L., M. Serrano, A. O. Henriques & U. Volker, (2005) Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of Bacillus subtilis. *Microbiology* 151: 399-420.

Stevens, C. M. & J. Errington, (1990) Differential gene expression during sporulation in Bacillus subtilis: structure and regulation of the spoIIID gene. *Mol Microbiol* 4: 543-551.

Stragier, P., C. Bonamy & C. Karmazyn-Campelli, (1988) Processing of a sporulation sigma factor in Bacillus subtilis: how morphological structure could control gene expression. *Cell* 52: 697-704.

Stragier, P., B. Kunkel, L. Kroos & R. Losick, (1989) Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science* 243: 507-512.

Sun, D. X., R. M. Cabrera-Martinez & P. Setlow, (1991) Control of transcription of the Bacillus subtilis spoIIIG gene, which codes for the forespore-specific transcription factor sigma G. *J Bacteriol* 173: 2977-2984.

Sun, Y. L., M. D. Sharp & K. Pogliano, (2000) A dispensable role for forespore-specific gene expression in engulfment of the forespore during sporulation of Bacillus subtilis. *J Bacteriol* 182: 2919-2927.

Suzuki, S., Y. Izawa, K. Kobayashi, Y. Eto, S. Yamanaka, K. Kubota & K. Yokozeki, (2000) Purification and characterization of novel transglutaminase from Bacillus subtilis spores. *Biosci Biotechnol Biochem* 64: 2344-2351.

Tatti, K. M., M. F. Shuler & C. P. Moran, Jr., (1995) Sequence-specific interactions between promoter DNA and the RNA polymerase sigma factor E. *J Mol Biol* 253: 8-16.

Tovar-Rojo, F., M. Chander, B. Setlow & P. Setlow, (2002) The products of the spoVA operon are involved in dipicolinic acid uptake into developing spores of Bacillus subtilis. *J Bacteriol* 184: 584-587.

Trach, K. A. & J. A. Hoch, (1993) Multisensory activation of the phosphorelay initiating sporulation in Bacillus subtilis: identification and sequence of the protein kinase of the alternate pathway. *Mol Microbiol* 8: 69-79.

Tullius, T., (1995) Homeodomains: together again for the first time. *Structure* 3: 1143-1145.

van Ooij, C. & R. Losick, (2003) Subcellular localization of a small sporulation protein in Bacillus subtilis. *J Bacteriol* 185: 1391-1398.

Vlamakis, H., C. Aguilar, R. Losick & R. Kolter, (2008) Control of cell fate by the formation of an architecturally complex bacterial community. *Genes Dev* 22: 945-953.

Wade, K. H., G. Schyns, J. A. Opdyke & C. P. Moran, Jr., (1999) A region of sigmaK involved in promoter activation by GerE in Bacillus subtilis. *J Bacteriol* 181: 4365-4373.

Wakeley, P., N. T. Hoa & S. Cutting, (2000a) BofC negatively regulates SpoIVB-mediated signalling in the Bacillus subtilis sigmaK-checkpoint. *Mol Microbiol* 36: 1415-1424.

Wakeley, P. R., R. Dorazi, N. T. Hoa, J. R. Bowyer & S. M. Cutting, (2000b) Proteolysis of SpolVB is a critical determinant in signalling of Pro-sigmaK processing in Bacillus subtilis. *Mol Microbiol* 36: 1336-1348.

Wang, L., J. Perpich, A. Driks & L. Kroos, (2007a) Maintaining the transcription factor SpoIIID level late during sporulation causes spore defects in Bacillus subtilis. *J Bacteriol* 189: 7302-7309.

Wang, L., J. Perpich, A. Driks & L. Kroos, (2007b) One perturbation of the mother cell gene regulatory network suppresses the effects of another during sporulation of Bacillus subtilis. *J Bacteriol* 189: 8467-8473.

Wang, S. T., B. Setlow, E. M. Conlon, J. L. Lyon, D. Imamura, T. Sato, P. Setlow, R. Losick & P. Eichenberger, (2006) The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* 358: 16-37.

Wojciak, J. M., J. Iwahara & R. T. Clubb, (2001) The Mu repressor-DNA complex contains an immobilized 'wing' within the minor groove. *Nat Struct Biol* 8: 84-90.

Wu, L. J. & J. Errington, (1998) Use of asymmetric cell division and spoIIIE mutants to probe chromosome orientation and organization in Bacillus subtilis. *Mol Microbiol* 27: 777-786.

Wu, L. J. & J. Errington, (2003) RacA and the Soj-Spo0J system combine to effect polar chromosome segregation in sporulating Bacillus subtilis. *Mol Microbiol* 49: 1463-1475.

Youngman, P., J. B. Perkins & R. Losick, (1984) Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in Bacillus subtilis or expression of the transposon-borne erm gene. *Plasmid* 12: 1-9.

Yu, Y. T. & L. Kroos, (2000) Evidence that SpoIVFB is a novel type of membrane metalloprotease governing intercompartmental communication during Bacillus subtilis sporulation. *J Bacteriol* 182: 3305-3309.

Zhang, B., R. A. Daniel, J. Errington & L. Kroos, (1997) Bacillus subtilis SpoIIID protein binds to two sites in the spoVD promoter and represses transcription by sigmaE RNA polymerase. *J Bacteriol* 179: 972-975.

Zhang, B., A. Hofmeister & L. Kroos, (1998) The prosequence of pro-sigmaK promotes membrane association and inhibits RNA polymerase core binding. *J Bacteriol* 180: 2434-2441.

Zhang, B. & L. Kroos, (1997) A feedback loop regulates the switch from one sigma factor to the next in the cascade controlling Bacillus subtilis mother cell gene expression. *J Bacteriol* 179: 6138-6144.

Zhang, B., P. Struffi & L. Kroos, (1999) sigmaK can negatively regulate sigE expression by two different mechanisms during sporulation of Bacillus subtilis. *J Bacteriol* 181: 4081-4088.

Zhang, J., P. C. Fitz-James & A. I. Aronson, (1993) Cloning and characterization of a cluster of genes encoding polypeptides present in the insoluble fraction of the spore coat of Bacillus subtilis. *J Bacteriol* 175: 3757-3766.

Zhang, J., H. Ichikawa, R. Halberg, L. Kroos & A. I. Aronson, (1994) Regulation of the transcription of a cluster of Bacillus subtilis spore coat genes. *J Mol Biol* 240: 405-415.

Zheng, L., R. Halberg, S. Roels, H. Ichikawa, L. Kroos & R. Losick, (1992) Sporulation regulatory protein GerE from Bacillus subtilis binds to and can activate or repress transcription from promoters for mother-cell-specific genes. *J Mol Biol* 226: 1037-1050.

Zheng, L. B., W. P. Donovan, P. C. Fitz-James & R. Losick, (1988) Gene encoding a morphogenic protein required in the assembly of the outer coat of the Bacillus subtilis endospore. *Genes Dev* 2: 1047-1054.

Zheng, L. B. & R. Losick, (1990) Cascade regulation of spore coat gene expression in Bacillus subtilis. *J Mol Biol* 212: 645-660.

Zhou, R. & L. Kroos, (2004) BofA protein inhibits intramembrane proteolysis of pro-sigmaK in an intercompartmental signaling pathway during Bacillus subtilis sporulation. *Proc Natl Acad Sci U S A* 101: 6385-6390.

Zhou, R. & L. Kroos, (2005) Serine proteases from two cell types target different components of a complex that governs regulated intramembrane proteolysis of pro-sigmaK during Bacillus subtilis development. *Mol Microbiol* 58: 835-846.

Zilhao, R., R. Isticato, L. O. Martins, L. Steil, U. Volker, E. Ricca, C. P. Moran, Jr. & A. O. Henriques, (2005) Assembly and function of a spore coat-associated transglutaminase of Bacillus subtilis. *J Bacteriol* 187: 7753-7764.