This is to certify that the

dissertation entitled

Characterization of the regulatory regions of cell-interaction-dependent genes in Myxococcus xanthus

presented by

Makda Fisseha

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Microbiology

Lee Kroos
Major professor

Date 11/16/95

# CHARACTERIZATION OF THE REGULATORY REGIONS OF CELL INTERACTION-DEPENDENT GENES IN *MYXOCOCCUS XANTHUS*

By

Makda Fisseha

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Microbiology

1995

ABSTRACT

CHARACTERIZATION OF THE REGULATORY REGIONS OF CELL INTERACTION-DEPENDENT GENES IN *MYXOCOCCUS XANTHUS*

By

Makda Fisseha

*Myxococcus xanthus* is a gram-negative, rod-shaped, soil bacterium that undergoes multicellular development when starved for nutrients. Development involves the coordinate movement of cells into aggregation centers, formation of fruiting bodies, and sporulation of cells within fruiting bodies. The coordination of this multicellular fruiting body formation and sporulation process requires at least five different cell-cell signaling interactions. These intercellular signaling interactions are also required for the proper expression of a temporally regulated set of genes.

Extracellular C-signaling is required for development starting at six hours after the initiation of the developmental process. Cells that are defective in C-signaling initiate development normally but fail to form fruiting bodies and spores. Expression of genes that normally begin after six hours of development is either reduced or abolished in C-signaling

defective cells indicating that C-signaling is required for gene expression after six hours of development.

To understand the mechanism by which extracellular C-signaling regulates gene expression during development, the upstream DNA regulatory elements of two fusions between developmentally regulated *M. xanthus* promoters and a promoterless *lacZ* gene were characterized.

The Tn*5* *lac* Ω4403 fusion, whose expression has an absolute requirement for C-signaling, is expressed from a promoter bearing some resemblance to *E. coli* promoters transcribed by $\sigma^{70}$-RNA polymerase. DNA between 80 and 72 nucleotides upstream of the transcriptional initiation site is required for the expression of this fusion in a developmentally regulated manner and DNA downstream of -80 is sufficient for C-signal-dependent activation of the promoter, indicating that proteins in addition to RNA polymerase may be required for C-signal-dependent expression of Ω4403.

Tn*5* *lac* Ω4499 is a developmentally regulated fusion whose expression only partially requires extracellular C-signaling. C-signal dependent expression of this fusion requires 3.2 kb of DNA upstream of the fusion point. Tn*5* *lac* Ω4499 appears to be inserted into an operon that encodes enzymes involved in secondary metabolism.

**TO**

Almaz Zewde for her love and for teaching me that hard work and at principles should not be compromised.

Haeran and Yonas Fisseha for their love, friendship, their high standards which have been a source of inspiration, and Yonas for coming to my rescue in my times of computer crises.

## ACKNOWLEDGEMENTS

Many people helped me in my endeavors in graduate school and my gratitude and appreciation goes out to all of them. I have to especially thank Lee Kroos with whom I have been honored to work. I thank him for teaching my the art of science and for all his patience, kindness and understanding. I would also like to thank Mary, Mark, David, and Heidi Kroos for their friendship throughout the years.

I would like to thank my father, Fisseha Haile-Meskal, whose work and interest in science have always been an inspiration. Richard Halberg was both a colleague and friend whose humor, advice, and camaraderie I will always remember and value. Sijie Lu shared with me not only his cloning expertise but also his enthusiasm for science, which made the late hours in lab fun. I would like to thank both him and Mimi Lu for their friendship and kindness. I also thank Jamie White, the pioneer Myxo worker at Michigan State University, for his sincerity and for his offer of a (risky) canoe excursion. I have shared the joys and travails of working with Myxo with Janine Brandner who has given me many invaluable advice in my experiments and in my writing, and for which I am grateful. I thank Tong Hao for her friendship, for her interest in my project and for continuing with the mobility shift experiment. Dvora Biran has shown me that trying new experiments can be not only challenging but also exciting. I thank her for this and for all her encouragement. I also thanks Jennifer Loconto who has made the lab livlier, Hiroshi Ichikawa for being a kind and conscientious lab mate, which made working in the lab very pleasant, Bin Zhang for his candor and for allowing me to laugh at his absent mindedness,

and Michelle Anderson for all her helpful suggestions.

Among my family and friends I would like to especially thank Eleni Gabre-Medhin for showing me that where there is a will there is a way. I also thank Rawl, Saleela and Nisha Hollingsworth for their warm friendship and for welcoming me into their family, Rena, Ron and Arden Azarbarzin for all the good times I have had with them. Thanks also to Lilu Tesfa, Roya Mavaddat and Ghislaine Lydon for their special friendship.

## TABLE OF CONTENTS

**Chapter 3. Characterization of the DNA upstream of a partially C-signal-dependent operon**

## LIST OF TABLES

## LIST OF FIGURES

**Appendices**

## LIST OF ABBREVIATIONS

| | |
|---|---|
| Ap | ampicillin |
| bp | base pair |
| C | cytidine |
| $CaCl_2$ | calcium chloride |
| CTP | cytosine-5'-triphosphate |
| dCTP | deoxycytosine-5'-triphosphate |
| DEPC | diethyl pyrocarbonate |
| dGTP | deoxyguanosine-5'-triphosphate |
| DNA | deoxyribonucleic acid |
| DTT | dithiothreitol |
| EDTA | (ethylenedinitriol)tetraacetic acid |
| G | guanosine |
| kb | kilobase pairs |
| kDa | kilodalton |
| Km | kanamycin |
| LB | Luria-Bertani |
| µg | microgram |
| µl | microliter |
| µM | micromolar |
| mM | millimolar |
| M | molar |

| | |
|---|---|
| nM | nanomolar |
| PCR | polymerase chain reaction |
| RNA | ribonucleic acid |
| SDS | sodium dodecylsulfate |
| Tc | tetracycline |
| TDT | terminal deoxynucleotide transferase |
| v | volume |

# INTRODUCTION

*Myxococcus xanthus* is a gram-negative, rod-shaped, soil bacterium that has a multicellular form of existence. Upon nutrient depletion, cells undergo a complex developmental process which involves the movement of cells into aggregation centers and the formation of fruiting bodies. Within fruiting bodies, some of the cells lyse, whereas others mature into ovoid spores known as myxospores.

Both the regulation of morphogenesis and the transcriptional regulation of gene expression during development require intercellular signaling. To begin to understand how one extracellular signaling event, known as C-signaling, controls gene expression, the DNA regulatory regions of two developmental genes whose expression is regulated by C-signaling was characterized.

In chapter one, a review of cell-cell interactions in *M. xanthus* is presented and is compared with cell-cell interactions in other systems.

Chapter two describes the characterization of the upstream DNA regulatory region of Ω4403, a fusion of Tn*5* *lac* with a developmentally regulated gene whose expression has an absolute requirement for C-signaling. This fusion appears to be with a gene encoding a protease, and its expression requires DNA more than 72 base pairs upstream of the transcriptional initiation site. This suggests that activation proteins may be involved in the expression of Ω4403. Studies showing that the position in the chromosome affects expression of this fusion are also presented in this chapter. The contents of this chapter

will be submitted to the Journal of Bacteriology with Monica Gloudemans (a former graduate student in our lab) and Ronald E. Gill (University of Colorado) as second and third authors, respectively. Monica Gloudemans was responsible for cloning the 8.5 kb of DNA upstream of Ω4403 into pREG1666 and pREG1175 and testing it for promoter activity. Monica also generated the 4.0 and 2.0 kb deletions and tested them for promoter activity, and performed the S1-nuclease analyses. Ron Gill constructed pREG1666 and pREG1727, which I used extensively, and which have not been described adequately in the literature.

Chapter three describes the characterization of the DNA regulatory region of a second Tn*5* *lac* fusion, Ω4499, which has only a partial requirement for extracellular C-signaling. This fusion appears to be with a gene encoding a cytochrome P-450 enzyme, based on deduced amino acid sequence similarities. This gene appears to be part of an operon whose transcription requires DNA more than 2.1 kb upstream of the Ω4499 insertion site. The operon appears to encode at least one other enzyme, an NADH-dependent reductase.

The final chapter contains the conclusions and significance of these results. The regulatory region of the absolutely C-signaling dependent gene requires sequences upstream of what is typically recognized by RNA polymerase, and the same may hold true for the regulatory region of the operon whose expression is partially C-signal dependent operon. These studies provide the basis for identifying regulatory proteins that mediate the C-signal-dependent response.

# Chapter 1

# Literature Review

Differentiation and development are important processes that have been widely studied in a variety of multicellular organisms, rangeing from the simple eukaryote *Dictyostelium discoideum* to *Caenorhabditis elegans*, *Drosophila melanogaster*, and embryos. *Myxococcus xanthus* is a gram-negative, rod-shaped, soil bacterium that has a multicellular form of existence. The highly cooperative processes of multicellular feeding, swarming, and fruiting in *M. xanthus* resemble the developmental processes of more complex multicellular organisms in their requirement for cell migration and intercellular communication. Because *M. xanthus* is readily amenable to genetic and biochemical analyses, and because its developmental process is relatively short and simple, it provides a good model system for studying how cell-cell interactions control multicellular developmental and cellular differentiation.

**Social growth and motility.** During vegetative growth on a solid surface, *M. xanthus* cells glide as swarms toward nutrient sources and secrete bacteriolytic, cellulolytic, and other digestive enzymes to hydrolyze macromolecules (Kaiser & Kroos, 1993). Growth of *M. xanthus* is cell is density dependent, as indicated by experiments in which growth of cells on casein as the sole nutrient source increased proportionally as the cell density was raised in liquid culture (Rosenberg et al., 1977). This cell density dependent growth is attributed to the increased concentration of nutrients available for growth that results from the increased concentration of extracellular hydrolytic enzymes at high cell concentrations.

The multicellular nature of growth and development of *M. xanthus* is also evident in its motility. *M. xanthus* cells move in large groups known as swarms. Within the swarm, movement of cells in small groups is known as social (S) motility. Occasionally, individual cells briefly move outside of a swarm before returning, and this is known as adventurous (A) motility. These two types of motility are controlled by different systems that function largely independently. The S-motility system involves at least ten genes

(Shimkets, 1990). Cells defective in any of the S-motility genes fail to move in clusters but are able to move as individual cells (Shimkets, 1990). The A-motility system involves at least 23 gene products (Fink & Zissler, 1989; Hodgkin & Kaiser, 1977; Sodergren & Kaiser, 1983). Defects in any of the A-motility genes results in the failure of cells to move individually; however, these cells can still move in groups. A third locus that controls motility is the *mgl* locus. Mutations in the *mgl* locus render cells completely nonmotile. Thus, cells that are nonmotile both as individual cells and in swarms have mutations either in *mgl* or in both an A- and a S-motility gene (A-S-) (Hodgkin & Kaiser, 1977; Hodgkin & Kaiser, 1979b). The *mgl* locus consists of two genes, *mglA* and *mglB* (Stephens et al., 1988). The predicted amino acid sequence of *mglA* shows significant similarity to a class of GTP-binding proteins, and the region of similarity spans the putative GTP binding and hydrolysis sites (Hartzell & Kaiser, 1991). The product of the *mglB* gene appears to be important for the synthesis and/or stability of the MglA protein since mutations in *mglB* that do not affect *mglA* transcription have lower levels of the MglA protein. Since *lacZ* expression of transcriptional fusions of A- or S-motility genes with Tn*5* *lac* was not affected in *mgl* strains, it is unlikely that the *mgl* gene products are involved in regulating the transcription of A- or S- motility genes (MacNeil et al., 1994).

Motility of *M. xanthus* cells is dependent upon cell-cell interactions since increasing cell density also increases motility (Kaiser & Crosby, 1983). Further evidence that cell-cell interactions are essential for motility is that cells with defects in some motility genes can be transiently induced to be motile by placing them in contact with wild-type cells (Kaiser & Crosby, 1983).

**Social development.** When *Myxococcus* cells are starved at a sufficiently high density on a solid surface they initiate a developmental process. Early in development cells move synchronously, arranging themselves in a series of equidistant ridges separated by regions devoid of cells. At this stage, cells move in a pulsating manner and are said to be

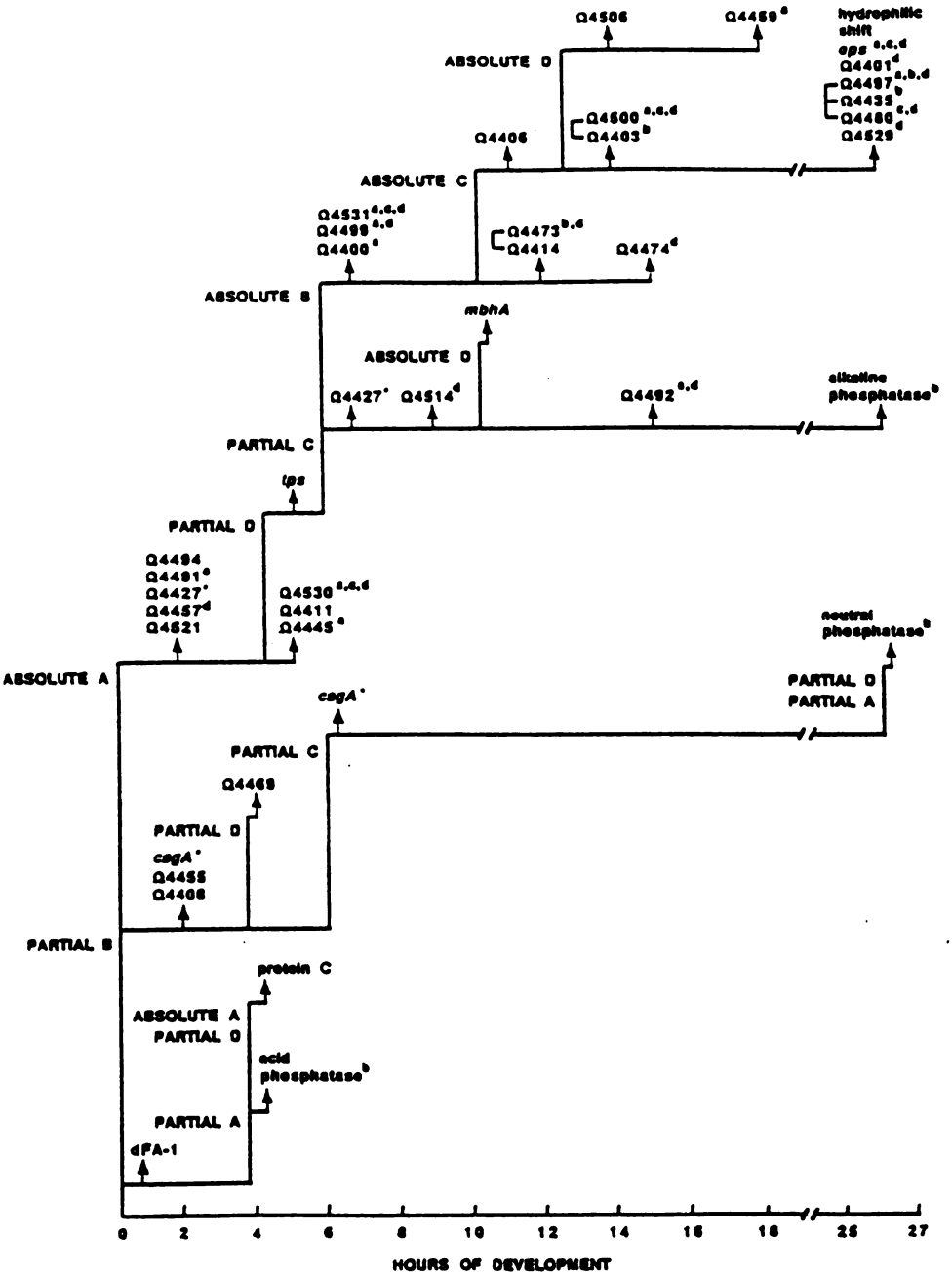
rippling since the ridges resemble ripples on the surface of water (Shimkets, 1990). Sager and Kaiser (Sager & Kaiser, 1993a) have proposed that there is no net movement of cells during rippling. They suggest that the appearance of propagating waves is due to the back-and-forth movement of cells as they reverse the direction of their gliding upon end-to-end contact with cells in adjacent ridges.

The second morphological feature of development is the movement of cells into aggregation centers where they form structures known as fruiting bodies. Fruiting bodies consist of about $10^5$ cells and within a fruiting body some of the cells lyse, whereas a small proportion of cells forms spherical spores, known as myxospores, which are resistant to harsh environmental conditions such as freezing, heat, dessication and UV irradiation (Shimkets, 1990). The morphological changes that occur during development are also accompanied by the coordinated expression of many developmentally regulated genes (Kuspa et al., 1986) and the synthesis of new proteins. At least 15 new proteins that appear as bands on protein gels were observed in extracts prepared from developing cells (Inouye et al., 1979b). Among the most abundant proteins shown to be expressed during development are the spore coat proteins S (found on the outer surface of spores) (Inouye et al., 1979a) and S-1 (found inside spores) (Teintze et al., 1985), which are the products of the *tps* and *ops* genes respectively. These proteins are 90% identical at the DNA sequence level (Inouye et al., 1983a; Inouye et al., 1983b), but are expressed at different times during development. The *tps* gene is expressed starting at about 5 hours into development, whereas the *ops* gene is expressed at about 24 hours into development (Inouye et al., 1983a; Inouye et al., 1983b; Downard & Zusman, 1985; Downard et al., 1984). Other proteins that are expressed during development include myxobacterial hemagglutanin (encoded by *mbhA* ), which is expressed 10 hours into development (Romeo & Zusman, 1991), protein C, a spore coat protein that is expressed early in development (McCleary et al., 1991), and phosphatases (Weinberg & Zusman, 1990).

Multicellular development in *M. xanthus* is coordinated by at least five extracellular signals known as the A-, B-, C-, D-, and E- signals. These extracellular signaling systems were identified by isolating conditional nonsporulating mutants that fail to undergo development (Hagen et al., 1978). When allowed to co-develop with wild-type cells, these mutants develop normally, indicating that they are defective in producing but not perceiving extracellular developmental signals. In a similar extracellular complementation experiment, development of these mutants was rescued when mixed in certain pair-wise combinations. These complementation experiments identified four classes of cell-cell signaling interactions (the A-, B-, C- and D-signaling interactions), and more recently a fifth signaling class, E-signaling, has been identified (Downard et al., 1993).

Kroos et. al. (Kroos & Kaiser, 1987; Kroos et al., 1986) demonstrated that extracellular signaling interactions regulate the expression of developmental genes using a Tn*5* *lac* promoter probe to generate transcriptional fusions of *lacZ* to developmentally regulated genes and assaying for the production of developmental ß-galactosidase in cells that were defective in an extracellular signaling interaction. These and later studies showed that genes expressed during the early hours of development require the A- and B- signaling interactions but do not require the C- and D- signaling interactions (Figure 1.1). Genes that are expressed starting at about four hours into development require the D-signaling interaction in addition to the earlier signal, and genes that are expressed after six hours require have an additional requirement for the C-signaling interaction. The requirement for the E-signaling interaction also appears to be early in development since the expression of *tps* (which is expressed at about 5 hours) is reduced in E-signaling deficient cells (Downard et al., 1993). These analyses suggested that the extracellular signaling interactions lie on the same developmental pathway, but also showed that genes within a given temporal class can exhibit different levels of dependence on a signaling event.

**Figure 1.1** Dependence of developmental markers on A, B, C, and D, factors. Each marker is placed above an arrow positioned along the x axis according to the time at which its expression normally begins during development. Position along the y axis indicates the marker's dependence on cell-cell interactions, as determined by measuring its expression in *asg*, *bsg*, *csa*, and *dsg* mutants. For example, a marker whose expression is reduced in an asg mutant exhibits partial dependence on A factor (abbreviated "partial A"). Similarly, a marker whose expression whose expression is abolished in a *bsg* mutant exhibits absolute dependence on B factor (abbreviated "absolute" B). Dependences are cumulative, so a higher position along the y axis indicates greater dependence on cell-cell interactions. Superecript a, d, c, or d means that the marker has not been tested in *asg*, bsg, *csg*, or *dsg* mutant cells, respectively. Tn*5* *lac* insertions are indicated by Ω followed by a four digit number, and brackets connect insertions that probably lie in the same transcription unit. Reprinted with permission from authors. Kaiser, D. and Kroos, L. (1993) In *Intercellular signalling*. (Dworkin, M. and Kaiser, D., eds), 257-283, American Society for Microbiology.


Q4506
Q4459[a]
hydrophilic shift
ops[a,c,d]
Q4401[d]
Q4497[b]
Q4435
Q4486[c,d]
Q4529[d]
ABSOLUTE D
Q4600[a,c,d]
Q4403[b]
Q4406
ABSOLUTE C
Q4531[a,c,d]
Q4489[a,d]
Q4400[a]
Q4473[b,d]
Q4414
Q4474[d]
ABSOLUTE B
mbhA
ABSOLUTE D
Q4427[c] Q4514[d]
Q4492[a,d]
alkaline phosphatase[b]
PARTIAL C
tps
PARTIAL D
Q4494
Q4491[a]
Q4427[c]
Q4457[d]
Q4521
Q4530[a,c,d]
Q4411
Q4445[a]
neutral phosphatase[b]
ABSOLUTE A
PARTIAL D
PARTIAL A
cagA[c]
PARTIAL C
Q4469
PARTIAL D
cagA[c]
Q4455
Q4408
PARTIAL B
protein C
ABSOLUTE A
PARTIAL D
acid phosphatase[b]
PARTIAL A
dFA-1
0 2 4 6 8 10 12 14 16 18 25 27
HOURS OF DEVELOPMENT

**A-signaling.** A-signaling is required early in development. A-signaling mutants fail to form compact aggregates, are defective in sporulation and fail to express genes that are expressed early during development (Hagen et al., 1978; LaRossa et al., 1983; Kuspa et al., 1986). Mutations affecting A-signaling map to three genetic loci, *asgA*, *asgB*, and *asgC* (Kuspa & Kaiser, 1989). These loci have been cloned and sequenced (Plamann et al., 1995; Plamann et al., 1994; Davis et al., 1995).

The deduced amino acid sequence of *asgA* shows a high degree of similarity to proteins of the two-component signal transduction systems. These systems are widely used in prokaryotes to regulate cellular responses to environmental changes (Parkinson, 1993). They regulate a variety of responses including nitrogen regulation, nitrogen fixation, chemotaxis, sporulation, transformation, competence, pathogenicity, gliding and flagellar motility, membrane transport and intermediary metabolism (Stock et al., 1990). A two-component regulatory system typically consists of a sensor protein (a histidine kinase) and response regulator protein. The sensor has a C-terminal transmitter domain coupled to an N-terminal input domain, whereas the response regulator typically consists of an N-terminal receiver module coupled to one or more C-terminal output domains (Parkinson, 1993). In most cases, the sensor protein is localized in the cytoplasmic membrane, and environmental stimuli detected by the input domain cause autophosphorylation at a histidine residue in the transmitter domain. The activated sensor protein then phosphorylates the response regulator at its receiver domain (typically an aspartate residue). The activated response regulator mediates the response (e.g., controls expression of one or more genes via its DNA-binding activity).

The AsgA protein contains two domains that have a high degree of amino acid identity to the transmitter domain of the sensor protein and the receiver domain of the response regulator (Plamann et al., 1995). This configuration is characteristic of a group of the two-component signal transduction proteins such as VirA and FrzE, which contain

both the sensor and response regulator in a single protein (Stock et al., 1990). Unlike these proteins, however, neither the sensor nor the response regulator of AsgA contain the input or output domains. Thus, it has been proposed (Plamann et al., 1995) that AsgA may function in a phosphorelay-type signal transduction mechanism similar to that used in the early stages of development in *Bacillus subtilis*.

The deduced protein product of the *asgB* gene also appears to be a regulatory protein based on the presence of a helix-turn-helix motif near the C-terminus, and its high degree of amino acid similarity to region 4.2 of major sigma factors (Plamann et al., 1994) which is involved in recognizing the -35 sequence of promoters. It has, therefore, been hypothesized that AsgB may recognize sequences similar to the -35 hexamer, TTGACA, and that it may also be a component of a phosphorelay system that regulates gene(s) involved in A-signal production.

The *asgC* mutation localizes to the gene encoding the putative major vegetative sigma factor of *M. xanthus*, *rpoD* (Davis et al., 1995), which may also be involved in the expression of genes required for A-factor production. Upstream of *rpoD*, other genes encoding proteins involved in the initiation of DNA synthesis and protein synthesis were identified (Davis et al., 1995). These genes, which include *dnaG* (encoding DNA primase) and *rpsU* (encoding ribosomal protein S21), have been proposed to make up the macromolecular synthesis operon that is conserved among many different groups of bacteria .

The A-factor itself has been purified from media conditioned by developing cells by monitoring the expression of an early expressed Tn*5* *lac* fusion marker, Ω4521, as an assay to identify fractions containing A-factor activity (Kaplan et al., 1991). Since Ω4521 is the earliest expressed Tn*5* *lac* fusion that is dependent on A-signaling, it provides a useful assay for the earliest appearance of A-factor. The A-factor isolated from conditioned media consisted of amino acids, peptides, and proteases (Plamann et al., 1992; Kuspa et

al., 1992b). Since single amino acids could rescue normal aggregation and sporulation of A-signaling mutant cells to wild-type levels, it is thought that the A-signal consists primarily of amino acids, peptides and proteases that are connected by a pathway in which proteases degrade proteins into peptides and finally into amino acids (Kuspa et al., 1992b).

Based on the observation of Kuspa et al. (1992a) that a threshold level of amino acids was required to rescue development, it was suggested that the concentration of extracellular amino acids (which increases proportionally to cell density) may serve as a mechanism for monitoring the density of starving cells. This model is supported by the finding that wild-type cells fail to develop when their density falls below the predicted density required for the production of the threshold level of A factor. Development of wild-type cells at this low density is rescued by the addition of amino acids or peptides (Kuspa et al., 1992b).

Other components of the A-signaling pathway were identified by isolating second-site suppressors of *asgB* that restore expression of an A-signal dependent gene (Kaplan et al., 1991). The *sasA* (suppressor of *asg*) gene product appears to be a member of the ATP-binding cassette (ABC) family of transport proteins (Guo et al., 1995), and the *sasB* gene product is similar to the histidine kinase family of proteins (Yang & Kaplan, 1995).

The emerging model for the A-signaling pathway is that starvation is sensed by Asg proteins (such as AsgA, AsgB and AsgC) which induce the expression of genes involved in A factor production and secretion. When cells reach the appropriate cell density, monitored by the concentration of A-factor, cells proceed to the next stage of development. Similar observations have been made in *B. subtilis* where the initiation of starvation is at least partially dependent on an extracellular peptide or peptides (known as the extracellular differentiation factor, EDF-A) (Kroos & Cutting, 1994). EDF-A is sensed by an ATP-binding cassette (ABC) transport protein and is used to measure cell density that functions in the decision whether to initiate endosporulation (Grossman & Losick, 1988;

Rudner et al., 1991). Perhaps *M. xanthus* SasA senses A-factor and functions in a similar manner. In *Streptomyces coelicolor* an extracellular peptide, SapB, was shown to serve as an extracellular signal during development (Willey et al., 1993). The mating behavior of *Enterococcus facealis* (Clewell & Weaver, 1989; Dunny, 1990) and of *Saccharomyces cerevisiae* (Herskowitz, 1989) depend on extracellular peptide signals.

In addition to the high cell density requirement for the initiation of the developmental process, cells must also be starved. The initial nutrient depletion sensed by the Asg proteins has been proposed to affect intracellular levels of (p)ppGpp. M. Singer and D. Kaiser (Singer & Kaiser, 1995) suggest that *M. xanthus* responds to starvation by integrating two starvation signals, intracellular (p)ppGpp concentration and extracellular A-signal concentration. These two signals may initiate and coordinately control entry into the developmental process. According to this model, nutritional limitations induce (intracellular) (p)ppGpp production. The intracellular (p)ppGpp regulates a number of cellular functions, including the expression of starvation dependent genes such as genes involved in the production of A-signal. The accumulation of the A-factor serves as a population starvation signal. Thus, when the threshold level of A-signal is attained, cells initiate entry into development. This model integrates the sensing of starvation in single cells with the coordination of the response to starvation within the population of cells.

**B-signaling.** Another signaling system required for development is B-signaling, which is required within two hours after the onset of development. B-signaling defective cells fail to form fruiting bodies and spores and do not express early developmental genes. Some B-signaling mutations map to the *bsgA* gene (Gill et al., 1988a). The *bsgA* gene product, BsgA, is an ATP-dependent cytoplasmic protease with a high degree of amino acid similarity to the Lon protease of *E. coli* and *Bacillus brevis* (Gill et al., 1993). Thus the role of BsgA in development may be to modulate the half-life of protein(s) involved in producing the B-signal. The activity of some regulatory proteins is modulated by

proteolysis. The CII protein of bacteriophage λ is a well known example. Proteolytic cleavage inactivates the LexA repressor, causing induction of the SOS response in *E. coli* (Ptashne, 1987). In *B. subtilis*, activation of the development-specific sigma factors $\sigma^E$ and $\sigma^K$ requires the proteolytic cleavage of precursor proteins. The proteolytic activity of BsgA may affect a regulatory protein and/or the proteins of the B-factor. However, the nature of the B-signal is not yet known.

**D-signaling.** D-signaling mutant (*dsg*) cells arrest development after about four hours. These mutants fail to complete aggregation, are defective in sporulation, and fail to express Tn*5* *lac* fusions whose expression normally begins about 5 to 6 hours into development. *dsg* cells are also defective in their vegetative colony morphology due to defects in motility but their vegetative growth is unaffected (Cheng & Kaiser, 1989b; Kalman et al., 1994). The two *dsg* point mutations identified map to the same locus, *dsgA* (Cheng & Kaiser, 1989a). Sequence analysis showed that the *dsgA* gene product is very similar to translation initiation factor IF3 of *E. coli* and *Bacillus stearothermophilus* (Cheng & Kaiser, 1989b), suggesting that *dsgA* mutant cells are defective in the translation of a gene or genes required for D-factor production. Although it is not known what the D-factor itself is, it has been proposed to be a mixture of fatty acids since fatty acids rescue the development of *dsgA* mutant cells (Rosenbluh & Rosenberg, 1989). An alternative explanation for this rescue is that fatty acids may be increasing membrane permeability and facilitating the transfer of other molecule(s) that are the D-signal.

In addition to the cellular requirement for DsgA during development, *dsgA* is also essential for cell viability (Cheng & Kaiser, 1989b). IF3 is normally required for the initiation of translation, and since there is only one copy of *dsgA* in *M. xanthus*, a null mutant cannot be obtained. The role of IF3 during translation initiation is in the selection of the AUG, GUG, or UUG translation initiation codons. The translation initiation codon

for the gene encoding IF3 is AUU, and it has been suggested to be involved in autoregulation (Butler et al., 1987; Gold et al., 1984). A low level of IF3 in the cell causes less stringent selection of the translation initiation codon by the 30S ribosomal subunit, and translation initiation can occur at rare codons, such as AUU. This allows translation of the IF3 gene when IF3 is in low abundance; and when the level of IF3 is high, selection of the initiation codon is stringent and translation is prevented from the rare AUU initiation codon. It has been hypothesized that in *M. xanthus*, the *dsgA* mutant gene may encode an IF3 that is more stringent in its selection of a translation initiation codon. Thus, the mutant IF3 does not recognize rare translation initiation codons, and would fail to translate some vegetative as well as developmental genes.

**E-signaling.** Cells defective in E-signaling also fail to form fruiting bodies and spores. Expression of an early developmental gene (*tps*) is reduced in these cells, indicating that E-signaling affects early events. The mutation affecting E-signaling maps to the *esg* locus (Downard et al., 1993). The *esg* locus consists of two ORFs with similarities to the a and b subunits of the E1 component of the branched-chain keto acid dehydrogenase (BCKAD), a multienzyme complex (Toal et al., 1995). BCKAD is involved in the synthesis of branched-chain fatty acids (BCFAs) from branched chain amino acids (BCAAs). During vegetative growth, BCFAs are synthesized and incorporated into membrane phospholipids. However, since the *esg* mutation reduces but does not abolish synthesis of these fatty acids, alternative routes must exist for their synthesis. Furthermore, *esg* cells grown vegetatively in the presence of short-chain fatty acids are rescued for development, indicating that BCFAs synthesized during vegetative growth are used later in cell-cell signaling . Thus, one or more of the fatty acids may be the E-signal. Alternatively, these fatty acids may function less directly in E-signaling, for example, they may modify the activity of the protein that is transmitted between cells.

**C-signaling.** C-signaling is one of the best characterized signaling systems in *M. xanthus*. All of the mutations affecting C-signaling map to a single gene known as *csgA* (Shimkets & Asher, 1988) (Shimkets et al., 1983). C-signaling mutant cells initiate aggregation but fail to form tight mounds and do not sporulate. Unlike wild-type cells which form compact translucent mounds by 12 hours, *csgA* cells are still in larger, less compact mounds at 18 hours into development (Kim & Kaiser, 1990a). In addition to defects in fruiting body formation and sporulation, *csgA* cells fail to form ripples and they have reduced or abolished expression of genes that are normally expressed after six hours into development (Kroos & Kaiser, 1987; Shimkets & Kaiser, 1982b).

Defects in C-signaling can be rescued by co-development of *csgA* cells with wild-type cells or with mutant cells in a different signaling group. Additionally, many compounds, including several carbohydrates and amino acids that are biochemical precursors to peptidoglycan such as N-acetylglucosamine, N-acetylmuramic acid, D-alanine and diaminopimelate can induce *csgA* cells to sporulate (Shimkets & Kaiser, 1982b; Shimkets, 1990). However, the actual nature of the C-signal has yet to be established.

Extensive effort has been devoted to identifying and characterizing the C-signal. Kim and Kaiser showed that a 17 kDa protein tightly associated with cells rescues fruiting body formation and sporulation of *csgA* cells (Kim & Kaiser, 1990b; Kim & Kaiser, 1990a). In addition to this result, the following observations indicate that C-signaling involves a short-range signal whose transmission is restricted to contiguous cells. First, when *csgA* responder cells were separated from wild-type C-factor donor cells by a 0.45 μm nucleopore membrane, the *csgA* cells could not be rescued for development, indicating that the C-factor is not diffusible (Kim & Kaiser, 1990a). Second, extracts made by extensive washing of wild-type cells at various developmental stages also failed to rescue development of *csgA* cells, whereas extracts made from lysed developing wild-type cells restored both fruiting body formation and sporulation to *csg* cells (Kim & Kaiser, 1990b).

Cloning of the *csgA* gene and initial sequence analysis suggested that it encodes a 17.7 kDa protein (Shimkets et al., 1983; Hagen & Shimkets, 1990), consistent with the results that a 17 kDa protein purified from developing cells rescues development of *csgA* mutant cells. The N-terminal amino acid sequence of the purified 17 kDa protein was the same as the deduced N-terminal amino acid sequence of *csgA*. Moreover, polyclonal antibodies raised against a *lacZ-csgA* fusion protein reacted with the purified 17 kDa C-factor protein. Thus, it appears that the purified 17 kDa protein is encoded by the *csgA* gene.

Recent studies have shown that the *csgA* gene product, CsgA, may actually be larger than previously reported. Extending the *csgA* ORF 5' to another possible translation initiation codon increases the size of the predicted protein to 24 kDa (Lee & Shimkets, 1994). The amino acid sequence of this protein is 19-39 % similar to the sequences of a family of short-chain alcohol dehydrogenases. These alcohol dehydrogenases have a wide variety of substrates and are involved in the production of many different signaling molecules including steroids, prostaglandins, and nodulation factors in plants (Baker, 1994; Persson et al., 1991). In addition to amino acid sequence similarity, several results are consistent with *csgA* encoding a protein in the alcohol dehydrogenase group. First, one class of *csgA* suppressors, *socA* (suppressor of *csg*), identified an operon that has a gene, *socA1*, whose deduced amino acid sequence is also similar to the family of short-chain alcohol dehydrogenases. Second, site-directed mutagenesis of the predicted active site of the CsgA alcohol dehydrogenase causes a Csg- phenotype (Lee et al., 1995a). Third, NAD+, a coenzyme predicted to bind short-chain alcohol dehydrogenases, stimulates a MalE-CsgA fusion protein to rescue development of *csgA* cells (Lee et al., 1995a). Thus, CsgA may be an enzyme and it may be involved in the production of C-factor rather than being the C-factor itself.

Several genes that circumvent the need for the C-signal transduction have been

isolated as second-site suppressors of a *csgA* mutant. The *socA* cistron encodes three proteins (Lee & Shimkets, 1994). In addition to *socA1* (whose deduced amino acid sequence is similar to alcohol dehydrogenases), the predicted protein product of *socA2* has a 48% amino acid sequence identity with the *Proteus vulgaris frd* gene product, whose function is to anchor fumarate reductase to the membrane. The third gene, *socA3*, contains the Tn*5* supressor mutation. However, the deduced amino acid sequence of *socA3* does not show any similarity with known proteins. It has been proposed that the wild-type *socA3* gene product negatively regulates the *socA* operon (Lee & Shimkets, 1995). Thus, cells containing the *socA* suppressor mutation are unable to produce SocA3, resulting in overproduction of SocA1. Since the *socA1* gene product belongs to the same group of alcohol dehydrogenases as CsgA, increased production of SocA1 is thought to overcome the *csgA* defect (Lee & Shimkets, 1995). However, even though a transposon insertion in *socA3* restores sporulation to *csgA* mutant cells, the spores produced by these double mutants germinate prematurely and are unstable upon prolonged plating on starvation medium. This indicates that suppression is incomplete and may explain why *csgA* has been retained through evolution.

*socD500* is a second mutation that suppresses the sporulation defect of *csgA* cells under normal developmental conditions (low nutrient levels, high cell density and presence of a solid surface at 32°C). *socD500* also allows sporulation of *csg+* or *csgA* cells at 15°C in the presence of nutrients (Rhie & Shimkets, 1991). In addition to these phenotypic changes, at 15°C *socD500* activates expression of *tps*, whose expression precedes that of *csgA*, and of *ops*, whose expression is after that of *csgA*. This indicates that both early and late gene expression are affected by *socD500*. Furthermore, *socD500* restores sporulation to *csgA* cells in the absence of nutrients without restoring aggregation, suggesting that the *socD500* mutation may be in a component of the nutrition sensory mechanism that has a regulatory function at several points in the developmental pathway.

The *socD500* gene encodes a 61.6 kDa histidine protein kinase (Lee et al., 1995b), and the mutations in *socD* that lead to the suppression of *csgA* map to the N-terminal sensor domain of the protein. It has thus been hypothesized that an input signal binds SocD and perpetuates growth, and mutations in this regulatory protein trigger sporulation. Consistent with this model, null mutations in *socD500* could not be obtained (Lee et al., 1995b) suggesting that SocD is required for growth.

Different concentrations of purified CsgA trigger distinct morphogenetic and transcriptional responses (Kim & Kaiser, 1991). When purified 17 kDa polypeptide was added to starving *csgA* cells at a concentration of 0.8 nM, aggregation and the expression of a *lacZ* fusion that is normally expressed earlier in development and is only partially dependent on C-signaling was rescued. At a concentration of 1 nM or more, the purified 17 kDa polypeptide rescued aggregation, sporulation and the expression of a *lacZ* fusion that is normally expressed late in development and has an absolute requirement for C-signaling (Kim & Kaiser, 1990a). This indicates that *M. xanthus* cells can discern and respond to small changes in C-factor concentrations. Similarly, Li et al. showed that increased expression of *csgA in vivo* over the course of development entrains the natural sequence of morphogenetic changes during development (Li et al., 1992).

The regulation of the *csgA* gene itself appears to be complex. During growth, *lacZ* expression in *csg* cells containing a *csgA-lacZ* transcriptional fusion is low, suggesting that *csgA* is expressed at very low levels in vegetatively growing cells (Kim & Kaiser, 1991). As cells enter the developmental stage, *lacZ* expression increases by about two-fold. When starving *csgA* cells (containing the *csgA-lacZ* fusion) are mixed with either an equal number of wild-type cells or with 1 nM of purified C-factor, *lacZ* expression increases by about four-fold . This suggests that the initial increase in C-factor production triggers the positive regulation of CsgA synthesis and the aggregation of cells. As cells progress through development, aggregation leads to an increase in local C-factor concentration due

to the increase in cell density and increased efficiency of C-factor transmission (see below). At about 18-20 hours spore formation is initiated and the level of CsgA decreases due to lysis of cells.

Based on the autoregulation of *csgA* and the requirement for a threshold level of C-factor for the induction of the different developmental stages, Kim and Kaiser (Kim & Kaiser, 1991) proposed that the C-factor may be functioning as both a morphogenetic trigger and as a developmental timer whose increasing concentration during development triggers different morphogenetic and transcriptional changes. According to this model, low levels of C-factor, present early in development, are sufficient for the initiation of aggregation and the expression of early developmental genes. As cells aggregate, the local concentration of C-factor increases, inducing the production of a higher level of C-signal and, subsequently, inducing sporulation and late gene expression. As a morphogenetic signal, C-signal appears to allow cells to sense when they have moved into a fruiting body and then triggers their differentiation into spores.

An interesting aspect of *M. xanthus* development is its requirement for motility. Nonmotile mutant cells (A-S- or *mgl*-) have similar defects in development as *csgA* null mutants; they fail to complete aggregation, they do not sporulate and do not fully express developmental markers that are normally turned on between five and nine hours into development (Kroos et al., 1988). This suggests that nonmotile cells may be defective in the same signaling interactions as *csgA* mutant cells (Kim & Kaiser, 1990a). Unlike *csgA* mutants, however, nonmotile mutants are proficient in the production of C-signal. Development of *csgA* cells was restored by adding cell-free extracts of developing *mglA* cells, demonstrating that *mglA* cells produce the C-signal. Addition of purified CsgA to non-motile cells restores sporulation and expression of developmentally regulated *lacZ* fusions, demonstrating that *mglA* cells can perceive the C-signal (Kim & Kaiser, 1990a). On the other hand, aggregation was not restored, indicating that the motility defect was not

overcome by the addition of purified CsgA. Mixing *mglA* cells(which do produce CsgA) either with wild-type cells or with *csgA* cells does not restore sporulation or expression of developmental *lacZ* fusions in *mglA* cells to wild-type levels (Kim & Kaiser, 1990a), supporting a model in which both donor and responder cell motility is required for C-signal transmission.

The role of motility in development appears to be to allow cells to aggregate and increase the local cell density since increasing the initial cell density of starving *mglA* cells increases sporulation efficiency (Kroos et al., 1988). However, even a ten-fold increase in the initial density of the *mglA* cells over the concentration of wild-type cells that is normally required for development increased the sporulation efficiency to only 1 % of the wild-type levels (and the expression of developmental *lacZ* fusions did not increase above background levels). These results indicate that a mere increase in cell density is not sufficient for restoring development in *mglA* cells. Kim and Kaiser therefore suggested that C-factor transmission may require a critical spatial orientation of cells which is acheived after cells move into the dense, highly aligned organization (of cells) in a nascent fruiting body. Indeed, scanning electron microscopic observations show that rod-shaped cells within a fruiting body are densely aligned in parallel rows (Kim & Kaiser, 1990a; O'Connor & Zusman, 1989). The movement of cells is thought, therefore, to align cells in a pattern that greatly increases the local cell density and efficiency of C-factor transmission. In support of this model, manipulating the orientation of nonmotile cells on the surface of starvation agar so that the cells are aligned in parallel rescues the developmental defect, and the expression of late developmental Tn*5* *lac* fusions (Kim & Kaiser, 1990a). Similar alignment of *csgA* cells rescues neither development nor the expression of late developmental markers, supporting the idea that motility is required to mediate C-signaling.

C-signaling in *M. xanthus* has characteristics that are similar to signaling processes

in eukaryotes. The most striking similarity is with cAMP signaling during the starvation-induced development of *Dictyostelium discoideum*. *D. discoideum* is a common soil protozoan that grows unicellularly in the presence of nutrients, but undergoes multicellular development when nutrients are depleted (Kaiser, 1986). Development in *D. discoideum* entails aggregation of cells into a mass of about $10^5$ cells, forming a nascent structure known as a slug. The cells that form the slug are motile, and later in development the slug migrates and forms a fruiting body consisting of two types of differentiated cells. The stalk consists of cells which secrete cellulose and eventually mature into nonviable stalk cells containing vacuoles and thick cellulose walls. Atop the stalk, the majority of cells differentiate into dormant spores which initiate vegetative growth when nutrients are available.

Initiation of development in *Dictyostelium* requires signaling by cAMP. Upon starvation adenylate cyclase activity is induced, increasing levels of both intracellular and extracellular cAMP. Production of cAMP is autoregulatory, with the binding of extracellular cAMP to the cAMP receptor further inducing the cellular adenylate cyclase activity. The cAMP positive feedback loop is, however, tempered by an adaptive chemosensory system that is responsive to the level of receptor-bound cAMP. In addition to adaptation, the cAMP level is also reduced by phosphodiesterase activity. Together, the adenylate cyclase feedback loop, adaptation and phosphodiesterase activity cause oscillations in the levels of cAMP. These oscillations coordinate movement of cells in a way that resembles the C-signal-dependent rippling observed in the early stages of *M. xanthus* development.

Clearly, despite their evolutionary distance, interesting parallels can be drawn between the developmental processes of *Myxococcus* and *Dictyostelium*. In both cases starvation induces intercellular signaling (C-signaling in *M. xanthus* and cAMP signaling in *D. discoideum*), directed cell movement into aggregates, and differentiation of vegetative

cells into environmentally resistant spores. Directed cell movement and the positive feedback loop of cAMP production in *Dictyostelium* are analogous to the C-signal mediated movement of cells into aggregates and the autoregulation of of C-factor production (Kim, 1991). This suggests that elucidation of the mechanisms of multicellular development that occurs in *Myxococcus* may contribute to our understanding of developmental processes in other multicellular organisms.

The organization of *M. xanthus* cells within a fuiting body also offers a simple model system for studying pattern formation and spatial regulation of gene expression. The nascent fruiting body of *M. xanthus* consists of an outer hemispherical domain of highly organized, densely packed, motile, rod-shaped cells that are aligned in parallel rows (Sager & Kaiser, 1993b; O'Connor & Zusman, 1989; Kaiser et al., 1985). Cells within this outer domain move in a bidirectional stream orbiting the inner domain which consists of relatively less ordered, loosely packed nonsporulating cells (Sager & Kaiser, 1993b). Because spores are nonmotile, they are passively transported towards the center of the fruiting body by the driving force of the rod-shaped cell that move in circular orbits within the outer domain (Sager & Kaiser, 1993a). Since cells in the outer domain of fruiting bodies are at high cell density, ordered and motile, C-signaling is most efficient in this domain. Differentiation and expression of C-signal-dependent *lacZ* fusions, therefore, initiates in the outer domain (Sager & Kaiser, 1993a). As the undifferentiated rod-shaped cells continue to circle the inner domain, and proceed through development, the differentiating spores lose their motility and accumulate in the inner domain. Accordingly, expression of developmentally regulated *lacZ* fusions is higher in the inner domain as these cells have been expressing the fusion for a longer time.

Thus, the short-range, spatially restricted, C-signaling system that regulates both the spatial and temporal expression of developmental genes in *M. xanthus* provides a good model system for studying the mechanism by which short-range signals are used in

differentiation of a homogeneous population of cells. Perhaps the best understood case in which the importance of short-range signaling in regulating differentiation has been demonstrated is vulval development of *Caenorhabditis elegans*. In addition to long-range diffusible signals, vulval precursor cells respond to short-range signals that determine their fates. These signals are responsible for the formation of precise spatial patterns of three cell types that develop into the vulva (Sternberg & Horvitz, 1989). Thus, an understanding of the mechanims by which C-signaling regulates differentiation in a fruiting body may shed new light on the mechanisms involved in pattern formation in eukaryotes. Additionally, short-range signaling is also important in the interaction between plant and animal hosts and microbes. These species-specific interactions establish the outcome of the interactions as beneficial or detrimental (Clark et al., 1992).

The long-term goal of our studies is to unravel the molecular mechanism(s) by which extracellular C-signaling regulates gene expression during starvation-induced development in *M. xanthus*. *M. xanthus* has a simple genetic system which is amenable to manipulation, and many genetic tools are available that facilitate these studies. Biochemical studies are possible as large quantities of cells can be grown and developed synchronously for isolation of proteins. Thus, *M. xanthus* provides a simple and tractable multicellular system for studying the mechanism by which intercellular signaling regulates gene expression.

Regulation of gene expression in response to environmental cues is mediated by a variety of mechanisms in prokaryotes. Typically, regulation occurs at the transcriptional level. Many bacteria possess multiple sigma subunits of RNA polymerase and use alternative forms of RNA polymerase holoenzyme to express different sets of genes in response to a wide variety of environmental signals, including heat and cold shock, osmotic pressure, nitrogen limitation and starvation. A well studied system in which alternative $\sigma$ factors have been shown to regulate gene expression is the sporulation process

of *B. subtilis* (Haldenwang, 1995). Starvation-induced sporulation in *B. subtilis* involves transcription of developmentally regulated sets of genes. A cascade of σ-factors, that are activated at different stages during development at the transcriptional as well as post-transcriptional levels, transcribe genes in a spatially and temporally regulated manner. At least two post-transcriptional activation mechanisms activate sigma factors in *B. subtilis*; proteolytic removal ofenzy inhibitory sequences at the N-terminus (the pro-region) and competition for binding of inhibitory proteins (anti-sigma factors). Thus, expression of sporulation genes is fine-tuned at different levels by regulating the expression as well as the activation of regulatory proteins.

Multiple sigma factors have also been identified in *M. xanthus* (Inouye, 1990; Apelian & Inouye, 1990; Apelian & Inouye, 1993). The *sigA* gene product, $\sigma^A$, is thought to be the major sigma factor in vegetatively growing cells and has amino acid sequence similarity to $\sigma^{70}$ of *E. coli* and $\sigma^{43}$ of *B. subtilis* (Inouye, 1990). At least three sigma factors are produced during development, $\sigma^C$, the *sigC* gene product, is expressed early in development (Apelian & Inouye, 1993), and $\sigma^B$, the *sigB* gene product, is expressed late in development (Apelian & Inouye, 1990). Furthermore, there is evidence for a third developmental sigma factor, which has similarity to the $\sigma^{54}$ family, and appears to be involved in early developmental gene expression (Keseler & Kaiser, 1995). Thus gene expression during development in *M. xanthus* also entails the use of alternative sigma factors for differential gene expression.

The regulation of many bacterial genes is mediated by transcriptional activator proteins. In genes transcribed by $\sigma^{54}$ RNA polymerase, activator proteins bind upstream of the promoter and activate transcription by catalyzing the formation of a tertiary complex. In some cases, these activation sequences are far removed from the promoter, and DNA

looping mediates their interaction with the promoter. Several *M. xanthus* genes have been identified that appear to be transcribed by $\sigma^{54}$-type RNA polymerase and which require upstream activation sequences. The *mbhA* gene, which is expressed early in development and encodes myxobacterial hemagglutinin, is transcribed from a $\sigma^{54}$-type promoter and requires cis-acting sequences between 86 and 276 bp upstream of the promoter (Romeo & Zusman, 1991). Expression of the developmentally regulated Tn*5* *lac* Ω4521 fusion also appears to be from a $\sigma^{54}$-type promoter, and requires sequences upstream for transcription activation.

Other genes are also present in *M. xanthus* that require upstream activation sequences for transcription. A DNA segment located between 131 and 311 bp upstream of the *ops* transcriptional start site activate not only *ops* transcription but also transcription of *tps* promoter which is located about 2 kb downstream of *ops* (Kil et al., 1990). The activation of *tps* by this DNA was orientation independent, and thus appears to be mediated by sequences that function much like eukaryotic enhancers. The *ops* UAS segment has also been shown to have a DNA-binding activity (Downard & Kroos, 1993), suggesting that proteins that activate transcription may bind in this region. Expression of *tps* is also activated by two UASs regions between 82 and 110 and 173 to 375 bp upstream of the its transcriptional start site (Downard & Kroos, 1993).

The promoter region of *csgA* also requires a fairly large upstream sequence to modulate its expression in response to nutrient levels, peptidoglycan components and B-signaling (Li et al., 1992). At least 400 bp upstream from the start of transcription initiation were required for maximum expression of *csgA*. In the presence of low levels of nutrients, sequences extending to about 930 bp were required for the developmentally regulated expression of *lacZ*.

Another widely used mechanism of gene regulation in prokaryotes is the two-

component signal transduction system. Chemotaxis in *E. coli* employs this system with several sensor proteins regulating the rate of swimming and tumbling in the presence of attractants and repellants, and *M. xanthus* proteins similar to these chemotaxis proteins have been shown to regulate reversal of gliding movement. Mutations in these regulatory proteins affect fruiting body formation and sporulation.

Similarly, mechanisms that employ transcriptional or post-transcriptional activation of sigma factors or two-component systems may be involved in C-signal-mediated gene expression in *M. xanthus*. On the other hand, novel mechanisms may be involved in coordinating extracellular signaling with gene expression. Interestingly, eukaryotic-type protein serine/threonine kinases and tyrosine kinase(s) have been identified in *M. xanthus* (Munoz-Dorado et al., 1991; Frasch & Dworkin, 1995). Deletion of at least three different protein kinases, PK1, PK5, and PK6, (Udo et al., 1995; Zhang et al., 1995) has been demonstrated to affect development. It, therefore, seems possible that signal transduction systems similar to those in eukaryotes may be utilized by the C-signaling system.

# Chapter 2

## Characterization of the regulatory region of a cell interaction-dependent gene in *Myxococcus xanthus*

(The work described in this chapter will be submitted for publication with Monica Gloudemans, Ron Gill, and Lee Kroos as second, third, and fourth authors, respectively).

## Abstract

Ω4403 is the site of a Tn*5* *lac* insertion in the *Myxococcus xanthus* genome that fuses *lacZ* expression to a developmentally regulated promoter. Cell-cell interactions that occur during development, including C-signaling, are required for expression of Tn*5* *lac* Ω4403. We cloned DNA upstream of the Ω4403 insertion site, localized the promoter, and identified a potential open reading frame. Based on the deduced amino acid sequence, the gene disrupted by Tn*5* *lac* Ω4403 appears to encode a serine protease that is dispensable for development. The gene begins to be expressed between 6 and 12 h after starvation initiates development, as determined by measuring mRNA or ß-galactosidase accumulation in cells containing Tn*5* *lac* Ω4403. The putative transcriptional start site was mapped and sequences centered near -10 and -35 show some similarity to these regions of promoters transcribed by *E. coli* $\sigma^{70}$ RNA polymerase. However, deletions showed that an essential promoter element lies between -80 and -72, suggesting the possible involvement of an upstream activator protein. DNA downstream of -80 is sufficient for C-signal-dependent activation of this promoter. The promoter is not fully expressed when fusions are integrated at the Mx8 phage attachment site in the chromosome. Titration of a limiting factor by two copies of the regulatory region (one at the attachment site and one at the native site) can, in part, explain the reduced expression. We speculate that the remaining difference may be due to an effect of chromosomal position. These results provide a basis for studies aimed at identifying regulators of C-signal-dependent gene expression.

## Introduction

*Myxococcus xanthus* is a gram-negative soil bacterium that undergoes multicellular development (Dworkin & Kaiser, 1993). When starved at high cell density on a solid surface, cells move into aggregation centers where they construct a fruiting body comprised of about $10^5$ cells. Within the fruiting body cells differentiate into dormant, ovoid spores.

In addition to morphological changes, *M. xanthus* development involves a highly ordered program of gene expression controlled by cell-cell interactions (Kaiser & Kroos, 1993). At least five cell-cell signals, known as the A-, B-, C-, D- and E- signals, are required for normal development. Cells with mutations in genes needed for any of these cell interactions arrest development at a specific stage, but are rescued when they are co-developed with wild-type cells or with mutant cells from a different signaling group (Downard et al., 1984; Hagen et al., 1978; LaRossa et al., 1983).

To study the role of cell-cell interactions in controlling *M. xanthus* gene expression, Tn*5 lac* has been used to identify developmentally regulated genes. Tn*5 lac* is a transposable promoter-probe containing a promoterless *lacZ* gene inserted near one end (Kroos & Kaiser, 1984). Transposition of Tn*5 lac* into the *M. xanthus* chromosome can generate a transcriptional fusion between *lacZ* and an *M. xanthus* promoter. Among 2,374 Tn*5 lac* insertions, Kroos *et al.* (Kroos et al., 1986) found 29 in distinct transcriptional units whose activity increases during development. Dependence of developmental gene expression on cell-cell interactions was examined by monitoring ß-galactosidase expression of the *lac* fusions in cell-interaction mutants. A- and B-signaling were required for normal gene expression beginning at one to two hours into development (Gill & Cull, 1986; Kroos & Kaiser, 1987; Kuspa et al., 1986). D-signaling was first required at about four hours (Cheng & Kaiser, 1989a). C-signaling was required for normal expression of nearly all genes that begin to be expressed after six hours into development (Kroos & Kaiser, 1987;

Li & Shimkets, 1993). It has recently been shown that E-signaling is required early in development, as demonstrated by the lack of expression of an early developmental gene (*tps*) in an E-signaling mutant (Downard & Kroos, 1993).

We are focusing on the regulation of developmental gene expression by C-signaling for several reasons. First, the gene expression studies mentioned above show that the C-signal-mediated cell-cell interaction occurs later in development than the other signaling interactions. Since the the A-, B-, D- and E-signals are all required early in development, it may be difficult to decipher the role of each of these signals in controlling gene expression. In addition, nutritional downshift signals may also control early developmental gene expression. Another reason for focusing on C-signal-dependent gene expression is that the *csg* group of mutants, which are defective in C-signaling, are genetically simple and well-characterized. All of the *csg* mutations map to a single gene, *csgA*, whose product appears to mediate the cell-cell interaction (Hagen & Shimkets, 1990; Kim & Kaiser, 1990a; Shimkets & Asher, 1988; Shimkets et al., 1983). CsgA is associated with the cell surface membrane and/or extracellular matrix (Kim & Kaiser, 1990b; Shimkets & Rafiee, 1990). Upon solubilization and partial purification, CsgA at approximately 1 to 2 nM restores normal development to *csgA* mutant cells (Kim & Kaiser, 1990a). Recent results suggest that CsgA may function as an enzyme to generate an unknown signal molecule that is perceived by other cells (Lee et al., 1995b; Lee & Shimkets, 1995).

C-signaling is also interesting because it requires cell movement for transmission (Kim & Kaiser, 1990b; Kim & Kaiser, 1990a)(Kroos *et al.*, 1988). During development, cells move into alignment and become densely packed in the outer domain of a nascent fruiting body (Sager & Kaiser, 1993b). Sager and Kaiser (Sager & Kaiser, 1993a) have proposed that end-to-end and/or side-to-side contacts between cells in the outer domain permit efficient C-signaling, triggering sporulation. Consistent with this model, patches of spore precursors and expression of several C-signal-dependent *lacZ*

fusions is first observed in the outer domain (Sager & Kaiser, 1993a). Hence, expression of C-signal-dependent genes is regulated spatially, as well as temporally. C-signaling appears to couple morphogenesis of the fruiting body with expression of late genes and differentiation of cells into spores. In addition to coordinating the aggregation and sporulation behavior of *M. xanthus* cells, C-signaling is required for rippling (Sager & Kaiser, 1994; Shimkets & Kaiser, 1982a). Ripples form early in development when cells organize themselves into ridges that move as traveling waves. A rising level of CsgA has been proposed to entrain the developmental sequence of rippling, then aggregation, and finally sporulation (Li & Shimkets, 1988). A higher level of CsgA has been shown to be required to activate a later-expressed C-signal-dependent gene than an earlier-expressed gene (Kim & Kaiser, 1991).

To begin to understand how C-signaling regulates gene expression during *M. xanthus* development, we are characterizing the DNA regulatory regions of C-signal-dependent genes. Here, we report the characterization of the regulatory region of Ω4403, a gene identified by Tn*5* *lac* insertion (Kroos et al., 1986). Expression from Tn*5* *lac* Ω4403 starts at about 11-15 hours into development and has an absolute requirement for C-signaling (Kroos & Kaiser, 1987). We constructed deletions to localize the sequences required for promoter activity, determined the nucleotide sequence of the promoter region, and mapped the transcriptional start site. Promoter activity required sequences farther upstream of the transcriptional start site than is usually required for RNA polymerase binding, suggesting the involvement of an additional regulatory protein(s). Analysis of the minimum promoter region in *csgA* mutant cells showed that it requires C-signaling for activity. These studies lay the foundation to start identifying regulatory proteins required for C-signal-dependent expression of genes.

## Materials and Methods

**Bacterial Strains and plasmids.** Strains and plasmids used in this work are listed in Table 2.1.

**Growth and Development.** *Escherichia coli* cells were grown at 37°C in LB medium containing 50 μg ampicillin or 25 μg kanamycin per ml when necessary. *M. xanthus* was grown at 32-34°C in CTT medium (Hodgkin & Kaiser, 1977) in liquid cultures or on agar plates as described previously (Kroos et al., 1986). Forty μg of kanamycin or 12.5 μg of oxytetracycline per ml were used when required for selective growth. Fruiting body development was performed on TPM (10 mM Tris-HCl, pH 8.0, 1 mM $KH_2PO_4$, 8 mM $MgSO_4$, final pH 7.6) agar (1.5%) plates as described previously (Kroos et al., 1986).

**Molecular cloning and construction of plasmids.** Recombinant DNA work was performed using standard techniques (Sambrook et al., 1989). Plasmid DNA was prepared from *E. coli* DH5α or JM83. *M. xanthus* chromosomal DNA used for cloning was prepared as described previously (Laue & Gill, 1994). *M. xanthus* DNA upstream of Tn*5 lac* Ω4403 was cloned by digesting chromosomal DNA from strain DK4368 with *Xho*I and ligating into *Xho*I-digested pGEM-7Zf, resulting in pMES003 (Table 1). Restriction fragments of *M. xanthus* DNA from pMES003 were gel-purified and ligated into vectors as indicated in Table 1. Vectors were digested with the same restriction enzymes used to produce the fragments, except as indicated below.

To test *M. xanthus* DNA fragments for promoter activity, we constructed pREG1666 and its derivative pREG1727. pREG1666 was derived from pREG1175 (Gill et al., 1988a) as follows. A 3 kbp *Sma*I fragment from phage Mx8 (Stellwag et al., 1985) was ligated into *Sma*I-digested pUC18 (Yanisch-Perron et al., 1985), then excised

Table 2.1 Bacterial Strains and Plasmids

| Strain or plasmid | Relevant characteristics | Source or reference |
|---|---|---|
| *E.coli* | | |
| DH5α | ϕ80 Δ*lacZ* ΔM15 Δ*lacU169 recA1 endA1 hsdR17 supE44 thi-1 gyrA relA1* | Hanahan, D., 1983 |
| JM83 | *ara* Δ*lac-pro strA thi* ϕd*lacZ* ΔM15 | Messing, D., 1979 |
| *M. xanthus* | | |
| DK1622 | Wild-type | Kaiser, D., 1979 |
| DK4368 | Tn*5 lac* ($Km^r$) Ω4403 | Kroos et al., 1986 |
| DK4499 | Tn*5 lac* ($Km^r$) Ω4499 | Kroos et al., 1986 |
| JW103,104, 107 | *attB*::pREG1666 | Jamie White |
| MMF1727-7, -9, -11 | *attB*::pREG1727 | This work |
| MES005, 008, and 012 | *attB*::pMES004 | This work |
| MMF71-6, -15, and -22 | *attB*::pMF71-1 | This work |
| MES034, 036, 047 | *attB*::pMES116 | This work |
| MES053, 064, 077 | *attB*::pMES115 | This work |
| MMF31-3, -7, and -9 | *attB*::pMF31 | This work |
| MMF81-1, -8, and -32 | *attB*::pMF81 | This work |
| MMF100-6, -17, and -24 | *attB*::pMF100 | This work |
| MMF72-2, -10, and -32 | *attB*::pMF72 | This work |
| MMF62-4, -34 | *attB*::pMF62 | This work |

Table 2.1 (cont'd)

| | | |
|---|---|---|
| MMF52-35, -44, -45 | *attB*::pMF52 | This work |
| MMF301-3, -4, -5 | *attB*::pMF301 | This work |
| MMF200-5, -6, and -7 | *attB*::pMF200 | This work |
| MES014, 016, and 017, | *Ω4403*::pMES108 | This work |
| MSH1 | *Ω4403*::pSH1 | Stacie Hill |
| JPB07 | Tn*5 lac* ($Tc^r$) Ω4403 | Janine Brandner |
| MLK1-3, -4, and -5 | *Ω4403*::pLK1Tn5 lac ($Tc^r$) Ω4403 | This work |
| DK5208 | *csgA*::Tn5-132 ($Tc^r$) Ω205 | Shimkets, L., and Asher, 1988 |
| MES119 | *csgA*::Tn*5-132* ($Tc^r$) Ω205 *attB*::pREG1666 | This work |
| MMF100C-6, -17, -24 | *csgA*::Tn*5-132* ($Tc^r$) Ω205 *attB*::pMF100 | This work |
| MMF200C-4, -7, -11 | *csgA*::Tn*5-132* ($Tc^r$) Ω205 *attB*::pMF200 | This work |

| | | |
|---|---|---|
| pUC19 | $Ap^r$ *lacα* | Yanisch-Perron et al., 1985 |
| pGEM-7Zf | $Ap^r$ *lacα* | Promega |
| pREG1666/pREG1727 | $Ap^r$ $Km^r$ P1-*inc attP* | This work |
| pREG1175 | $Ap^r$ $Km^r$ P1-*inc* | Gill, R. and Bornemann, M., 1988 |
| pREG429 | $Ap^r$ $Km^r$ P1-*inc* | Gill et al. 1988 |
| pMES003 | $Ap^r$ (pGEM-7Zf), *M. xanthus* Ω4403, $Km^r$ (*aphII*), 16.7 kb *Xho*I fragment from DK4368 | This work |
| pMES004 | $Ap^r$ $Km^r$ P1-*inc attP* (pREG1666), 8.5 kb *Xho*I-*Bam*HI fragment from pMES003 | This work |

Table 2.1 (cont'd)

| | | |
|---|---|---|
| pMES110 | Ap$^r$ (pGEM-7ZF) 4.0 kb *Cla*I-*Bam*HI fragment from pMES003 | This work |
| pMES112 | Ap$^r$ (pUC19), 2.0 kb *Pst*I-*Bam*HI fragment from pMES003 | This work |
| pMES115 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 2.0 kb *Hin*dIII-*Bam*HI fragment from pMES112 | This work |
| pMES116 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 4.0 kb *Xho*I-*Bam*HI fragment from pMES110 | This work |
| pMF01 | Ap$^r$ (pUC19), 1.0 kb *Sal*I-*Bam*HI fragment from pMES112 | This work |
| pMF31 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 1.0 kb *Hin*dIII-*Bam*HI fragment from pMF01 | This work |
| pMF3.4 | Ap$^r$ (pUC19), 674 bp *Sph*I-*Bam*HI from pMF01 | This work |
| pMF81 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), 680 bp *Hin*dIII-*Bam*HI from pMF3.4 | This work |
| pMF072 | Ap$^r$ (pGEM-7ZF), *M. xanthus* 513 bp *Xho*I-*Bam*HI generated by PCR | This work |
| pMF72 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), *Xho*I-*Bam*HI fragment from pMF72 | This work |
| pMF062 | Ap$^r$ (pGEM-7ZF), *M. xanthus* 503 bp *Xho*I-*Bam*HI generated by PCR | This work |
| pMF62 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), *Xho*I-*Bam*HI fragment from pMF0062 | This work |
| pMF052 | Ap$^r$ (pGEM-7ZF), *M. xanthus* 493 bp *Xho*I-*Bam*HI generated by PCR | This work |
| pMF52 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), *Xho*I-*Bam*HI fragment from pMF0052 | This work |
| pMF301 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), *M. xanthus* 483 bp *Xho*I-*Bam*HI generated by PCR | This work |
| pMF0301 | Ap$^r$ (pGEM-7ZF), *M. xanthus* 483 bp *Xho*I-*Bam*HI from pMF301 | This work |

Table 2.1 (cont'd)

| | | |
|---|---|---|
| pMF200 | $Ap^r$ $Km^r$ P1-*inc attP* (pREG1727), 452 bp *Rsa*I-*Bam*HI fragment from pMF01 | This work |
| pMF100 | $Ap^r$ $Km^r$ P1-*inc attP* (pREG1727), 521 bp *Hae*II-*Bam*HI fragment from pMF01 | This work |
| pMF0100 | Apr (pGEM-7ZF), 521 bp *Xho*I-*Bam*HI fragment from pMF100 | This work |
| pMES108 | $Ap^r$ $Km^r$ P1-inc (pREG1175), 8.5 kb *Xho*I-*Bam*HI fragment from pMES003 | This work |
| pSH1 | $Ap^r$ $Km^r$ P1-*inc* (pREG429), 4 kb *Cla*I-*Bam*HI fragment from pMES110 | Stacie Hill |
| pLK1 | $Ap^r$ $Km^r$ P1-*inc* (pREG429), *M. xanthus* 17 kb *Eco*RI-*Bam*HI fragment from MSH1 | This work |
| pMF71-1 | $Ap^r$ $Km^r$ P1-*inc attP* (pREG1666), 17 kb *Bam*HI-*Eco*RI from pLK1 | This work |

as an *Eco*RI-*Bam*HI fragment and ligated into *Eco*RI-*Bam*HI-digested pREG1175. The resulting plasmid was digested with *Bam*HI and a 711 bp *Sau*3AI fragment from pRS577 (Simons et al., 1987) containing 4 copies of the T1 terminator of the *E. coli rrnB* operon was inserted. The *Sau*3AI end upstream of the terminators does not generate a *Bam*HI site, but the *Sau*3AI end downstream does, upon insertion into *Bam*HI-digested vector. A plasmid in which the unique remaining *Bam*HI site was proximal to *lacZ* was digested with *Bam*HI and a 49 bp *Bgl*II-*Bam*HI fragment containing multiple restriction sites was inserted. pREG1666 has the polylinker in the orientation that leaves the unique remaining *Bam*HI site proximal to *lacZ*. The polylinker fragment was generated by deleting the *Bgl*II site in the polylinker of pSP72 (Promega), digesting the resulting plasmid with *Xho*I and filling-in the ends with the Klenow fragment of DNA polymerase I, ligating on *Bgl*II linkers (5'-GAAGATCTTC-3'), and digesting with *Bgl*II and *Bam*HI. pREG1666 was used initially to test several *M. xanthus* fragments for promoter activity upon integration at the Mx8 phage attachment site (Mx8 *att*). To identify *M. xanthus* strains containing a single copy of the plasmid integrated at Mx8 *att*, the 3 kbp *Eco*RI-*Bam*HI fragment from pREG1666 containing Mx8 *attP* was labeled and used to probe Southern blots of *Eco*RI-*Xho*I-digested chromosomal DNA as described below. The desired strains show fragments of 2 and 10 kbp in length that hybridize the probe and fail to show a 4 kbp fragment corresponding to unrecombined *attP*. Occasionally, a fragment slightly less than 2 kbp was detected, rather than the expected 2 kbp fragment. We examined the structure of pREG1666 and found that the terminator-containing fragment was slightly larger than expected. We do not know the source of additional DNA. pREG1727 is identical to pREG1666, except the terminator-containing fragment is the expected size. pREG1727 was used subsequently to test *M. xanthus* fragments for promoter activity, but this did not solve the problem of occasional integrants (about 15%) showing a fragment slightly smaller than 2 kbp upon Southern blotting. We suspect that recombination within the

terminator-containing fragment accounts for this observation. We avoided integrants showing a fragment smaller than 2 kbp. In addition, we compared *lacZ* expression from pREG1666 or pREG1727 with no insert of *M. xanthus* DNA or with the 1.0 kbp Ω4403 upstream region inserted, and detected no significant difference between the two vectors.

pMF3.4 was constructed by digesting pMF01 with *Sph*I, which cuts once in the 1.0 kb *M. xanthus* insert and once in the multiple cloning site of pUC19 near the end of the insert more distal from the Ω4403 insertion site, releasing a fragment of about 400 bp. The 3.4 kb vector-containing fragment was gel-purified and ligated, generating pMF3.4.

To construct pMF100, the insert fragment was generated by digesting pMF01 with *Hae*II and making these ends blunt by digesting with mung-bean nuclease (Sambrook et al., 1989). After phenol-chloroform extraction and ethanol precipitation the DNA was digested with *Bam*HI and fractionated on an agarose gel. The supposed 523 bp fragment was purified and directionally cloned into pREG1727 that had been digested with *Hin*dIII, subjected to end-filling using the Klenow fragment of DNA polymerase I (Sambrook et al., 1989), digested with *Bam*HI and gel-purified. The *Hin*dIII-*Hae*II junction in pMF100 was sequenced by first subcloning the insert as a *Xho*I-*Bam*HI fragment into *Xho*I-*Bam*HI-digested pGEM-7Zf, generating pMF0100. The M13 pUC forward sequencing primer (United States Biochemical) was used to sequence this junction as descibed below. The sequence showed that mung-bean nuclease had over-digested the *Hae*II end by two base pairs. Thus, pMF100 contains a 521 bp insert rather than the 523 bp insert predicted by the location of the *Hae*II restriction site.

pMF200 was constructed by directional cloning of the 452 bp *Rsa*I-*Bam*HI fragment from pMF01 into pREG1727 that had a filled-in *Hin*dIII blunt-end and a *Bam*HI cohesive end (prepared as described above for pMF100).

pMF072, 062, and 052 contain PCR-generated fragments of Ω4403 upstream DNA. The regions between -72, -62, and -52 (relative to the putative Ω4403

transcriptional start site) and the *Bam*HI site at the left end of Tn*5 lac* were amplified using pMF01 as template. For this purpose, the downstream primer was the M13 pUC forward sequencing primer (United States Biochemical) that binds in pUC19 downstream of the *Bam*HI site, and the upstream primers were 34-mers that bind between 454 and 429, 444 and 419, 434 and 409 bp upstream of the Ω4403 insertion site. The 5' ends of the upstream primers had a 9 b sequence containing a *Xho*I site that does not bind to the template. The amplified fragments were digested with *Bam*HI and *Xho*I, gel-purified, and ligated into *Xho*I-*Bam*HI digested pGEM-7Zf, resulting in pMF72, 062, and 052. The inserts in these plasmids were sequenced. pMF301 also contains a PCR-generated fragment of Ω4403 upstream DNA. In this case, the upstream primer binds between 424 and 405 bp upstream of the Ω4403 insertion site and the downstream primer was the same as above. The amplified fragment was digested with *Xho*I and *Bam*HI, gel-purified, and ligated into *Xho*I-*Bam*HI-digested pREG1727, generating pMF301. To sequence the insert fragment, the *Xho*I-*Bam*HI fragment from pMF301was first cloned into *Xho*I-*Bam*HI digested pGEM-7Zf, generating pMF0301, and sequencing was performed on this plasmid.

To clone the 17 kb *Eco*RI-*Bam*HI fragment upstream of Ω4403, pSH1 (Table 1) was transduced into DK1622 with selection for kanamycin-resistance (Km$^r$). One of the Km$^r$ transductants, MSH1, had pSH1 integrated at the native site (designated *Ω4403* in Table 1) by homologous recombination, as verified by Southern blot analysis (data not shown). Chromosomal DNA from MSH1 was digested with *Eco*RI, which cuts *M. xanthus* DNA 17 kb upstream of the Ω4403 insertion site and cuts downstream at the distal end of the integrated vector. Religation at low DNA concentration to favor recircularization and transformation into *E. coli* DH5α with selection for Km$^r$ and ampicillin resistance (Ap$^r$) resulted in the cloning of pLK1. To test the 17 kb Ω4403 upstream segment in pLK1 for promoter activity, the plasmid was digested with *Eco*RI and the ends were filled-in

using the Klenow fragment of DNA polymerase I, followed by digestion with *Bam*HI. The 17 kb fragment was gel-purified and ligated to pREG1727 that had a filled-in *Hin*dIII blunt end and a *Bam*HI cohesive end (prepared as described above for pMF100), resulting in pMF71-1.

**DNA sequencing.** DNA fragments to be sequenced were cloned into either pUC19 or pGEM-7Zf and sequenced on both strands using synthetic oligonucleotide primers. Double-stranded sequencing was performed by the Sanger method (Sanger *et al.*, 1977) using the Sequenase kit (United States Biochemical). Ambiguities arising from premature termination were resolved using the protocol of Fawcett and Bartlett (Fawcett & Bartlett, 1990). Briefly, 1 µl of a reaction mixture containing terminal deoxynucleotide transferase (1 µM of each dNTP, pH 7.0, 2 units/µl terminal deoxynucleotide transferase, 1 X sequenase reaction buffer) was added to each of the termination reactions (7 µl total) and incubated at 37°C for 30 min. The reaction was terminated using 4 µl of stop buffer (United States Biochemical). 7-deaza dGTP reaction mixes (United States Biochemical) were used to resolve regions of compression. DNA and protein sequence analyses were performed using the University of Wisconsin Genetics Computer Group software package.

**Construction of *M. xanthus* strains.** Strains containing pREG1666 or pREG1727, or derivatives of these plasmids, integrated at the Mx8 phage attachment site (designated *attB* in Table 1) were constructed by P1 specialized transduction from the *rec*+ *E. coli* strain JM83 into wild-type *M. xanthus* strain DK1622 or the *csgA* mutant strain DK5208 as described previously (Gill et al., 1988a). For each plasmid, three transductants (listed in Table 1), each containing a single copy of the plasmid integrated at Mx8 *att*, were identified using Southern blot analysis (data not shown). Similarly, pMES108 was transduced into DK1622 and three transductants containing a single copy of the plasmid integrated at the native site (designated *Ω4403* in Table 1) by homologous

recombination were identified by Southern blotting (data not shown).

Strain JPB07 was constructed by transducing bacteriophage P1::Tn*5* *lac* (Tc$^r$) into DK4368 with selection for oxytetracyclin resistance (Tc$^r$). Screening for kanamycin-sensitive transductants identified JPB07 in which the Km$^r$ gene was replaced by the Tc$^r$ gene, as verified by Southern blot analysis (Brandner & Kroos). The maximum level of developmental *lacZ* expression from JPB07 (and similar, independent isolates) is consistently about two-fold lower than from DK4368 (Brandner & Kroos), and we do not know the reason for this difference. MLK1-3, -4, and -5 are Km$^r$, Tc$^r$ strains resulting from P1 specialized transduction (Gill et al., 1988a) of pLK1 from *E. coli* JM83 into *M. xanthus* JPB07.

**Southern blot analysis.** *M. xanthus* chromosomal DNA was isolated using the method of Avery and Kaiser (Avery & Kaiser, 1983). Three to 5 μg of DNA was digested with the appropriate restriction enzyme(s) and fractionated on a 0.5% agarose gel. DNA was transferred as described previously (Davis et al., 1980), except nylon membranes were used. Probe synthesis and hybridization of the membrane were carried out using the Genius kit (Boeringer Mannheim) according to the manufacturer's specifications.

**RNA isolation.** Total RNA was purified from DK4368 (which contains the Tn*5* *lac* Ω4403 fusion transcript) or DK4499 (which contains the wild-type Ω4403 transcript) using the protocol of Igo and Losick (Igo & Losick, 1986), but with the following modifications. Cells were harvested from exponentially growing cultures (50 ml of 5 x $10^8$ cells/ml) or from an equivalent number of cells at 24 h of development on TPM agar plates, washed with ice-cold TPM, and resuspended in 2 ml LETS buffer (100 mM LiCl, 10 mM EDTA, 10 mM Tris pH 7.8 with LiCl, 1 % SDS). The nucleic acid was resuspended in 200 μl of diethylpyrocarbonate (DEPC)-treated water (0.1% v/v) and treated with RNase-free DNase (Boeringer Mannheim) to eliminate DNA. DNase was removed by

extracting with phenol-chloroform and precipitating with ethanol. RNA was resuspended in DEPC-treated water and used for further analysis.

**S1-nuclease protection experiments.** Low resolution mapping of the 5' end of the Ω4403-associated transcript was performed as described by Burton et al. (Burton et al., 1983). Hybridization buffer, S1 nuclease mapping buffer, stop solution and formamide loading buffer were prepared as described previously (Sambrook et al., 1989). The mapping strategy was designed to determine whether the start site for transcription was located between the *Bam*HI and*Sal*I sites (0 to 1 kb) or the *Sal*I and*Pst*I sites (1 to 2 kb) upstream of the Ω4403 insertion. Plasmid pMES112 was digested with *Bam*HI or *Sal*I, phosphatase treated, 5' end-labelled with $^{32}$P-γATP and T4 polynucleotide kinase, and used as a probe. Developmental or vegetative RNA (20 to 50 μg) was precipitated with probe (0.5 μg) and the pellet was resuspended in 10 μl of hybridization buffer. After denaturing the nucleic acids at 85°C for 10 min, the samples were incubated at 53°C for 16 h. Unhybridized, single-stranded DNA and RNA was digested for 1 h at 37°C with S1 nuclease (Boehringer Manneheim , 25 to 250 units) in a final volume of 200 μl containing 1X S1 nuclease mapping buffer (Sambrook et al., 1989). The reaction was stopped by the addition of 40 μl of stop solution and was extracted with phenol-chloroform (150 μl). The samples were precipitated with ethanol, resuspended in formamide loading buffer, and the loaded onto a 5% polyacrylamide-8M urea gel. The protected products were separated by electrophoresis at 25 mA and visualized by autoradiography.

For quantitative S1 nuclease mapping, experiments, the procedure of Gilman and Chamberlin (Gilman & Chamberlin, 1983) was used. Yeast tRNA was added to some samples to maintain a constant total amount of input RNA (50 μg). *Bam*HI-digested pMES112 was 5' end-labelled and used as probe (0.5 μg). Hybridization, S1 nuclease digestion, and polyacrylamide gel electrophoresis conditions were identical to those described above. Signals were quantified using a Visage Digital Imager and fell within the

linear response range of the assay as determined with different amounts (0-100 μg) of RNA from DK4368 cells harvested at 18 h of development.

**Primer extension analysis.** Primer extension reactions were performed as described previously (Sambrook et al., 1989) using 20 μg of RNA. The oligonucleotide (5'-CGCCCCCATCAGCAACATCATGCCCACGGA-3') used for the primer extension analysis corresponds to a sequence approximately 100 bp downstream of the putative transcriptional start site mapped by S1 nuclease protection. The primer was end-labeled using T4 polynucleotide kinase and $^{32}$P-γATP, and purified as described previously (Ausubel et al., 1989). Primer extension products were electrophoresed at 60 W on a sequencing gel (6 % polyacrylamide, 5 M urea) in parallel with sequencing reactions performed with the same end-labelled primer.

## Results

**Cloning DNA upstream of Ω4403 and testing it for promoter activity.** To clone the putative promoter region upstream of the developmentally regulated Tn*5 lac* insertion Ω4403, we took advantage of a *Xho*I restriction site approximately 8 kb upstream of the site of insertion Ω4403 (Kroos et al., 1986) and a *Xho*I site in Tn*5 lac* about 9 kb from the left end (Figure 2.1). Since the *Xho*I site in Tn*5 lac* is downstream of the *aphII* gene (encoding aminoglycoside phosphotransferase, which confers kanamycin resistance), *Xho*I digestion of chromosomal DNA from *M. xanthus* containing Tn*5 lac* Ω4403 should yield a 17 kb fragment able to confer kanamycin resistance when cloned in *E. coli*. We isolated DNA from *M. xanthus* strain DK4368 containing Tn*5 lac* Ω4403, digested it with *Xho*I, ligated the fragments with *Xho*I-digested pGEM-7Zf, and transformed the mixture into *E. coli* strain DH5α, selecting for both ampicillin resistance of the vector and

kanamycin resistance of the desired insert. One transformant with a plasmid bearing an insert of the expected size was characterized further. Restriction mapping of the plasmid, pMES003, showed the expected patterns based on restriction sites in DNA upstream of Ω4403 that had been mapped by Southern blotting (Kroos et al., 1986) and on known restriction sites in Tn*5 lac* and the vector (data not shown). Figure 2.1 shows a restriction map of DNA upstream of Ω4403 based on these results.

To test the DNA upstream of Ω4403 for promoter activity, the *Xho*I-*Bam*HI restriction fragment from pMES003 that includes 8.5 kb of *M. xanthus* DNA and about 50 bp of the left end of Tn*5 lac* (Figure 2.1) was subcloned into *Xho*I-*Bam*HI-digested pREG1666 (Figure 2.2) to construct pMES004. Because the multiple cloning site in pREG1666 is upstream of the same *lacZ*-containing segment found in Tn*5 lac*, pMES004 contains Ω4403 upstream DNA fused to *lacZ* in exactly the same way as in the chromosome of *M. xanthus* strain DK4368, which has Tn*5 lac* inserted at site Ω4403

**Figure 2.1.** Physical map of the Ω4403 insertion region and summary of deletions tested for promoter activity. The top part shows restriction sites in Tn*5 lac* and the adjacent *M. xanthus* chromosome that were used in cloning experiments. The distance of restriction sites from the Tn*5 lac* Ω4403 insertion is give in kb. B, *Bam*HI; C, *Cla*I; E, *Eco*RI; S, *Sal*I; Sp, *Sph*I. The bottom part shows different segments of Ω4403 upstream DNA that were fused to *lacZ* to permit testing for promoter activity as described in the Materials and Methods. The maximum ß-galactosidase specific activity during a 48 h developmental time course of wild-type *M. xanthus* strain DK1622 derivatives containing single copies of the plasmids at Mx8 *att* are given as a percentage of the maximum activity observed for strain DK4368 containing Tn*5 lac* Ω4403. In each case, the maximum activity of three independent transductants was measured and the average is given.

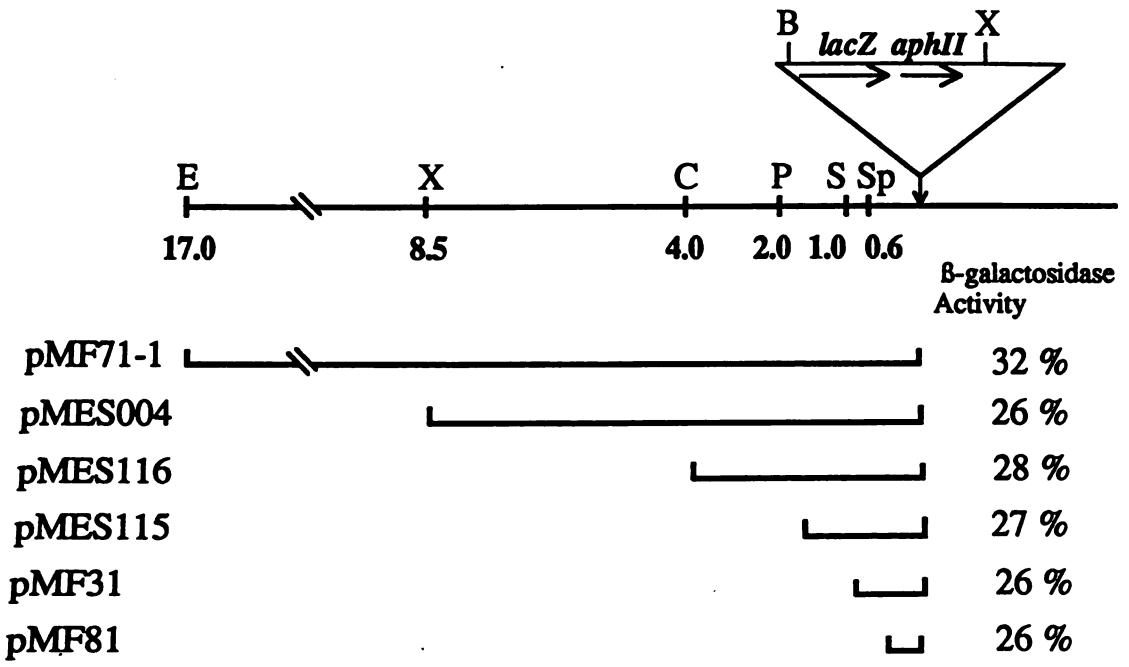
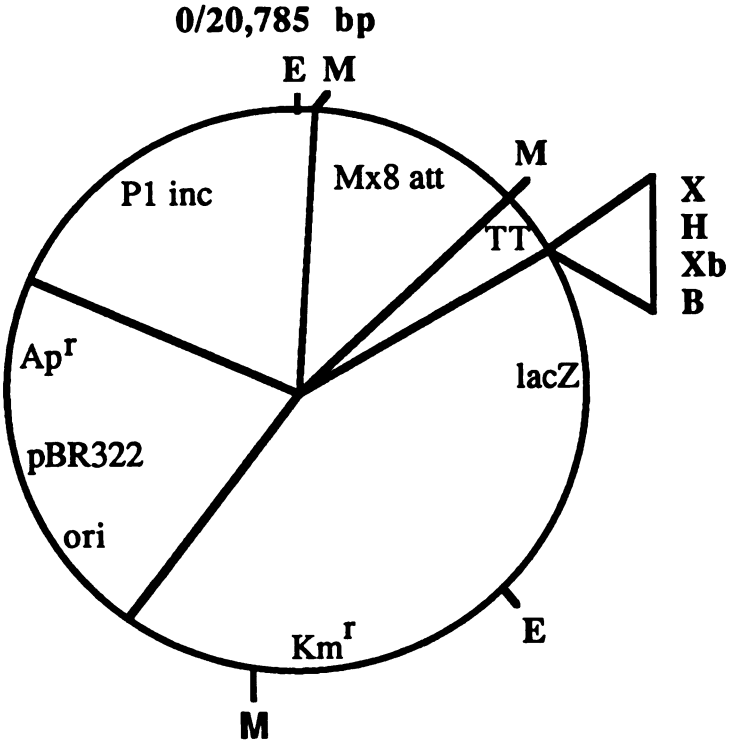

B
lacZ aphII
X
E
X
C
P
S
Sp
17.0
8.5
4.0
2.0
1.0
0.6
ß-galactosidase Activity
pMF71-1
32 %
pMES004
26 %
pMES116
28 %
pMES115
27 %
pMF31
26 %
pMF81
26 %

**Figure 2.2.** Map of the plasmid used to test for promoter activity. Unique cloning sites in this vector include B, *Bam*HI; H, *Hin*dIII; Xb, *Xba*I; X, *Xho*I. Additional restriction sites indicated are E, *Eco*RI and M, *Sma*I. This plasmid contains the Mx8 phage attachment site, Mx8 *att*, E.*coli* transcriptional terminators, TT, the promoterless *lacZ* gene, the *aphII* gene (conferring $Km^r$) with its own promoter, part of pBR322 including the origin of replication and the *bla* gene (conferring $Ap^r$), and a P1 incompatibility fragment, P1 *inc*.

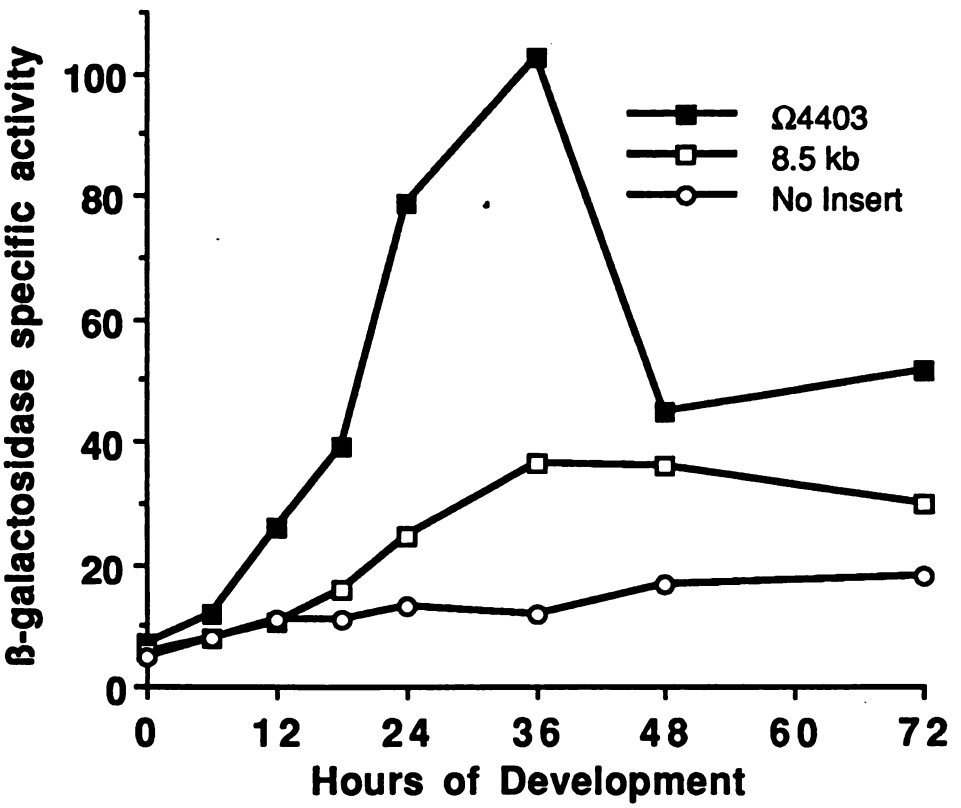

0/20,785 bp
E M
M
X
H
Xb
B
P1 inc
Mx8 att
TT
Ap^r
lacZ
pBR322
ori
Km^r
E
M

(Kroos et al., 1986). When pMES004 was transduced from *E. coli* strain JM83 into the wild-type *M. xanthus* strain DK1622 using bacteriophage P1 specialized transduction, it integrated into the chromosome efficiently at the Mx8 prophage attachment site (Mx8 *att*). Transductants containing a single copy of the pMES004 integrated at Mx8 *att* were identified by Southern blot hybridization (data not shown, see the Materials and Methods). Several of these transductants were assayed for ß-galactosidase activity during development. In parallel, developmental *lacZ* expression of *M. xanthus* strain DK4368 containing Tn*5* *lac* Ω4403 was measured. Also, as a negative control, expression of transductants containing a single copy of pREG1666 (with no *M. xanthus* DNA insert) integrated at Mx8 *att*, was examined. Figure 2.3 shows that transductants containing pMES004 integrated at Mx8 *att* expressed *lacZ* with similar timing as the strain containing Tn*5* *lac* Ω4403, but reached only about 25% of the maximum level (after subtracting the background expression observed for the negative control). These results indicate that the 8.5 kb Ω4403 upstream DNA has a promoter that is able to drive development-specific expression of a *lacZ* reporter gene.

To determine whether a higher level of promoter activity could be observed by cloning additional upstream DNA, we cloned 17 kb of DNA upstream of Ω4403 (see the Materials and Methods) and tested it for promoter activity as described above. Like the 8.5 kb segment, the 17 kb segment directed only about 25% of the maximum level of developmental *lacZ* expression seen with *M. xanthus* strain DK4368 containing Tn*5* *lac* Ω4403 (Figure 2.1). This result suggests that the lower activity of fusions integrated at Mx8 *att* as compared to the fusion created by insertion of Tn*5* *lac* at site Ω4403 is unlikely to be due to insufficient upstream DNA.

To determine the approximate location of the promoter in the 8.5 kb Ω4403 upstream DNA, smaller fragments were tested for promoter activity after fusion to *lacZ* and integration at Mx8 *att* as described above. Figure 2.1 shows that promoter activity

**Figure 2.3.** Expression of developmental *lacZ* under the control of the Ω4403 promoter. Developmental *lacZ* expression was determined as described previously (Kroos, et al., 1986) for *M. xamthus* strain DK4368 containing Tn*5 lac* Ω4403 (■) and for for three independently isolated transductants containing a single copy of the 8.5 kb Ω4403 upstream DNA fused to *lacZ* and integrated at Mx8 *att* (strains MES005, 008, and 012 in Table 1) (❑) or the vector with no insert DNA at Mx8 *att* (JW103, 104, and 107) (○). The average ß-galactosidase activity from three determinations for DK4368 or from one determination for each of three independent transductants is plotted. The units of ß-galactosidase specific activity are nanomoles of ONP (ortho nitrophenyl phosphate) minutes$^{-1}$ milligram of protein$^{-1}$.

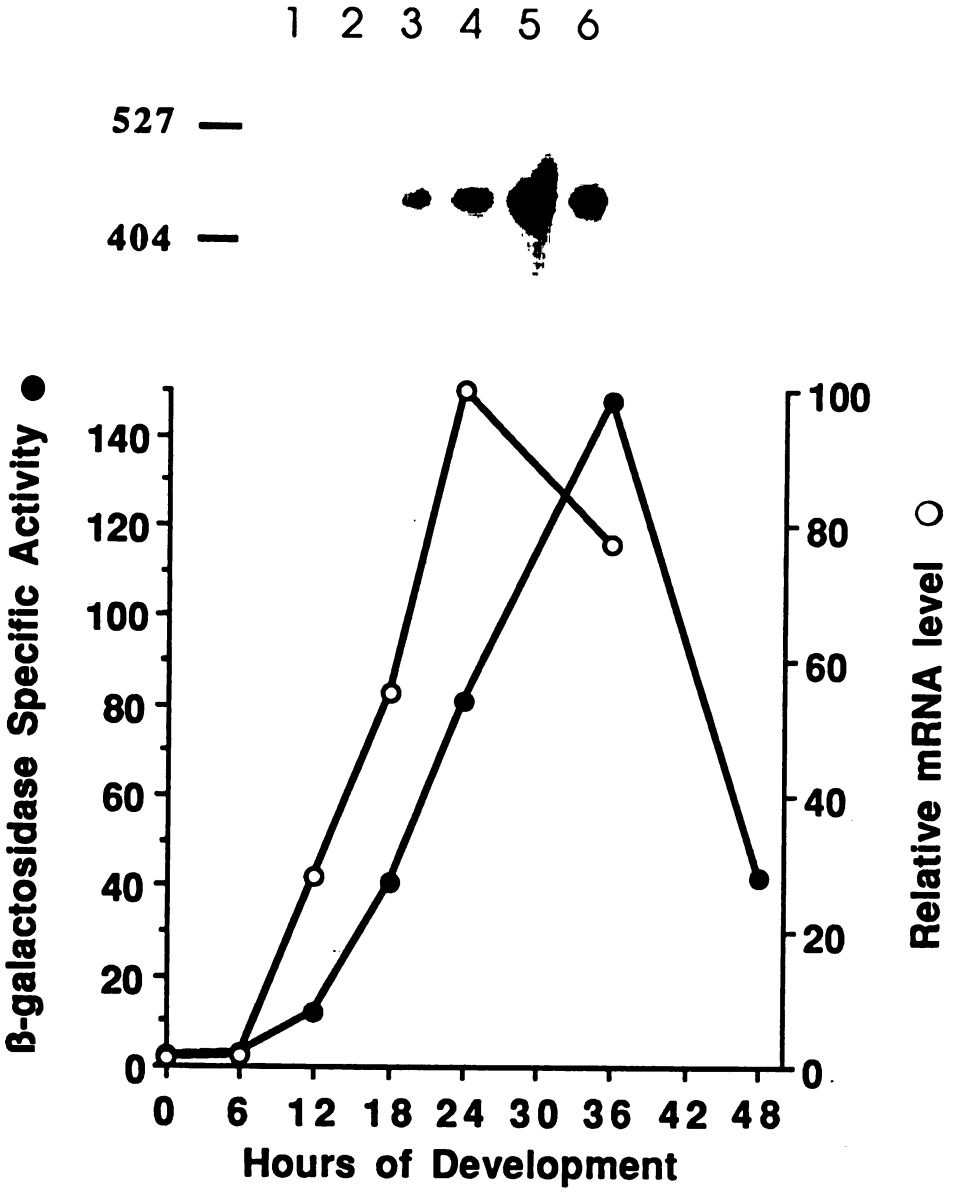

ß-galactosidase specific activity
100
80
60
40
20
0
0
12
24
36
48
60
72
Hours of Development
Ω4403
8.5 kb
No Insert

localized to 0.6 kb of DNA immediately upstream of Ω4403. No significant difference was observed in the timing or level of developmental *lacZ* expression for any of the fragments that showed promoter activity.

**Accumulation of the Tn*5* *lac* Ω4403 fusion mRNA during development.** Localization of promoter activity within 0.6 kb upstream of Ω4403 indicated that an RNA 5' end should map within this region. A probe labeled at the *Bam*HI site near the left end of Tn*5* *lac* (Figure 2.1) and including about 2 kb of *M. xanthus* DNA upstream of Ω4403, was hybridized to RNA from *M. xanthus* strain DK4368, which contains Tn*5* *lac* Ω4403 (Kroos et al., 1986), and subjected to S1 nuclease protection analysis. RNA from 24 h developing cells, but not from growing cells, protected a fragment of about 430 b in length. We used this assay to quantitate the relative amount of fusion mRNA from Tn*5* *lac* Ω4403 in strain DK4368 during development. Figure 2.4 shows that the fusion mRNA was first detected at 12 h of development and reached a maximum level at 24 h. ß-galactosidase activity was measured in samples collected in the same experiment. Figure 2.4 shows that accumulation of ß-galactosidase activity lagged slightly behind accumulation of fusion mRNA. We conclude that Tn*5* *lac* Ω4403 identifies a transcriptional unit with a putative start site approximately 430 b upstream of the *Bam*HI site in Tn*5* *lac*. The increase in the level of fusion mRNA during development appears to account for the observed developmentally regulated ß-galactosidase activity of strain DK4368. The increase in fusion mRNA probably results from activation of a developmentally regulated promoter, although we cannot exclude the possibility that the fusion mRNA is stabilized in a developmentally regulated fashion.

We were unable to detect any discrete species of fusion mRNA by Northern blotting of RNA from developing DK4368 cells (Appendix 1). Apparently, the large fusion mRNA is susceptible to degradation. The ability to map the 5' end of the fusion mRNA suggests that the S1 nuclease assay is more sensitive and/or may involve

**Figure 2.4.** Levels of fusion mRNA and ß-galactosidase specific activity fron Tn*5 lac* Ω4403 during development. The top part shows quantitative S1-nuclease protection analysis of developmental RNA from *M. xanthus* strain DK4368, performed as described in the Materials and Methods with 50 µg of RNA harvested from cells at 0, 6, 12, 18, 24, and 36 h of development in lanes 1-6, respectively. The numbers on the left indicate the length of markers in bases. The lower panel shows quantitation of the S1-nuclease protected products in the upper panel (○), as well as ß-galactosidase specific activity (●) of samples harvested in the same experiment.


1 2 3 4 5 6
527
404
ß-galactosidase Specific Activity ●
Relative mRNA level ○
0
20
40
60
80
100
120
140
0 6 12 18 24 30 36 42 48
Hours of Development

hybridization of probe to partial breakdown products of the fusion mRNA. Using *M. xanthus* cells that do not contain Tn*5 lac* Ω4403, and DNA immediately upstream of Ω4403 as the labeled probe, we detected a 1.5 kb transcript by Northern blotting of RNA isolated from 24 h developing cells, but not from growing cells (Appendix 1). Since the native Ω4403 mRNA has the same 5' end as the Tn*5 lac* Ω4403 fusion mRNA (see below), the 1.5 kb size of the native transcript suggests that transcription terminates about 1.1 kb downstream of the Ω4403 insertion site.

**DNA sequence of the Ω4403 upstream region.** Figure 2.5 shows the nucleotide sequence of 616 bp of *M. xanthus* DNA immediately upstream of the Ω4403 insertion site. An open reading frame (ORF) beginning with ATG at position 376 is preceded 6 nucleotides upstream by the sequence 5'-AGCAGG-3', which is complementary (except for 1 mismatch) to a sequence near the 3' end of *M. xanthus* 16S rRNA, suggesting it could serve as a ribosome binding site. The ORF exhibits codon preference typical of *M. xanthus* genes (Hagen & Shimkets, 1990), as well as a strong bias towards usage of guanine or cytosine at the third position of codons, which is also typical of GC-rich organisms like *M. xanthus* (Shimkets, 1990). The ORF remains uninterrupted for at least 80 amino acids, extending to the Ω4403 insertion site at the end of the sequenced region. The deduced amino acid sequence of the ORF shows 38% amino acid identity and 73% amino acid similarity over a 63 amino acid stretch to aqualysin I (Figure 2.6), a secreted protease of *Thermus aquaticus* (Terada et al., 1990). Both the ORF and aqualysin I exhibit a sequence highly similar to the sequence YIVGFK, which is conserved in subtilisin-type proteases and influences processing of precursor protein to the active protease (Jacobs et al., 1985; Lee et al., 1991). Based on these observations, we speculate that Tn*5 lac* Ω4403 disrupts the first (and perhaps only) gene in a developmentally regulated transcriptional unit, and that this gene encodes a subtilisin-type serine protease. The function of the putative protease remains unclear because *M. xanthus*

**Figure 2.5.** Nucleotide sequence of 616 bp directly upstream of Tn*5* *lac* Ω4403. The transcriptional start site is indicated by an arrow and the primer used for primer extension analysis is underlined. The hexanucleotide sequences in the -10 and -35 regions that resemble sequences recognized by *E. coli* $\sigma^{70}$ RNA polymerase are indicated by bold lettering. The putative start codon and a potential ribosome binding site are boxed in the deduced amino acid sequence of the Ω4403 partial open reading frame is shown below the nucleotide sequence. Restriction sites used to generate some deletions in the Ω4403 promoter region are shown. The numbers above the nucleotide sequence indicate the 5' end of deletions that were tested for promoter activity.

*Sph*I
GCATGCCCACCTGGTGAACGCAGCACGAGAGATGGGCATCAACGACATCCTGTCACGCCC

CGGGAAACGTCCACTGGAAGCAGTCTTGTACGAATGGGGAACCACCATTCGTCCTTGGTA
*Hae*II-80 -72 -62
GGTCATGGCCAGACATCGCGCCTTGAAGCGCCATGGCATGTTCAATCACGGACCGCCGTC
-52 -42 -11 +1
TCATCCCTCCGGG**TTGATT**CATGAATAAGCCGTTTTTGA**TGTACA**CCCGTTTTACCCATC

ACAACCATTTTCCTCTGGGTAAAACATGAATCGACACACACTGAAGCTGTTCGTCTCCGT

GGGCATGATGTTGCTGATGGGGGCGTGCGCATCGCGCAGGTGCCCTGCAAGGACCGCACG

GCCAGCAGGCCCACGATGGGCCAGAAGCAGACCGGCAAGTTCATCACTGTCCGAAAGAAG
M G Q K Q T G K F I T V R K K

ATTCCGGGTGAATACATTGTCGTCCTGAAGTCGCCCGCACAAAGCCTGGAACAGGTCGAG
I P G E Y I V V L K S P A Q S L E Q V E

GTCCAGCAAGCCACGACGAGCCTCATCACGGCTTACGGTGGCACCGCATTCGCGATGTAT
V Q Q A T T S L I T A Y G G T A F A M Y

GAGAATGCGTTGCGTGGTTTCGCGGCCAAGATGACGGAAGCCCAGGCGCGGGCCATGGCG
E N A L R G F A A K M T E A Q A R A M A

AACGACCCCGCAGCTG
N D P A A

**Figure 2.6.** Alignment of the amino acid sequence of the Ω4403 partial ORF and the N-terminal region of aqualysin I. Identical amino acids are indicated by a vertical line and similar amino acids are indicated by two dots. The conserved amino acid sequence in subtilisin-type proteases (Jacobs, et al., 1985) is shown in bold lettering above the corresponding regions.

```
                                        YIVGFK
Ω4403          1  MGQKQTGKFITVRKKIPGEYIVVLK- SPAQSLEQVEVQQATTSLI
                                           III  IIII : I  :  : I II  I  : :       : :I
Aqualysin I   36  WPKEAPVYGLDDPEAIPGRYIVVFKKGKGQSLLQGGI TT LQARLA

Ω4403         45  TAYGGTAFAMYENALRGFAAKMTEAQARAMANDPAA
                   : :  I  : :  :  I : : II : II I I: I :  : :   I :  : : I
Aqualysin I   81  PQ-GVVVTQA YTGALQGFAAEMAPQALEAFRQSPDVE
```

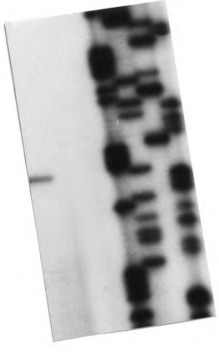
cells bearing Tn*5 lac* Ω4403 show no discernible growth or developmental defects (Kroos et al., 1986), (see below).

**Precise mapping of the Ω4403 mRNA 5' end.** Knowing the approximate location of the 5' end of the Tn*5 lac* Ω4403 fusion mRNA (Figure 2. 4) and the sequence of Ω4403 upstream DNA (Figure 2.5), we designed a primer to precisely map the 5' end of Ω4403 mRNA using primer extension analysis. The position of the primer is shown in Figure 2.5. Figure 2.7 shows that the 5' end of the native Ω4403 mRNA mapped to an adenine nucleotide 382 bp upstream of the Ω4403 insertion site. The 5' end of the Tn*5 lac* Ω4403 fusion mRNA also mapped to the same position (data not shown). This is about 430 bp upstream of the *Bam*HI site near the left end of Tn*5 lac* (Figure 2.1), in good agreement with the S1 nuclease protection result (Figure 2.4). The position of the putative transcriptional start site is shown in Figure 2.5. Inspection of the DNA sequence of the putative promoter region revealed no striking similarity to known *M. xanthus* promoters. The hexanucleotide sequences 5'-TTGATT-3' and 5'-TGTACA-3' in the -35 and -10 regions, respectively, bear some resemblance to the 5'-TTGACA-3' and 5'-TATAAT-3' sequences that are conserved in the corresponding regions of *E. coli* promoters transcribed by $\sigma^{70}$ RNA polymerase (Lisser & Margalit, 1993).

**Further deletion analysis of the Ω4403 promoter.** We tested smaller fragments of Ω4403 upstream DNA for promoter activity *in vivo* to determine whether a promoter exists at the location predicted by the 5' end mapping and, if so, to determine the extent of sequences required for promoter activity. Figure 2.5 shows the positions of 5' deletions that were generated by digestion with restriction enzymes or by PCR. The 3' end of each fragment was the same, being produced by *Bam*HI digestion at the site near the left end of Tn*5 lac* (Figure 2.1). Each fragment was inserted in the proper orientation into the vector shown in Fig. 2, and *M. xanthus* strains with a single copy of the plasmid at Mx8 *att* were tested for developmental *lacZ* expression. Eighty bp of DNA upstream of the putative

**Figure 2.7.** Primer extension analysis of Ω4403 mRNA. RNA was isolated from *M. xanthus* strain DK4499 (which makes the native Ω4403 mRNA) and subjected to primer extension analysis as described in the Materials and Methods. The same end-labelled primer was also used for sequencing of Ω4403 upstream DNA. A portion of the DNA sequence is indicated at the right. The starred adenine nucleotide indicates the putative transcriptional start site that was observed with RNA isolated from 24 h developing cells (lane 1), but not with RNA from vegetatively growing cells (lane 2).

1 2 A T C G

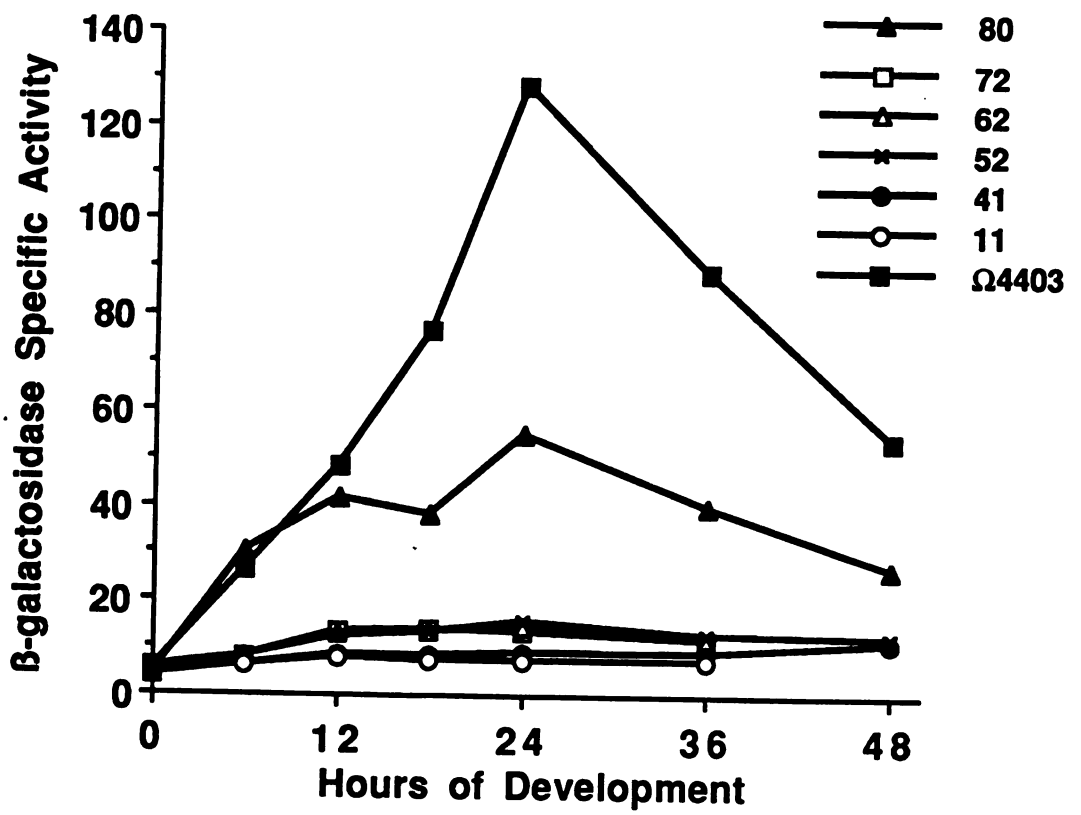


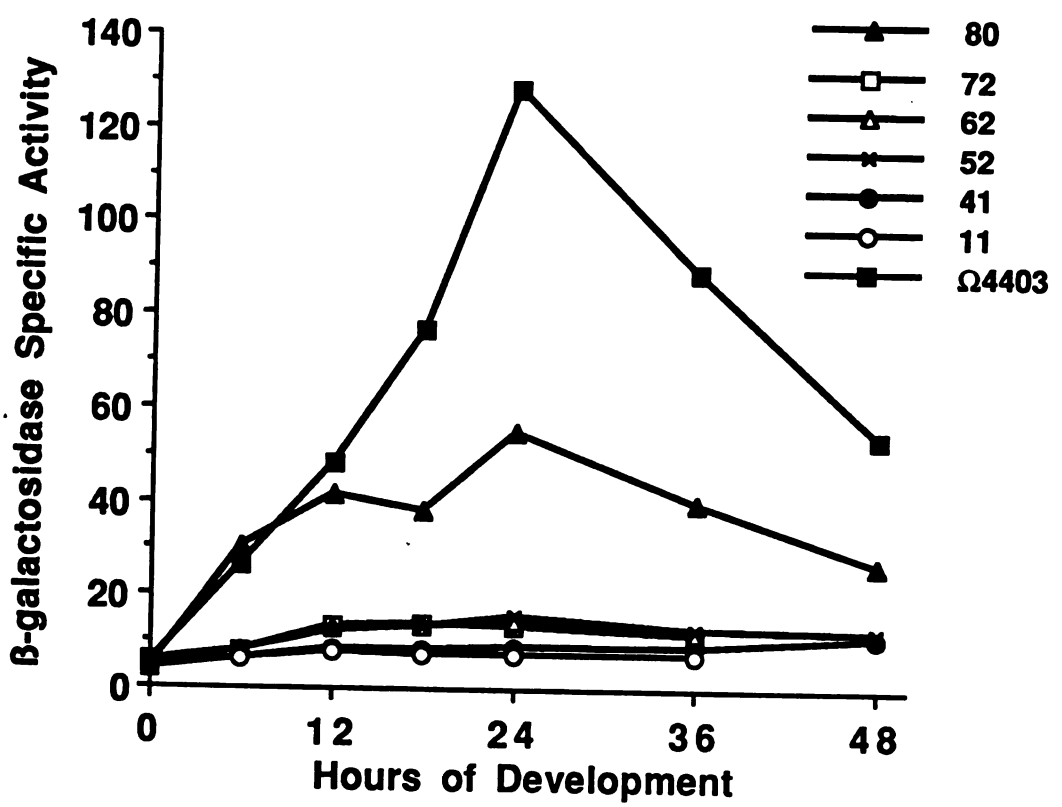
transcriptional start site was sufficient to promote *lacZ* expression with similar developmental kinetics and amount as observed for larger upstream segments (Figure 2.8). Developmental *lacZ* expression was abruptly abolished upon deletion to -72 or farther 3' (Figure 2.8). These results show that a promoter does exist in the predicted region, that 80 bp of DNA upstream of the transcriptional start site is sufficient for developmentally regulated expression of *lacZ* comparable to that observed with much larger upstream segments, and that DNA between -80 and -72 is critical for promoter activity.

**Possible explanations for the low level of expression from fusions integrated at Mx8 *att*.** Segments of Ω4403 upstream DNA fused to *lacZ* and integrated at Mx8 *att* promote only about 25% of the maximum level of developmental *lacZ* expression seen with *M. xanthus* strain DK4368 which contains Tn*5* *lac* Ω4403, regardless of whether the segment ends 80 bp upstream of the Ω4403 transcriptional start site (Figure 2.8) or much farther (up to 17 kb) upstream (Figures 2.1 and 2.3) and despite the identical nature of the fusion junctions. It seemed unlikely that the difference was due to a promoter or a promoter element (e.g., an enhancer) located farther upstream. A promoter element could be located downstream of the Ω4403 insertion site; however, this explanation seemed unlikely for two reasons. First, the downstream element would have to function over a distance of at least 12 kb (the size of Tn*5* *lac*) in order to account for the higher level of expression seen with the strain containing Tn*5* *lac* Ω4403. Second, the distance between the Ω4403 insertion site and the Mx8 *att* site is estimated to be 5-25 kb (Shimkets,), so the hypothetical downstream element would have to be incapable of functioning over this distance in order to account for the low level of expression from fusions integrated at *att*. These considerations led us to test other possible explanations.

A simple explanation for the low level of expression from fusions integrated at Mx8 *att* would be DNA instability. If the plasmid integrated at Mx8 *att* were lost or rearranged at a high frequency during development, it could explain the 75% reduction in *lacZ*

**Figure 2.8.** Deletion analysis of the Ω4403 promoter. Developmental *lacZ* expression was determined as described previously (Kroos, et al., 1986) for two or three independently isolated transductants containing a single copy of Ω4403 upstream DNA fused to *lacZ* and integrated at Mx8 *att* (see table 1 for strain designations). This 5' deletion series included 80 (▲), 72 (❑), 62 (Δ), 52 (✖), 42 (●), or 11 (○) bp of DNA upstream of the putative transcriptional start site. Points are the average ß-galactosidase activity for transductants of the same type. ß-galactosidase activity in DK4368 cells containing Tn*5* *lac* Ω4403 is also plotted (■). The units of enzyme activity are given in the Figure 2.3 legend.


ß-galactosidase Specific Activity
140
120
100
80
60
40
20
0
0
12
24
36
48
Hours of Development
80
72
62
52
41
11
Ω4403


ß-galactosidase Specific Actlvity
140
120
100
80
60
40
20
0
0
12
24
36
48
Hours of Development
80
72
62
52
41
11
Ω4403

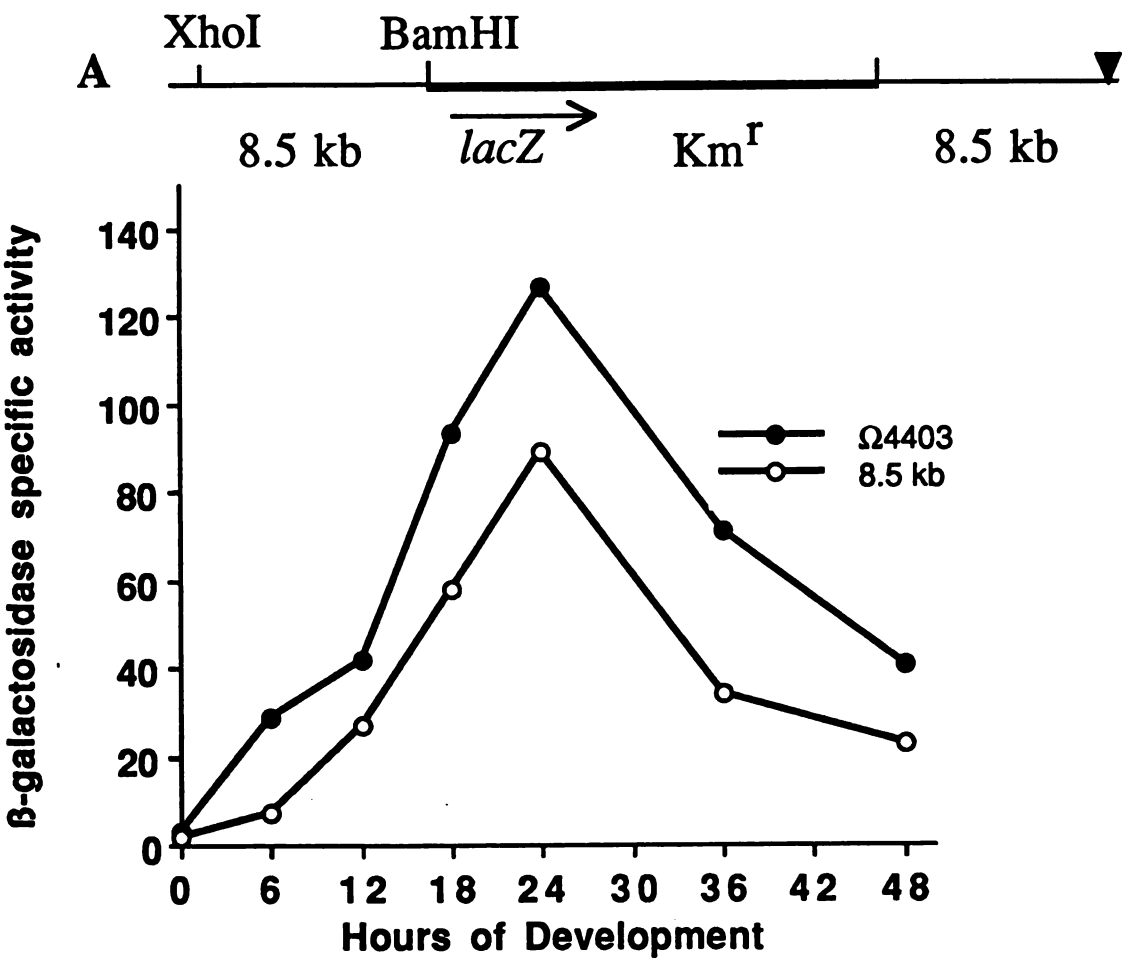
expression as compared to cells containing Tn*5* *lac* Ω4403. We isolated DNA from *M. xanthus* strain MMF31-7 which contains 1 kb of Ω4403 upstream DNA fused to *lacZ*, integrated at Mx8 *att*. DNA from cells grown in the presence of kanamycin (to maintain selection for the integrated plasmid) or from cells developing for 24 h in the absence of antibiotic was digested with *Xho*I and *Eco*RI, then subjected to Southern blot hybridization using a probe that detects fragments of 10 kb and 2 kb if the plasmid is integrated at Mx8 *att* (see the Materials and Methods). Equal amount of DNA from growing or developing cells resulted in hybridization signals of similar intensity at 10 kb and 2 kb, and fragments of other sizes were not detected (Appendix 2). These results show that a plasmid with Ω4403 upstream DNA integrated at Mx8 *att* is neither lost nor rearranged in a substantial proportion of developing cells. Also, after 3 days of development, all 500 heat- and sonication-resistant MMF31-7 spores we tested produced colonies of kanamycin-resistant cells.

Another simple explanation for the low level of expression from fusions integrated at Mx8 *att* would be that the integrated DNA hinders development. We examined fruiting body formation microscopically and measured the number of heat- and sonication-resistant spores produced (Kroos & Kaiser, 1987) from three strains: the wild-type strain DK1622, strain DK4368 containing Tn*5* *lac* Ω4403, and strain MMF31-7 containing 1 kb of Ω4403 upstream DNA fused to *lacZ* and integrated at Mx8 *att*. No significant differences were observed among the three strains (Appendix 3) indicating that a plasmid with Ω4403 upstream DNA integrated at Mx8 *att* does not inhibit development and supporting the previous finding that Tn*5* *lac* Ω4403 does not cause a developmental defect.

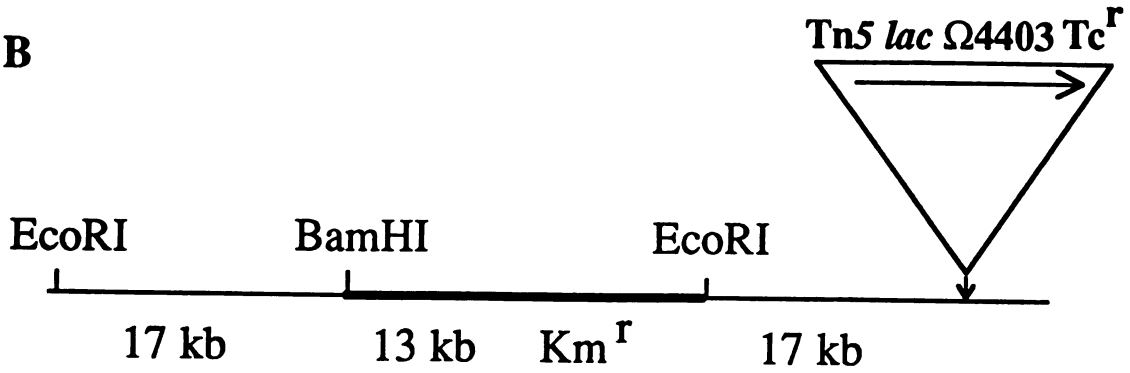
A third possible explanation for the low level of expression from fusions integrated at Mx8 *att* as compared to expression from Tn*5* *lac* Ω4403 would be that the transposon disrupts a gene whose product negatively regulates its own transcription. According to this model, expression from fusions integrated at Mx8 *att* is negatively regulated, directly or

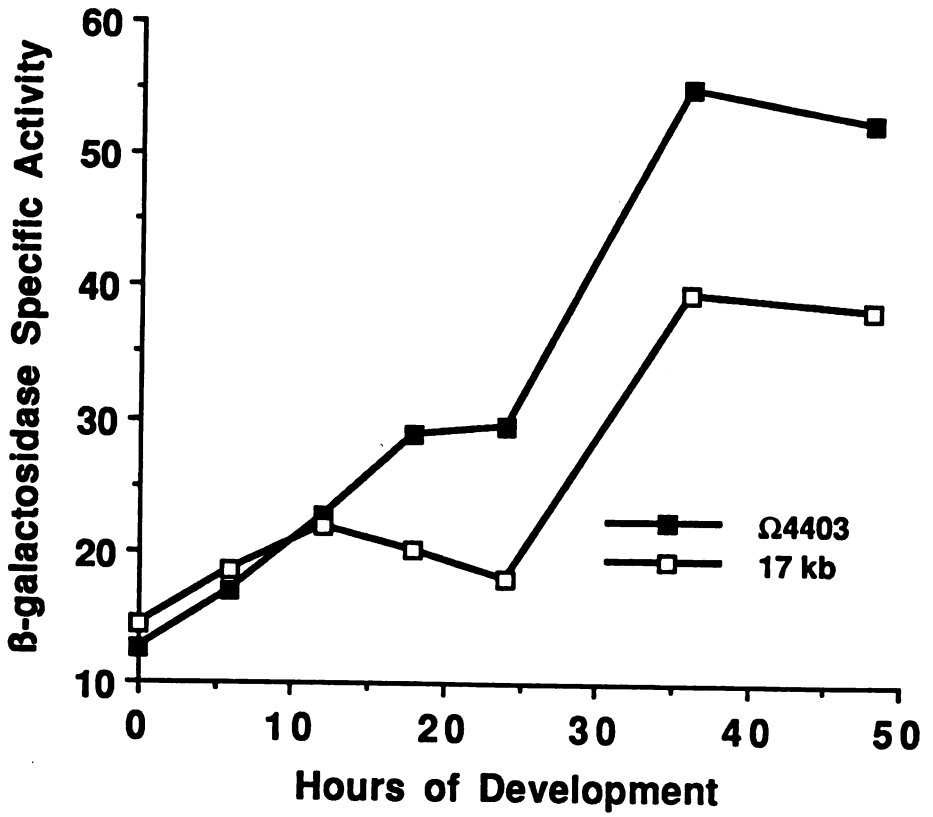
indirectly, by the protein produced by the intact gene at site Ω4403. Thony-Meyer and Kaiser (Kim & Kaiser, 1990a) observed negative autoregulation of *devR-lacZ* by a tandem copy of the intact *devRS* locus. We subcloned 8.5 kb of Ω4403 upstream DNA into a plasmid (pREG1175) (Gill & Bornemann, 1988b) much like the one shown in Figure 2.2, except lacking the Mx8 *attP* portion. Rather than integrating at Mx8 *att* upon transduction into the wild-type *M. xanthus* strain DK1622, the plasmid recombined into the chromosome via a single crossover at the homologous site, resulting in the structure depicted at the top of Figure 2.9A, which was verified by Southern blotting (data not shown). Strains with this structure have all the normal upstream sequence fused to *lacZ*, followed by 8.5 kb of Ω4403 upstream DNA fused to an intact copy of the Ω4403 gene. These strains expressed *lacZ* at about 70% the maximum level observed for strain DK4368 containing Tn*5* *lac* Ω4403 (Figure 2.9A). To determine whether the intact copy of the Ω4403 gene in tandem caused the 30% reduction in developmental *lacZ* expression, we constructed strains with the structure depicted at the top of Figure 2.9B. In this case, a plasmid (pLK1) with 17 kb of Ω4403 upstream DNA, but devoid of Mx8 *attP* and *lacZ*, was transduced into *M. xanthus* strain JPB07 containing Tn*5* *lac* Ω4403-Tc$^r$ (in which the kanamycin resistance gene was replaced with a tetracycline resistance gene to permit selection for the incoming plasmid). Transductants expressed developmental *lacZ* at about 70% the maximum level observed for strain JPB07 containing Tn*5* *lac* Ω4403-Tc$^r$ (Figure 2.9B), despite the lack of an intact Ω4403 gene. Hence, the Ω4403 gene product does not appear to be autoregulatory. The results shown in Figure 2.9 for two different types of strains with *lacZ* fusions integrated by homologous recombination at the native Ω4403 chromosomal site suggest that having two copies of Ω4403 upstream DNA titrates a limiting factor for Ω4403 expression. Because strains with fusions integrated at Mx8 *att* contain two copies of Ω4403 upstream DNA, a titration effect can partially account for the low level of expression observed from these strains as compared to the strain containing

**Figure 2.9.** Developmental *lacZ* expression from fusions at the native Ω4403 position in the chromosome. The arrangement of *M. xanthus* DNA (thin line) and plasmid vector DNA (thick line) is shown at the top in each panel. In panel A, the triangle indicates the site of the Ω4403 insertion as a reference point (Tn*5* *lac* Ω4403 is not present). The graph shows the average ß-galactosidase activity of three independently isolated transductants (strains MES014, 016, and 017) with this structure (○) and the activity of strain DK4368 containing Tn*5* *lac* Ω4403-Km$^r$ (●). In panel B, the triangle shows the position of Tn*5* *lac* Ω4403-Tc$^r$. The graph shows the average ß-galactosidase activity of three strains (MLK1-3, -4, and -5) with this structure (❑) and the activity of strain JPB07 containing Tn*5* *lac* Ω4403-Tc$^r$ (■). The units of enzyme activity are given in the Figure 2.3 legend.


A
XhoI
BamHI
8.5 kb
lacZ
Km^r
8.5 kb
ß-galactosidase specific activity
140
120
100
80
60
40
20
0
0
6
12
18
24
30
36
42
48
Hours of Development
Ω4403
8.5 kb

B

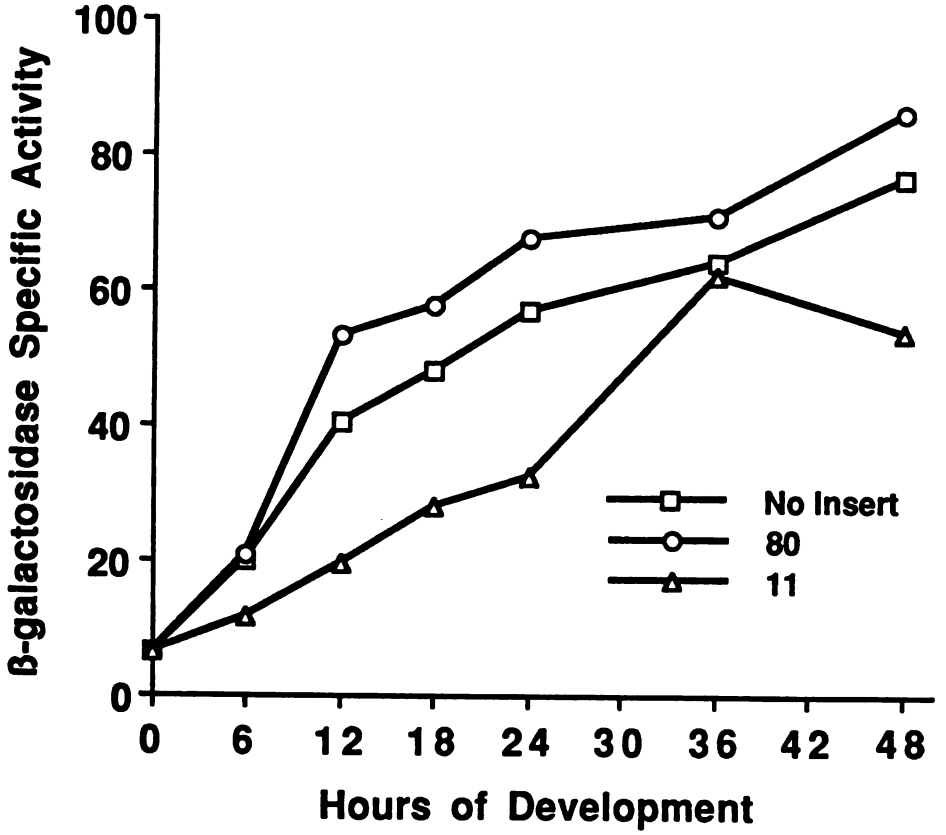

Tn5 lac Ω4403 Tc^r
EcoRI
BamHI
EcoRI
17 kb
13 kb
Km^r
17 kb

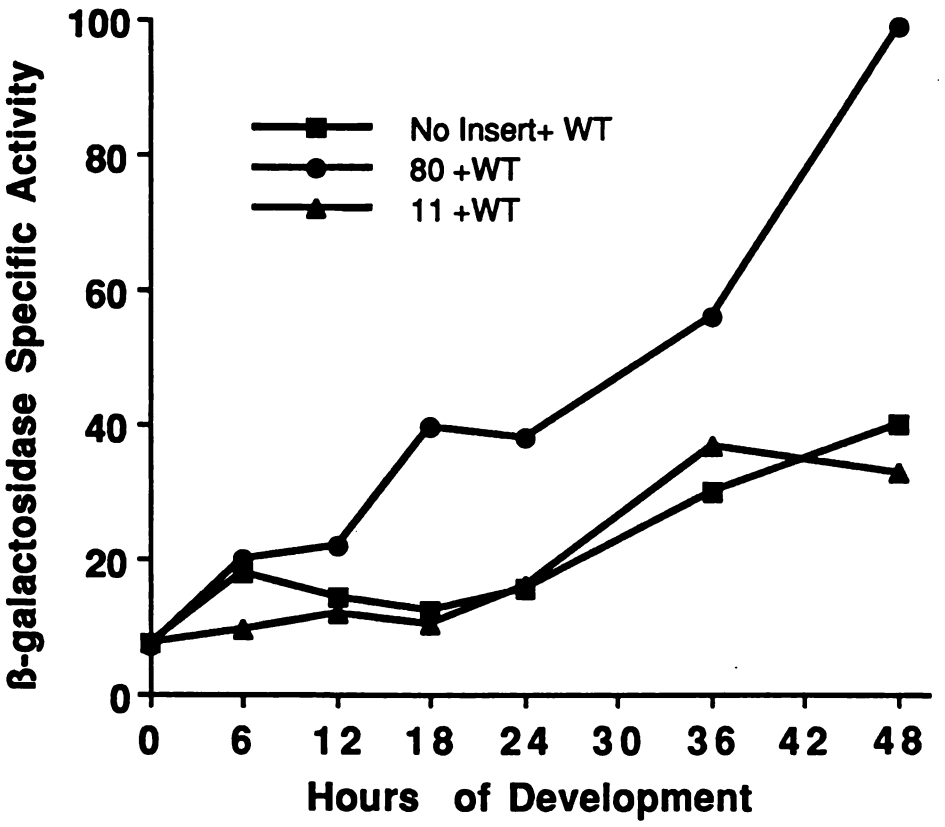

ß-galactosidase Specific Activity
60
50
40
30
20
10
0
10
20
30
40
50
Hours of Development
Ω4403
17 kb

Tn*5* *lac* Ω4403. Even so, expression from fusions integrated by homologous recombination at the native site was about 70% the maximum level observed for strains containing the corresponding Tn*5* *lac* insertion (Figure 2.9), whereas expression from fusions integrated at Mx8 *att* was only about 25% (Figures 2.1, 2.3, and 2.8). The remaining difference may be due to an effect of chromosomal position on expression of the fusions (see below).

**Dependence of expression on C-signaling.** Intercellular C-signaling is mediated by the product of *csgA* (Hagen & Shimkets, 1990; Kim & Kaiser, 1990a; Shimkets & Asher, 1988; Shimkets et al., 1983) and is required for the expression of many *M. xanthus* genes that begin to be expressed after 6 h into development (Kroos & Kaiser, 1987; Li & Shimkets, 1993). Introduction of a *csgA* mutation into cells containing Tn*5* *lac* Ω4403 abolished developmental *lacZ* expression (Kroos & Kaiser, 1987). To determine whether the minimum Ω4403 promoter region defined by our deletion analysis also exhibits dependence on C-signaling, we measured expression of fusions integrated at Mx8 *att* in *csgA* mutant cells. Surprisingly, we found that *csgA* cells with the vector alone (Figure 2.2; i.e., containing no insert of *M. xanthus* DNA) integrated at Mx8 *att* expressed *lacZ* at a considerable level and in a developmentally regulated fashion (Figure 2.10A). This result suggests that a developmentally regulated promoter lies upstream of *lacZ* in the vector or in adjoining *M. xanthus* DNA (after integration at Mx8 *att*). Expression from this promoter appears to be inhibited by C-signaling, since expression was observed in *csgA* mutant cells (Figure 2.10A) but not in wild-type cells (Figure 2.3). *csgA* cells containing the vector with different amounts of Ω4403 upstream DNA fused to *lacZ* and integrated at Mx8 *att* exhibit a similar pattern of developmental *lacZ* expression as the vector alone. Figure 2.10A shows the results for Ω4403 DNA extending to -80 or -11, and similar results were observed for DNA extending to -616 or about -8,000 (data not shown). In wild-type cells, the deletion to -80 retained Ω4403 promoter activity and the deletion to -11

**Figure 2.10.** Extracellular complementation of developmental *lacZ* expression from the Ω4403 promoter in *csgA* cells. (A) Developmental *lacZ* expression was measured as described previously (Kroos, et al., 1986) for three independently isolated transductants containing a single copy of Ω4403 upstream DNA extending to -80 (MMF100C-6, -17, -24) (○) or -11 (MMF200C-4, -7, -11) (△) fused to *lacZ* and integrated at Mx8 *att* in a *csgA* mutant background. Points represent the average. The ß-galactosidase activity of a similar strain (MES119) with no insert of Ω4403 upstream DNA (□) is also plotted. (B) The same strains as in panel A were co-developed with an equal number of wild-type DK1622 cells (that do not express ß-galactosidase) and the ß-galactosidase specific activity was determined as described previously (Kroos and Kaiser, 1987). The units of enzyme activity are given in the Figure 2.3 legend.


ß-galactosidase Specific Activity
100
80
60
40
20
0
0
6
12
18
24
30
36
42
48
Hours of Development
No Insert
80
11


100
80
60
40
20
0
ß-galactosidase Specific Activity
No Insert+ WT
80 +WT
11 +WT
0
6
12
18
24
30
36
42
48
Hours of Development

abolished activity (Figure 2.8). In *csgA* mutant cells, the two deletions exhibit similar expression as the vector alone Figure 2.10A). These results suggest that the Ω4403 promoter is inactive in *csgA* mutant cells.

To determine whether C-signaling could restore Ω4403 promoter activity in *csgA* mutant cells, we used wild-type cells as C-signal donors. *csgA* cells with a fusion integrated at Mx8 *att* were mixed with an equal number of wild-type strain DK1622 cells and allowed to co-develop. The specific activity of ß-galactosidase in the mixture was determined at different times of development. Figure 2.10B shows that expression from cells with Ω4403 DNA extending to -80 was considerably higher than expression from cells with Ω4403 DNA extending to -11, which was comparable to the vector alone control. C-signaling appears to stimulate the Ω4403 promoter that is present in cells with the deletion to -80 and absent from cells with the deletion to -11. *csgA* mutant cells containing the deletion to -11 or the vector alone expressed *lacZ* at a lower level when mixed with wild-type cells (Figure 2.10B) than when unmixed (Figure 2.10A), consistent with the idea that C-signaling inhibits expression of an upstream promoter that lies in the vector or in adjoining *M. xanthus* DNA (after integration at Mx8 *att*). If expression of this promoter in the vicinity of Mx8 *att* is inhibited by C-signaling, it seems possible that expression of other promoters (possibly including the Ω4403 promoter) might be inhibited (partially) when located at Mx8 *att* in wild-type cells (see the Discussion).

## Discussion

The analysis of the Ω4403 promoter reported here is part of a larger effort to understand gene regulation by intercellular C-signaling during *M. xanthus* development. Ω4403 is one of five Tn*5* *lac* insertions in distinct transcriptional units that have been

shown to depend absolutely on C-signaling for expression beginning at 10-18 h into development (Kroos & Kaiser, 1987). Expression of another eight Tn*5* *lac* insertions in different genes begins at 6-16 h into development and exhibits a partial requirement for C-signaling (Kroos & Kaiser, 1987). Two fairly well-characterized genes also appear to fall into the latter class; *mbhA*, which encodes a hemagglutinin (Cumsky & Zusman, 1979; LaRossa et al., 1983; Romeo & Zusman, 1991), and *csgA*, which encodes the mediator of C-signaling (Kim & Kaiser, 1991; Li et al., 1992). By characterizing the regulatory regions of several genes in each class, we hope to uncover common features of C-signal-dependent gene regulation. These features may suggest approaches to identify regulatory proteins involved in C-signal-mediated gene expression.

To facilitate characterization of promoters identified by Tn*5* *lac* insertions, we constructed a plasmid designed specifically for this purpose. A multiple cloning site present in the plasmid upstream of the same promoterless *lacZ* segment found in Tn*5* *lac* (Kroos & Kaiser, 1984) permits construction of fusions identical to those created by a Tn*5* *lac* insertion. This allows direct comparison between the promoter activity of a cloned segment and expression from the native promoter (with all the normal upstream DNA) as detected by the chromosomal Tn*5* *lac* insertion. Upstream (relative to *lacZ*) of the multiple cloning site is a series of factor-independent transcriptional terminators (Simons et al., 1987) designed to prevent transcription from upstream vector or chromosomal (after plasmid integration) sequences from reading through the cloned segment and expressing *lacZ*. Other features of the plasmid include a bacteriophage P1incompatibility segment that mediates cointegrate formation for efficient transfer of plasmid from *E. coli* to *M. xanthus* *via* specialized P1 transduction (Gill et al., 1988a), and a segment from myxophage Mx8 that promotes site-specific recombination at the phage attachment site (Stellwag et al., 1985; Stephens & Kaiser, 1987). Upon transduction of the plasmid into *M. xanthus*, we estimate that >85% of the $Km^r$ transductants contain a single copy of the plasmid integrated

at Mx8 *att*, which is easily determined by Southern blotting. Hence, the plasmid makes it simple to test a DNA segment for promoter activity under well-defined conditions of copy number and chromosomal position.

The plasmid just described allowed us to test a series of deletions and localize Ω4403 promoter activity to within 0.6 kb upstream of the insertion site (Figure 2.1). S1 nuclease analysis of RNA from developing cells identified a single 5' end in this region, providing evidence for a single promoter with a start site located about 380 bp upstream of Tn*5* *lac* Ω4403 (Figure 2.4). The promoter appeared to be inactive in growing cells and 6 h developing cells since no signal was observed upon S1 nuclease analysis of RNA from these cells and *lacZ* expression from Tn*5* *lac* Ω4403 remained at a low background level. The promoter was clearly active by 12 h into development, as evidenced by accumulation of fusion mRNA (detected by the S1nuclease assay) and ß-galactosidase activity from Tn*5* *lac* Ω4403. The simplest explanation of these results is that the Ω4403 promoter is developmentally regulated at the level of transcription initiation. We cannot exclude the possibility that the Ω4403 promoter is active in growing cells and/or early developing cells, and that the fusion mRNA is rapidly degraded. According to this model, the rise in ß-galactosidase activity from Tn*5* *lac* Ω4403 in 12 h developing cells would result from stabilization of the fusion mRNA. It seems unlikely *a priori* that a short (382 b) segment of *M. xanthus* RNA at the 5' end could stabilize the long (at least several kb) fusion mRNA in a developmentally regulated fashion. Indeed, we were unable to detect any discrete species of fusion mRNA in developing cells by Northern blotting, suggesting that it is unstable. The native Ω4403 mRNA may be more stable, since a 1.5 kb species was detected by Northern blotting of RNA from developing cells lacking Tn*5* *lac* Ω4403. Several *M. xanthus* genes that appear to be developmentally regulated at the level of transcription initiation produce transcripts with unusually long half-lives (Downard et al., 1984; Nelson & Zusman, 1983; Romeo & Zusman, 1992).

The DNA sequence of the Ω4403 upstream region revealed an ORF that has striking similarity to aqualysin I, a secreted protease of *Thermus aquaticus* (Terada et al., 1990). The ORF could extend farther upstream, beyond the predicted translational start at position 376 (Figure 2.5). Potential (GTG) start codons are located at positions 325 and 340, but neither of these exhibits a satisfactory ribosome binding site and the longer ORFs do not exhibit the codon preference and third codon position GC bias typical of *M. xanthus* genes. If translation starts at the ATG at position 376, the ORF would encode 80 amino acids and be interrupted downstream by the Tn*5 lac* Ω4403 insertion. A 63 amino acid stretch within the ORF shows high similarity to aqualysin I (Figure 2.6), and our sequence ends a few amino acids after this stretch. It will be interesting to clone DNA downstream of Tn*5 lac* Ω4403 and determine whether the similarity to aqualysin I can be extended. The 1.5 kb Ω4403 mRNA could encode a polypeptide similar in size to the aqualysin I precursor (51 kDa) (Terada et al., 1990), if it were monocistronic. Activation of the aqualysin I precursor involves self-processing of N- and C-terminal pro-sequences (Terada et al., 1990). The Ω4403 ORF is similar to the N-terminal pro-sequence of aqualysin I, and both contain a sequence highly similar to the sequence YIVGFK; which is conserved in subtilisin-type proteases (Jacobs et al., 1985). Amino acid substitutions in the YIVVFK sequence of aqualysin I enhanced processing of the precursor protein to the active protease, but deletion of this six amino acid sequence abolished accumulation of precursor and active enzyme (Lee et al., 1991). Unlike aqualysin I, the Ω4403 ORF does not appear to encode an N-terminal signal sequence, which is found in many proteins that are translocated across the inner membrane of gram-negative bacteria (Pugsley, 1993). Perhaps the Ω4403 gene product is not secreted. Alternatively, the protein might exit the cytoplasm by a mechanism that does not require a signal peptide, as appears to be the case for some extracellular and periplasmic developmental proteins of *M. xanthus* (Guespin-Michel et al., 1993).

If the Ω4403 genes does encode a developmentally regulated subtilisin-type

protease as we have speculated, this protease does not seem to be essential for aggregation or sporulation. The insertion of Tn*5 lac* at codon 81 of the putative protease gene would probably abolish function, yet the *M. xanthus* strain containing Tn*5 lac* Ω4403 aggregates normally and produces a normal number of heat- and sonication-resistant spores (measured after 3 days of development). Of course, the Ω4403 gene product may have subtle effects on aggregation and/or spore viability that escaped detection. It is also possible that the Ω4403 gene product is functionally redundant with another *M. xanthus* protein(s).

Our principal interest in the Ω4403 gene is its mechanism of regulation. Inspection of the DNA sequence immediately upstream of the putative Ω4403 transcriptional start site revealed hexanucleotide sequences in the -35 and -10 regions with some similarity to sequences found in these regions of *E. coli* (Lisser & Margalit, 1993) and *B. subtilis* (Helmann, 1995) promoters transcribed by the major vegetative RNA polymerase containing $\sigma^{70}$ and $\sigma^{43}$, respectively. However, the 20 bp spacing between the two hexanucleotide sequences in the Ω4403 case is more than the optimal spacing. *B. subtilis* $\sigma^{43}$ RNA polymerase tolerates spacings of 22 and 21 bp between the -35 and -10 regions of the *spoIIG* and *spoIIE* promoters, respectively, but transcription requires an activator protein, SpoOA-phosphate (Baldus et al., 1994; York et al., 1992). A similar mechanism might regulate the Ω4403 promoter, since *M. xanthus* contains a gene, *sigA*, predicted to encode the major sigma factor, and SigA is very similar to *B. subtilis* $\sigma^{43}$ (and *E. coli* $\sigma^{70}$), including the domains that interact with the -35 and -10 regions of promoters (Inouye, 1990). Alternatively, the Ω4403 promoter might be recognized by a development-specific form of RNA polymerase. Two developmentally regulated sigma factors have been described in *M. xanthus*, *sigB* and *sigC*, and more are likely to exist (Apelian & Inouye, 1990; Apelian & Inouye, 1993). Neither SigB nor SigC appears to be required for Ω4403 transcription because null mutations in these *sig* genes do not prevent developmental

expression of Tn*5* *lac* Ω4403-Tc$^r$ (Brandner & Kroos, ).

Deletion analysis showed that DNA between -80 and -72 is critical for Ω4403 promoter activity. Although RNA polymerase does interact with DNA upstream of the -35 region of some promoters (Busby & Ebright, 1994), more typically a required sequence element upstream of the -35 region is indicative of a binding site for a transcriptional activator protein (Collado-Vides et al., 1991). Transcriptional activators often bind to palindromic DNA sequences (Pabo & Sauer, 1992). The region critical for Ω4403 promoter activity has two copies of the palindromic sequence 5'-CATG-3' separated by 1 bp (Figure 2.5); however, the -80 deletion that retains promoter activity (Figure 2.8) replaces the CA of the upstream palindrome with GC from the vector. Nucleotide substitutions will be required to establish the importance of the palindromic sequence, or the -35 and -10 regions, for Ω4403 promoter activity. Mutations that reduce promoter activity could provide a means to identify regulatory proteins by isolating suppressor mutants with restored promoter activity. We are currently pursuing biochemical approaches to identify proteins that bind to the Ω4403 regulatory region.

The Ω4403 promoter is not unusual among developmentally regulated *M. xanthus* genes in requiring an upstream element. All such genes examined so far exhibit this requirement (Downard & Kroos, 1993). In some cases multiple elements appear to act over distances of several kb from the transcriptional start site (Kil et al., 1990). In the case of Ω4403, however, our results show that 80 bp of DNA upstream of the transcriptional start site (Figure2.8), or 17 kb of upstream DNA, direct the same level of developmentally regulated gene expression, providing no evidence for additional upstream promoter elements.

It seems unlikely that a promoter element lies more than 17 kb upstream of Ω4403, or can still function even when located downstream of the 12 kb Tn*5* *lac* element, yet cells containingTn*5* *lac* Ω4403 express 3- to 4-fold more ß-galactosidase activity during

development than any strains we tested with fusions integrated at Mx8 *att*. We ruled out several possible explanations for this difference. The fusion-containing plasmid integrated at the phage attachment site did not appear to be unstable. Niether was it a hindrance to the normal developmental process. Tn*5 lac* Ω4403 did not disrupt a gene whose product negatively autoregulates, as was observed for Tn*5 lac* Ω4414 (Thony-Meyer & Kaiser, 1993). Cells with one copy of the Ω4403 regulatory region fused to *lacZ* and another copy in tandem fused to the intact gene at the native site expressed *lacZ* at about 70% the maximum level observed for cells containing Tn*5 lac* Ω4403 (Figure 2.9). This suggests that titration of a limiting positive-acting factor by a second copy of the Ω4403 regulatory region in cells with fusions integrated at Mx8 *att* can, in part, account for their reduced expression as compared to cells containing Tn*5 lac* Ω4403 (with only one copy of the regulatory region).

An effect of chromosomal position may also hinder expression from some fusions located at Mx8 *att*. Our results show that a putative promoter in the vicinity of the phage attachment site is down-regulated by C-signaling (Figure 2.10). Perhaps this effect is not confined to the putative promoter, but is also observed for some promoters integrated at Mx8 *att*. In particular, it seems to be difficult to observe full expression at Mx8 *att* of late developmental genes that depend absolutely on C-signaling. Li and Shimkets (Li & Shimkets, 1988) observed about 50% as much developmental ß-galactosidase activity in cells with 1.3 or 11 kb of Ω4435 upstream DNA fused to *lacZ* and integrated at Mx8 *att* as compared with cells containing Tn*5 lac* Ω4435. Similarly, 1.3 or 10 kb of Ω4459 upstream DNA inserted in the plasmid we described here and integrated at Mx8 *att* drives only about 10-20% as much developmental *lacZ* expression as seen with cells containing Tn*5 lac* Ω4459 (Brandner et al.). Also, Thony-Meyer and Kaiser(Thony-Meyer & Kaiser, 1993) reported poor expression of *devRS* at Mx8 *att* as compared to the native site. These authors invoked differential condensation of the chromosome during development as a

possible explanation of the positional effects. Given that Tn*5 lac* Ω4403 is within 5-25 kb of Mx8 *att* (Shimkets, ), the differential condensation would have to be highly localized. Perhaps this involves a silencing mechanism similar to that used in yeast mating type switching (Dillin & Rine, 1995). Genes that are expressed earlier during development and/or that depend less strongly on C-signaling for expression seem to escape inhibition. Cells with Ω4499 (Fisseha & Kroos) or Ω4514 (Hao & Kroos) upstream DNA inserted in the plasmid shown in Figure 2 and integrated at Mx8 *att* show a comparable level of developmental *lacZ* expression as cells containing the corresponding Tn*5 lac* insertions. Taking all these observations together, we speculate that C-signaling leads to an altered chromosomal state of DNA integrated at the Mx8 phage attachment site and DNA nearby, partially inhibiting expression of genes that depend strongly on C-signaling. The Ω4403 promoter was sufficiently active upon integration at Mx8 *att* to permit deletion analysis (Figures 2.1 and 2.8). For promoters less active at Mx8 *att*, integration of plasmids by homologous recombination to produce structures like that depicted in Figure 2.9B might be the preferred approach.

Unexpectedly, our plasmid vector alone, with no insert of *M. xanthus* DNA, expressed *lacZ* upon integration at Mx8 *att* in *csgA* mutant cells (Figure 2.10A). A promoter that is active in developing *csgA* mutant cells must lie upstream of *lacZ* in the vector or in adjoining *M. xanthus* DNA after integration at Mx8 *att*. We did not map this promoter; however, it is likely to be located upstream of the multiple cloning site present in the vector. The same sequences downstream of the multiple cloning site in the plasmid are also present in Tn*5 lac* (Kroos & Kaiser, 1984), and many Tn*5 lac* insertions express only a low background level of ß-galactosidase when present in *csgA* mutant cells (Kroos & Kaiser, 1987). The multiple cloning site itself cannot be the origin of the transcription because it is almost completely deleted in some of the plasmids that contain Ω4403 upstream DNA inserts and show a similar pattern of *lacZ* expression as the vector alone.

Neither the transcriptional terminators engineered into the vector nor inserts of Ω4403 upstream DNA up to 8.5 kb in size seemed to impede the transcription, suggesting that the RNA polymerase is in an antitermination mode. Interestingly, this transcription is negatively regulated by C-signaling. *csgA*+ cells express only a low background level of ß-galactosidase when the vector is present at Mx8 *att* (Figure 2.3). Moreover, mixing *csgA*+ cells with *csgA* mutant cells containing the vector at Mx8 *att* inhibits *lacZ* expression from the *csgA* cells (Figure 2.10B).

The ability of *csgA*+ cells to donate C-signal and substantially reduce *lacZ* expression from *csgA* mutant cells containing the vector alone integrated at Mx8 *att* allowed us to demonstrate that the minimal Ω4403 promoter, as defined by deletions, is responsive to C-signaling. We conclude that elements needed for C-signal-dependent expression of Ω4403 lie between -80 and +382. Our deletion analysis showed that an element critical for Ω4403 promoter activity lies between -80 and -72. A simple model is that this DNA element interacts with a transcriptional activator protein whose synthesis or activity is controlled directly or indirectly by C-signalling. The challenge now is to identify the regulatory protein(s).

# Chapter 3

# Characterization of the DNA upstream of a C-signal-dependent operon

## Abstract

Tn*5* *lac* Ω4499 is a fusion of a promoterless *lacZ* gene to a developmentally regulated promoter in the *Myxococcus xanthus* chromosome. Expression of Tn*5* *lac* Ω4499 in wild-type cells starts about six hours after the initiation of development and peaks at about 24 hours of development. Normal expression from the Ω4499 promoter requires extracellular C-signaling, which is mediated by the product of the *csgA* gene. A lesion in *csgA* reduces *lacZ* expression from Tn*5* *lac* Ω4499 to approximately 25 % of the expression in wild-type cells (Kroos & Kaiser, 1987).

Based on *in vivo* deletion analyses of the Ω4499 upstream region, regulatory elements required for expression of *lacZ* in a C-signal-dependent manner were localized to between 2.1 and 3.2 kb upstream of the insertion. DNA sequence analysis showed that there are two open reading frames (ORF1 and ORF2) in the 2.1 kb of upstream DNA. ORF1 is similar to NADH-putidaredoxin reductase from *Pseudomonas putida* and ORF2 is similar to cytochrome P-450 from *Streptomyces griseolus*. The DNA sequence between ORF1 and ORF2 shows no similarity to transcriptional termination sequences or to known promoter sequences from *M. xanthus* or other bacteria. Together with RNA analyses, these results suggest that Ω4499 is inserted into a developmentally regulated operon whose transcription initiates more than 2.1 kb upstream of the Ω4499 fusion point; however, the multicistronic RNA appears to be very susceptible to degradation.

## Introduction

*Myxococcus xanthus* is a gram-negative, soil bacterium that undergoes multicellular development when nutrient sources are depleted. The developmental process is complex and involves the movement of cells into mounds known as fruiting bodies and the differentiation of some cells into ovoid myxospores. This developmental process is coordinated by cell-cell interactions, five of which have been identified so far. These cell-cell interaction systems are known as the A-, B-, C-, D-, and E- signaling interactions(Hagen et al., 1978; Downard et al., 1993). Our studies are focused on understanding how extracellular C-signaling regulates the expression of developmental genes. The relative genetic simplicity of the C-signaling interaction make it an attractive model system for studying cell-cell interaction-regulated gene expression. All of the mutations that affect C-signaling map to a single genetic locus known as *csgA*(Shimkets & Asher, 1988; Shimkets et al., 1983). The *csgA* gene product is well-characterized and encodes an alcohol dehydrogenase (Lee & Shimkets, 1994). Extracellular addition of a purified *csgA* gene product rescues the developmental defect of C-signaling defective cells (Shimkets & Rafiee, 1990; Kim & Kaiser, 1990a).

The C-signaling interaction is required starting at about six hours after the onset of development, well separated from the time at which the other signaling systems are first required. Genes whose expression begins after six hours following the initiation of development require C-signaling at least partially. To investigate how extracellular C-signaling regulates gene expression during development, the upstream regulatory region of Tn*5* *lac* Ω4499, a transcriptional fusion of a promoterless *lacZ* gene to a developmentally regulated transcriptional unit, whose expression has only a partial requirement for C-signaling, was examined. This fusion was one of 29 fusions to developmentally regulated transcriptional units that were identified by Kroos et al. (Kroos et al., 1986). These

fusions were generated by transposition of a promoterless Tn*5* *lac* promoter probe into the *M. xanthus* chromosome and identification of transductants that increased *lacZ* expression during development but not during vegetative growth. The promoter region of one other gene, *csgA*, whose expression is stimulated by C-signaling, has been characterized (Li et al., 1992). *csgA* requires a fairly large upstream sequence to modulate its expression in response to nutrient levels, peptidoglycan components and B-signaling. At least 400 bp upstream from the transcription initiation site was required for the maximum expression of *csgA* in the absence of nutrients. In the presence of low levels of nutrients, sequences farther upstream (~930 bp) were required for proper developmental expression.

Deletion analyses were used to show that promoter elements required for C-signal-dependent expression are found between 3.2 and 2.1 kb upstream of the Ω4499 insertion point. Transcription from the Ω4499 promoter generates a transcript of at least 3.5 kb which appears to be unstable. Analysis of the 2.1 kb DNA sequence directly upstream of the fusion suggests that the Ω4499 insertion is in an operon that encodes enzymes involved in metabolizing alternative carbon sources. In addition to localizing the promoter region, these studies provide information on the type of genes that are expressed during development.

## Materials and Methods

**Bacterial Strains and Plasmids.** Strains and plasmids used in this work are listed in Table 3.1.

**Growth and Development.** *Escherichia coli* cells were grown at 37°C in LB medium containing 50 μg of ampicillin or 25 μg of kanamycin per ml when necessary. *M. xanthus* was grown in CTT (1 % casitone, 10 mM Tris.HCl, 1 mM $KH_2PO_4$-$K_2HPO_4$, 8

mM $MgSO_4$, pH 8.0) medium in liquid cultures or on agar plates as described previously (Kroos et al., 1986). 40 μg of kanamycin or 12.5 μg oxytetracycline per ml were used when required for selective growth. Fruiting body development was performed on TPM (10 mM Tris, pH 8.0, 1 mM $KH_2PO_4$, 8 mM $MgSO_4$, final pH 7.6) agar plates (1.5 % Bacto-agar) as described previously (Kroos et al., 1986).

**Molecular cloning and construction of plasmids.** Recombinant DNA work was performed using standard techniques (Sambrook et al., 1989). Plasmid DNA was prepared from *E. coli* strains JM83 or DH5α. *M. xanthus* DNA used for cloning was prepared as described previously (Laue & Gill, 1994). To test the Ω4499 upstream DNA for promoter activity, fragments were subcloned into the promoter-testing plasmid, pREG1666 or pREG1727 (Chapter 2).

To construct pMF1, pREG4499 was digested with *Xho*I, and the 22.3 kb fragment was purified from a low melting point agarose gel and religated at a low DNA concentration.

pMF500 was constructed by digesting pMF0051 with *Nco*I and making the ends blunt by digesting with mung-bean nuclease (Sambrook et al., 1989). After phenol-

Table 3.1 Bacterial strains and plasmids

| Strain or plasmid | Relevant characteristics or genotype | Source or reference |
|---|---|---|
| *E.coli* | | |
| DH5α | ϕ80 Δ*lacZ* ΔM15 Δ*lacU169 recA1 endA1 hsdR17 supE44 thi-1 gyrA relA1* | Hanahan, D., 1983 |
| JM83 | *ara* Δ*lac-pro strA thi* fd*lacZ* ΔM15 | Messing, D., 1979 |
| *M. xanthus* | | |
| DK1622 | Wild-type | Kaiser, D., 1979 |
| DK4368 | Tn*5 lac* ($Km^r$) Ω4403 | Kroos et al., 1986 |
| DK4499 | Tn*5 lac* ($Km^r$) Ω4499 | Kroos et al., 1986 |
| JW103,104, 107 | *attB*::pREG1666 | Jamie White |
| MMF1727-7, -9, -11 | *attB*::pREG1727 | This work |
| MMF21-1, -2, -4, -10 | *attB*::pMF21 | This work |
| MMF51-60, -70,-80, | *attB*::pMF51 | This work |
| MMF41-25, 40, 42 | *attB*::pMF41 | This work |
| MMF521-6, -16, -20 | *attB*::pMF521 | This work |
| MMF500-3, -24, -38, -39 | *attB*::pMF500 | This work |
| DK5208 | *csgA*:: Tn*5-132* ($Tc^r$) Ω205 | Shimkets, L. and Asher, 1988 |
| MES119 | *csgA*:: Tn*5-132* ($Tc^r$) Ω205*attB*::pREG1666 | This work |
| MMF21C-5, -15, -16, -50 | *csgA*:: Tn*5-132* ($Tc^r$) Ω205*attB*::pMF21 | This work |
| MMF51C-7, -23, -36, -40 | *csgA*:: Tn*5-132* ($Tc^r$) Ω205*att:B*::pMF51 | This work |

Table 3.1 (cont'd)

| | | |
|---|---|---|
| pGEM-7Zf | Ap$^r$ *lacα* | Yanisch-Perron et al., 1985 |
| pUC19 | Ap$^r$ *lacα* | Promega |
| pREG1666/pREG1727 | Ap$^r$ Km$^r$ P1-*inc attP* | chapter 2 |
| pIP110 | Ap$^r$ (pGEM-7Zf), *M xanthus* Ω4499, Km$^r$ (*aphII*), 20.4 kb *Hin*dIII fragment from DK4499 | Il Sun Park |
| pREG4499 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 9.4 kb *Hin*dIII-*Bam*HI fragment from pIP110 | Beta Borer |
| pMF002 | Ap$^r$ (GEM-7Zf) 4.5 kb *Sma*I-*Bam*HI from pIP110 in *Sma*I-*Bam*HI | This work |
| pMF21 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 4.5 kb *Xba*I-*Bam*HI fragment from pMF002 | This work |
| pMF0051 | Apr (pGEM-7Zf), 3.2 kb *Pst*I-*Bam*HI from pMF002 | This work |
| pMF51 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1666), 3.2 kb *Hin*dIII-*Bam*HI fragment from pMF0051 | This work |
| pMF41 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), 3.0 kb *Sma*I-*Xho*I fragment from pMF002 | This work |
| pMF2.3 | Ap$^r$ (pUC19), 2.1 kb *Nco*I-*Bam*HI fragment from pMF0051 in pUC19 | This work |
| pMF521 | Ap$^r$ Km$^r$ P1-*inc attP* (pREG1727), 2.1 kb *Hin*dIII-*Bam*HI fragment from pMF2.3 | This work |
| pMF1 | Ap$^r$ Km$^r$ P1-*inc* attP (pREG1666) 1.5.kb *Xho*I-*Bam*HI fragment from pREG4499 | This work |
| pMF400 | Ap$^r$ (pGEM-7Zf), 1.5 kb *Xho*I-*Bam*HI fragment from pMF1 | This work |
| pMF060-1 | Ap$^r$ (pGEM-7Zf), *M. xanthus* 1.3 kb *Sph*I-*Sph*I from pMF400 | This work |
| pMF030-26 | Ap$^r$ (pGEM-7Zf), *M. xanthus* 300 bp *Nco*I-*Xho*I from pMF2.3 | This work |

Table 3.1 (cont'd)

| | | |
|---|---|---|
| pMF040-110 | $Ap^r$ (pGEM-7Zf), *M. xanthus* 350 bp *Nco*I-*Nco*I from pMF 2.3 | This work |

chloroform extraction and ethanol precipitation, the DNA was digested with *Bam*HI and fractionated on a low melting point agarose gel. The 1.8 kb fragment was purified and directionally cloned into pREG1727 that had been digested with *Hin*dIII, subjected to end-filling using the Klenow fragment of DNA polymerase I (Sambrook et al., 1989), digested with *Bam*HI and gel-purified.

pMF2.3 was constructed by digesting pMF0051 to completion with *Pst*I followed by a partial digest with *Nco*I. The 4.8 kb fragment which contains pUC19 and the 2.1 kb *Nco*I-*Bam*HI fragment was purified from a low melting point agarose gel, digested with mung-bean nuclease (Sambrook et al., 1989) and religated.

Plasmids pMF030-26, 040-10, and 060-1 were constructed in pGEM-7Zf to serve as templates for the synthesis of riboprobes. In each case, the insert fragment was ligated into the multiple cloning site downstream of the SP6 promoter in the proper orientation so that transcription by SP6 RNA polymerase would produce a complement of the putative Ω4499 message. To make pMF040-10, the 350 bp *Hin*dIII-*Nco*I fragment from pMF2.3 (containing *M. xanthus* DNA between 2.1 and 1.8 kb upstream of Ω4499) was gel-purified and ligated into pGEM-7Zf that had been digested with *Hin*dIII and dephosphorylated as described previously (Sambrook et al., 1989). Following ligation, the incompatible ends were made blunt by digesting with mung-bean nuclease. The DNA was extracted with phenol and chloroform and religated. Similarly, pMF030-26 was made by gel-purifying the 300 bp *Nco*I-*Xho*I fragment from pMF2.3 (containing *M. xanthus* DNA between 1.9 and 1.6 kb upstream of Ω4499) and ligating it into *Xho*I-digested and dephosphorylated pGEM-7Zf. Following the ligation, the incompatible ends were made blunt by digesting with mung-bean nuclease and the DNA was extracted with phenol and chloroform, ethanol precipitated and religated.

**Construction of *M. xanthus* strains.** Strains containing pREG1666 or pREG1727, or derivatives of these plasmids, integrated at the Mx8 phage attachment site (designated *attB* in Table 3.1) were constructed by P1 specialized transduction from the *rec+ E. coli* strain JM83 into wild-type *M. xanthus* strain DK1622 or the *csgA* mutant strain DK5208 as described previously (Gill et al., 1988a). For each plasmid, three or four transductants (listed in Table 3.1), each containing a single copy of the plasmid integrated at Mx8 *att*, were identified using Southern blot analysis (data not shown).

**Southern blot analysis.** *M. xanthus* chromosomal DNA was isolated using the method of Avery and Kaiser (Avery & Kaiser, 1983). Three to 5 μg of DNA was digested with the appropriate restriction enzyme(s) and fractionated on a 0.5% agarose gel. DNA was transferred onto a nylon membrane (Fisher Scientific) using the alkaline transfer method described by Sambrook et.al.(Sambrook et al., 1989). Probe synthesis and hybridization of the membrane was carried out using the Genius kit (Boeringer Mannheim) according to the manufacturer's specifications.

**Expression of ß-galactosidase in *M. xanthus*.** Strains were grown to the logarithmic stage (5 x $10^8$ cells/ml), resuspended to 5 x $10^9$ cells/ml in TPM, and plated for development as described previously (Kroos et al., 1986). Assays for ß-galactosidase expression were performed as described by Kroos et.al. (Kroos et al., 1986).

**RNA isolation.** Total RNA was purified from DK4368 (which contains the wild-type Ω4499 transcript) using the protocol of Igo and Losick (Igo & Losick, 1986), but with the following modifications. Cells were harvested from exponentially growing cultures (50 ml of 5 x $10^8$ cells/ml) or from an equivalent number of cells at 24 h of development on TPM agar plates, washed with ice-cold TPM, and resuspended in 2 ml LETS buffer (100 mM LiCl, 10 mM EDTA, 10 mM Tris pH 7.8 with LiCl, 1 % SDS). The nucleic acid was resuspended in 200 μl of diethylpyrocarbonate (DEPC)-treated water

(0.1% v/v) and treated with RNase-free DNase (Boeringer Mannheim) to eliminate DNA. DNase was removed by extracting with phenol-chloroform and precipitating with ethanol. RNA was resuspended in DEPC-treated water and used for further analysis.

**Northern Blot Analysis.** Northern blot analysis was carried out as described previously (Ausubel et al., 1989) except with the following modifications. RNA was fractionated on a 0.6% agarose gel containing 1.1% (v/v) formaldehyde and transferred onto a nitrocellulose membrane (Schleicher & Schull). The membrane was baked for 2 h at 80°C. Riboprobes were synthesized from pMF400, pMF030-26, pMF040-10, and pMF060-1 linearized with restriction enzymes *Sph*I, *Xba*I, *Eco*RI and *Aat*II, respectively. These restriction enzymes cut in the multiple cloning site downstream (relative to the SP6 promoter) of the insert fragment. Riboprobes were synthesized using the SP6 riboprobe kit from Promega and [$\alpha^{32}$P]CTP. All of the run-off transcripts except the one derived from *Sph*I- digested pMF400 were expected to be of the same length as the insert DNA. *Sph*I cuts pMF400 about 230 bp upstream of *Bam*HI near the junction of *M. xanthus* DNA and Tn*5 lac* Ω4499. Therefore, the expected size of the run-off transcript was about 230 bp. Samples of the riboprobes were electrophoresed on 5% polacrylamide, 7 M urea gels to verify the sizes. Prehybridization was in 50% formamide, 6 X SSPE, 0.5% SDS, 5 X Denhardt's solution, 100 µg/ml sheared salmon sperm DNA for 2-3 h at 42°C, and hybridization was overnight at 56°C. The hybridization solution was the same as the prehybridization solution except that it contained 1-2 x $10^7$ CPM of the appropriate riboprobe. Following hybridization, the membranes were washed at 65°C for 30 minutes each time once in 2 X SSC 0.1% SDS, 1 X SSC 0.1% SDS and two times in 0.1 X SSC, 0.1% SDS. Membranes were exposed to X-ray films at 70°C with intensifying screens.

**DNA sequencing.** Plasmids pMF0051 and pMF060-1 were used for double-stranded sequencing of the 2.1 kb of DNA upstream of Ω4499. Sequencing was carried out on both strands using synthetic oligonucleotide primers. Double-stranded sequencing

was performed by the Sanger method (Sanger et al., 1977) using the Sequenase kit (United States Biochemical). Ambiguities arising from premature termination were resolved using the protocol of Fawcett and Bartlett (Fawcett & Bartlett, 1990). Briefly, 1 µl of a reaction mixture containing terminal deoxynucleotide transferase (1 µM of each dNTP, pH 7.0, 2 units/µl terminal deoxynucleotide transferase, 1 X sequenase reaction buffer) was added to each of the termination reactions (7 µl total) and incubated at 37°C for 30 min. The reaction was terminated using 4 µl of stop buffer (United States Biochemical). 7-deaza dGTP reaction mixes (United States Biochemical) were used to resolve regions of compression. DNA and protein sequence analyses were done using the University of Wisconsin Genetics Computer Group software package.

## Results

**Analysis of the Ω4499 upstream region for promoter activity.** To identify the promoter region upstream of the partially C-signal-dependent Tn*5 lac* fusion Ω4499, the upstream DNA had previously been cloned (Il-Seon Park, unpublished results) using *Hin*dIII restriction sites located approximately 9.4 kb upstream of the Ω4499 insertion (Kroos et al., 1986) and downstream of the *aphII* gene (which encodes aminoglycoside phosphotransferase, and confers resistance to kanamycin) in Tn*5 lac* (Figure 3.1A shows a partial restriction map). Thus, *Hin*dIII digestion of chromosomal DNA from *M. xanthus* containing Tn*5 lac* Ω4499 with *Hin*dIII generates a fragment of about 20.4 kb containing the gene for kanamycin resistance. Chromosomal DNA isolated from DK4499 cells was digested with *Hin*dIII, ligated into *Hin*dIII-digested pGEM-7Zf and transformed into *Escherichia coli* strain DH5α. Transformants were selected for resistance to ampicillin (conferred by the vector) and kanamycin (conferred by the cloned

*M. xanthus* DNA). One transformant carrying the plasmid with Ω4499 upstream DNA, pIP110, was identified and analyzed by restriction digestion using sites that had previously been mapped in *M. xanthus* and in Tn*5 lac* (Kroos et al., 1986) (Figure 3.1A). pIP110 contains approximately 9.4 kb of Ω4499 upstream DNA and 11 kb of DNA from Tn*5 lac*.

To test the DNA upstream of Ω4499 for promoter activity, a *Sma*I-*Bam*HI fragment that includes about 50 bp from the left end of Tn*5* and 4.5 kb of DNA directly upstream of Ω4499 (Figure 3.1A) was subcloned into the multiple cloning site of pREG1666. This creates a fusion of the upstream DNA with a promoterless *lacZ* gene in exactly the same configuration found in DK4499. This plasmid, pMF21, was introduced into wild-type *M. xanthus* strain DK1622, by P1 specialized transduction. Kanamycin-resistant *M. xanthus* transductants contain pMF21 integrated into the chromosome. Since pREG1666 contains an Mx8 phage attachment site (Mx8 *att*) the plasmid integrates either

**Figure 3.1 A.** Physical map of the Ω4499 insertion region. The restriction sites in the Tn*5 lac* and the adjacent *M. xanthus* DNA that were used in cloning experiments are shown. The distance between restriction sites andthe point of Tn*5 lac* Ω4499 insertion is given in kb. B, *Bam*HI; H, *Hin*dIII; M, *Sma*I; N, *Nco*I; P, *Pst*I; S, *Sph*I; X, *Xho*I

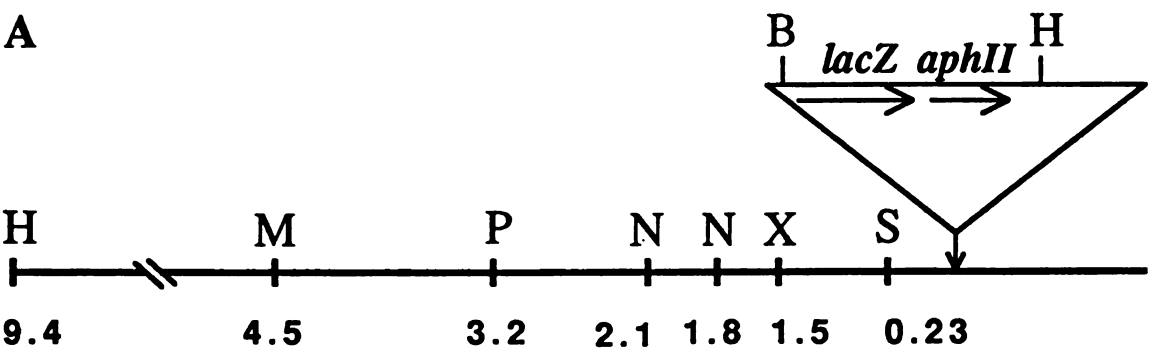
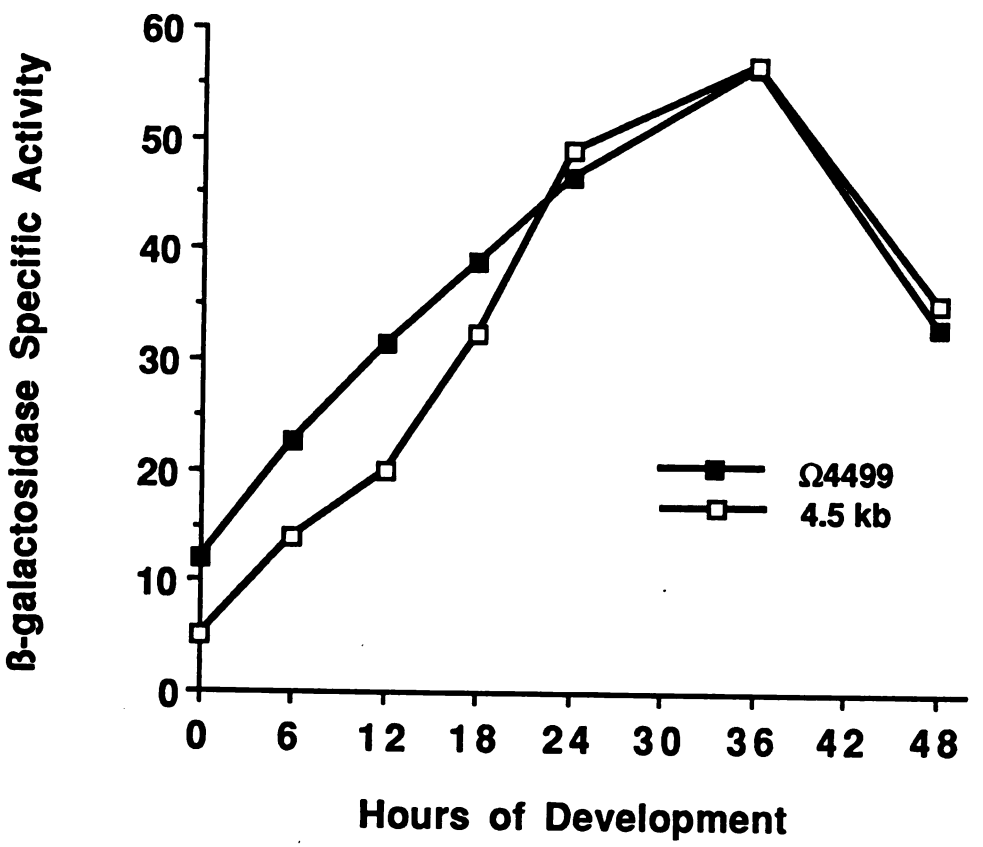

A
B
lacZ aphII
H
H
M
P
N
N
X
S
9.4
4.5
3.2
2.1
1.8
1.5
0.23

**Figure 3.1 B.** Developmental *lacZ* expression from the Ω4499 promoter. ß-galactosidase specific activity during development was determined as described previously (Kroos, et al., 1986). The ß-galactosidase activity from a single determination of *M. xanthus* strain DK4499, containing Tn*5* *lac* Ω4499, is shown (■). The average *lacZ* expression from three independently isolated transductants containing 4.5 kb of Ω4499 upstream DNA fused to *lacZ* and integrated at the Mx8 *att* site in single copy is plotted (❑). The units of ß-galactosidase specific activity are nanomoles ONP (ortho nitrophenyl phosphate) minutes$^{-1}$ milligram of protein$^{-1}$.

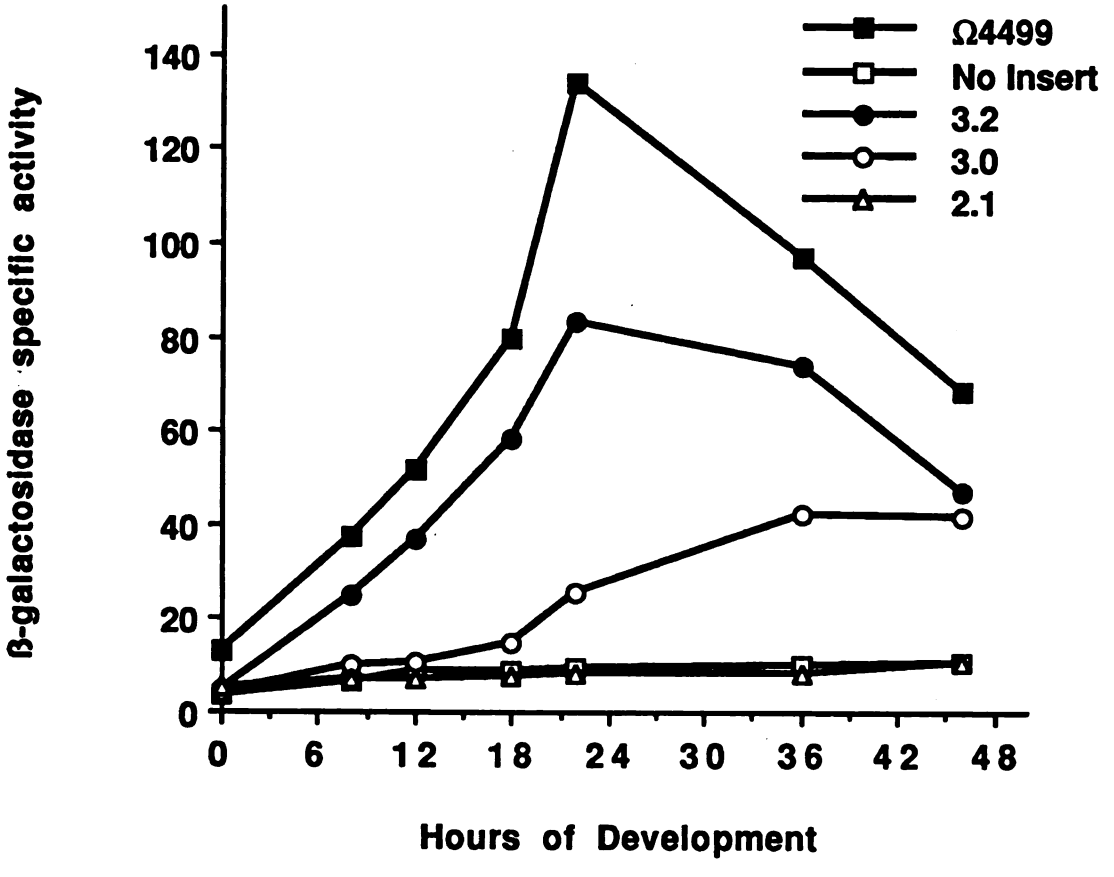
B


ß-galactosidase Specific Activity
60
50
40
30
20
10
0
0
6
12
18
24
30
36
42
48
Hours of Development
Ω4499
4.5 kb

by site-specific recombination at the *attB* site in the chromosome or by homologous recombination through the cloned upstream region. Transductants containing a single copy of pMF21 integrated at Mx8 *att* were identified by Southern blot analysis and several of these were analyzed for the production of ß-galactosidase during development. Figure 3.1B shows that 4.5 kb of Ω4499 upstream DNA directs *lacZ* expression with similar timing and level as observed for DK4499. This indicates that 4.5 kb of DNA upstream of Ω4499 contains all of the regulatory elements required for proper expression of *lacZ* in a developmentally regulated fashion.

To further localize the region upstream of Ω4499 that is required for proper expression of developmental *lacZ*, deletions of the upstream DNA were generated and tested for promoter activity in the same way as described above. Figure 3.2 shows that 3.2 kb of Ω4499 upstream DNA directs the expression of *lacZ* with the same timing as in DK4499, indicating that the 3.2 kb region contains a developmentally regulated promoter. The maximum level of expression from this region was about 65% of the level in DK4499 (after subtracting background expression in the strains containing the vector alone with no insert). These results indicate that the DNA between 4.5 and 3.2 kb upstream of Ω4499 contains a promoter(s) and/or promoter element(s) that elevates expression to the level observed for DK4499 cells. Further deletion of the upstream DNA to 2.1 kb or less abolished developmental ß-galactosidase production, indicating that essential promoter elements lie more than 2.1 kb upstream of the Ω4499 insertion point. The 3.0 kb region between 4.5 and 1.5 kb upstream of Ω4499 directed expression of about 30% as much ß-galactosidase activity as seen with DK4499 cells, and the timing of expression from this DNA segment was delayed by about 12 hours. This result is consistent with a developmentally regulated promoter being located between 2.1 and 3.2 kb upstream of Ω4499 but shows that DNA within 1.5 kb immediately upstream of Ω4499 is also

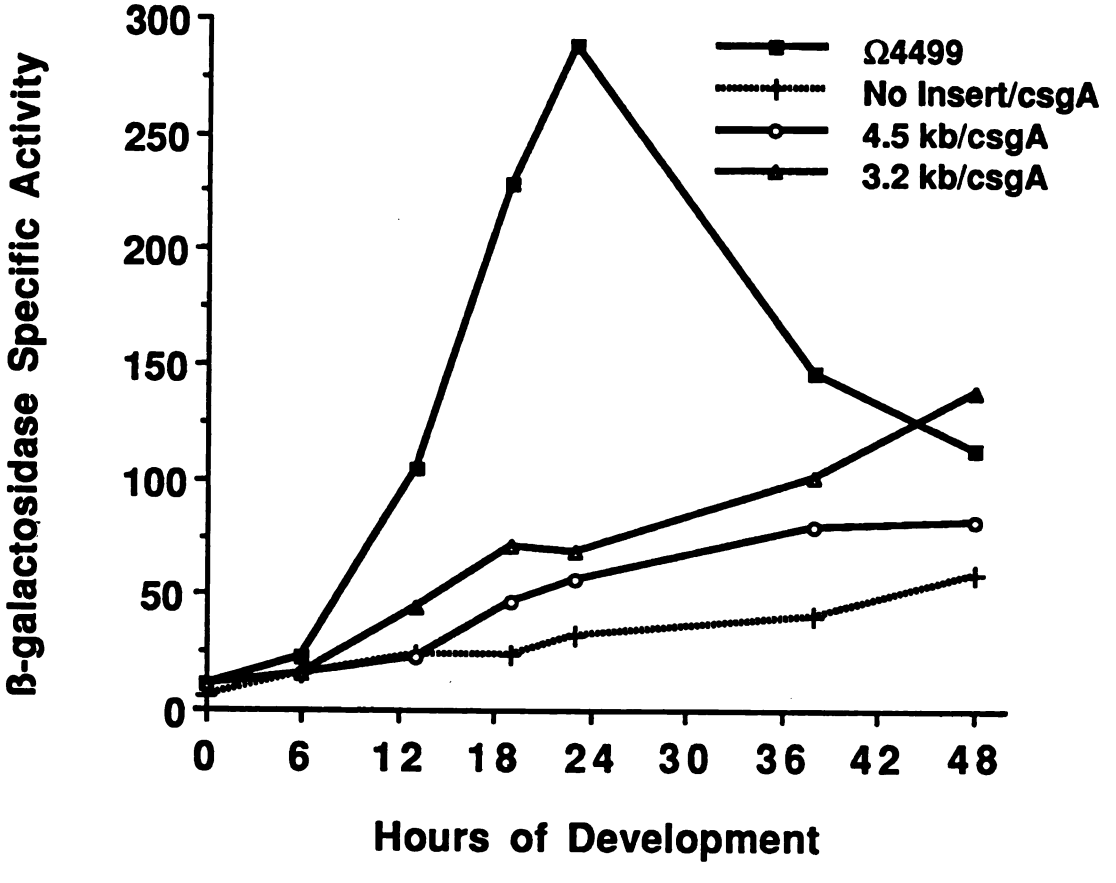
**Figure 3.2.** Deletion analysis of the Ω4499 promoter region. Developmental *lacZ* expression was determined as described previously (Kroos, et al., 1986) for transductants containing a single copy of either different segments of Ω4499 upstream DNA fused to *lacZ*, or the vector alone with no insert DNA (❑), at the Mx8 *att* site. The DNA segments tested for promoter activity included 3.2 (●) or 2.1 kb (Δ) directly upstream of Ω4499 and 3.0 kb (○) between *Sma*I and *Xho*I (Figure 3.1 A). *lacZ* expression from DK4499 cells containing Tn*5* *lac* Ω4499 is also shown (■). The units of enzyme activity are given in the legend of Figure 1B. Points represent the average for three or more independently isolated transductants, except the DK4499 data was a single determination made in the same experiment.


Ω4499
No Insert
3.2
3.0
2.1
ß-galactosidase specific activity
0
20
40
60
80
100
120
140
0
6
12
18
24
30
36
42
48
Hours of Development

required for the normal timing and levels of expression.

**C-signal dependent expression of Ω4499.** The Tn*5 lac* Ω4499 fusion is one of many developmentally regulated fusions whose expression requires extracellular C-signaling (Kroos & Kaiser, 1987). C-signaling is mediated by the product of the *csgA* gene (Shimkets & Rafiee, 1990; Kim & Kaiser, 1990a). *lacZ* expression in *csgA* mutant cells containing the Tn*5 lac* Ω4499 fusion is reduced to about 25% the level observed in wild-type cells, indicating that the expression of Ω4499 is partially dependent on extracellular C-signaling (Kroos & Kaiser, 1987). To determine whether *lacZ* expression from the cloned regions that have promoter activity exhibit C-signal dependency, fusions of the 4.5 and 3.2 kb upstream regions with *lacZ* were integrated at the Mx8 *att* site of *csgA* cells and tested for developmentally regulated promoter activity. Figure 3.3 shows that *lacZ* expression from these fusions is only 10-20% the level observed for wild-type cells containing Tn*5 lac* Ω4499 measured in parallel, indicating that the cloned promoter region is partially dependent on C-signaling for proper expression. The negative control strain, which contains pREG1666 alone with no *M. xanthus* insert DNA integrated at Mx8 *att,* reached about 50 units of ß-galactosidase activity at 48 hours of development. This activity is attributed to read-through transcription from a fortuitous promoter in the vector or in the adjacent *M. xanthus* DNA. Since *lacZ* expression did not rise above above 15 units in wild-type cells containing pREG1666 at Mx8 *att*, this read-through transcription must normally repressed by C-signaling (Chapter 2). The expression of *lacZ* in transductants containing the 4.5 and 3.2 kb fusions was greater than in transductants containing the vector alone indicating that there is some expression from these regions in *csgA* cells, just as was observed for *csgA* mutant cells containing Tn*5 lac* Ω4499.

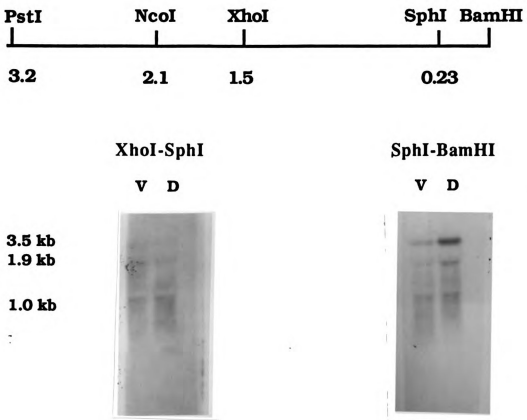
**Figure 3.3.** Developmental *lacZ* expression in *csgA* mutant cells. ß-galactosidase activity was measured as described previously for cells containing fusions of 4.5 (○) or 3.2 kb (Δ) of Ω4499 upstream DNA to *lacZ* integrated in single copy at the Mx8 *att* site. Points represent the average for three or more independently isolated transductants. A single determination of expression from csgA mutant *cells* containing a single copy of the vector alone (with no insert DNA) integrated at Mx8 *att* (+) and from DK4499 cells containing Tn*5* *lac* Ω4499 (■) was made in the same experiment.


ß-galactosidase Specific Activity
300
250
200
150
100
50
0
0
6
12
18
24
30
36
42
48
Hours of Development
Ω4499
No Insert/csgA
4.5 kb/csgA
3.2 kb/csgA

**Analyses of the Ω4499 transcript.** To determine the size of developmental transcripts produced from the Ω4499 region and to start localizing their 5' ends Northern blot analyses were performed. RNA was isolated from cells (containing the wild-type Ω4499 gene) that were either grown vegetatively or that were developed for 24 hours on starvation agar.

A 3.5 kb transcript was detected by riboprobes synthesized from pMF060-1 (linearized by digesting with *Aat*II) and from pMF400 (linearized by digesting with *Sph*I) (Figure 3.4). These riboprobes are complementary to the 0.23 kb region directly upstream of Ω4499 (detected by the pMF400 riboprobe), and to the region between 0.23 and 1.5 kb upstream of Ω4499 (detected by the pMF060-1 riboprobe). Two additional transcripts of about 1.9 and 1.0 kb were also detected using these riboprobes. These may be degradation products of the larger transcript. The intensity of all three transcripts was greater in RNA isolated from 24 hour developing cells compared to RNA isolated from vegetatively growing cells. This suggests that the Ω4499 transcript accumulates more in developing cells, as might be expected from the *lacZ* fusion results. The greater accumulation of transcripts in developing cells may be due to increased synthesis and/or stability of the mRNA. Riboprobes synthesized from DNA more than 1.5 kb upstream of the Ω4499 insertion point (using pMF030-26 and pMF040-10, data not shown) did not detect any of these transcripts. These results suggest that the 5' end(s) of the Ω4499 transcripts may lie within 1.5 kb upstream of Ω4499. Alternatively, it is possible that the native Ω4499 transcript is unstable at its 5' end.

**Figure 3.4.** Northern blot analysis of the Ω4499 transcript. RNA was isolated from vegetatively growing cells (lanes 1) and from 24 hour developing cells (lanes 2). Riboprobes were synthesized as described in the Materials and Methods, and used to detect right ward transcripts from the regions indicated above each autoradiogram. The sizes of the transcripts are indicated on the left.


PstI
NcoI
XhoI
SphI
BamHI
3.2
2.1
1.5
0.23
XhoI-SphI
V D
SphI-BamHI
V D
3.5 kb
1.9 kb
1.0 kb

**DNA sequence analysis of the Ω4499 upstream region.** Since the size and organization of open reading frames in DNA upstream of Ω4499 might help distinguish between alternative models about the location of the transcriptional start site(s), 2.1 kb of DNA upstream of Ω4499 was sequenced. Figure 3.5 shows the DNA sequence of this region. Two ORFs were identified based on the strong bias for guanine and cytosine nucleotides at the third position of codons (expected for G+C rich organisms like *M. xanthus*), and the usage of codons preferred by *M. xanthus*. Figure 3.6 shows the third position GC bias and codon preference within the 2.4 kb of Ω4499 upstream DNA. The sequence between 2.1 and 2.4 kb was determined in one strand and contains some ambiguities. ORF2 begins 939 bp upstream of the Ω4499 insertion point (shown in Figure 3.6 as the open bar in the first reading frame) and remains open downstream to the point of fusion with Tn*5* *lac* Ω4499. Starting at about nucleotide 1200, the third position GC bias and the codon preference are higher than would be expected for random occurrences (indicated in Figure 3.6 by the upper and lower graphs, respectively), indicating that there may be an ORF (designated ORF1 in Figure 3.5) in this region. The absence of a translational initiation codon for ORF1 suggests that translation may start beyond 2.1 kb upstream of the Ω4499 insertion point. A translation stop codon is present at nucleotide 1066 (102 bp upstream of ORF2), and delimits the N-terminal end of ORF1. However, the sequence between 1066 and 1200 (Figure 3.5) does not show a bias for G and C in the third position of codons (around nucleotide 1200 to 1400 in Figure 3.6). The 300 bp region between 2.4 and 2.1 kb upstream of Ω4499 (Figure 3.6) exhibits poor third position GC bias, usage of many rare codons, and several stop codons in the first reading frame. However, the same region in the third reading frame has a relatively high frequency of GC in the third position of codons and a high frequency of using codons preferred by *M. xanthus*, and remains open until it meets ORF1 This suggests that a sequencing error may have shifted the reading frame in the sequence analysis. Hence, the translation initiation

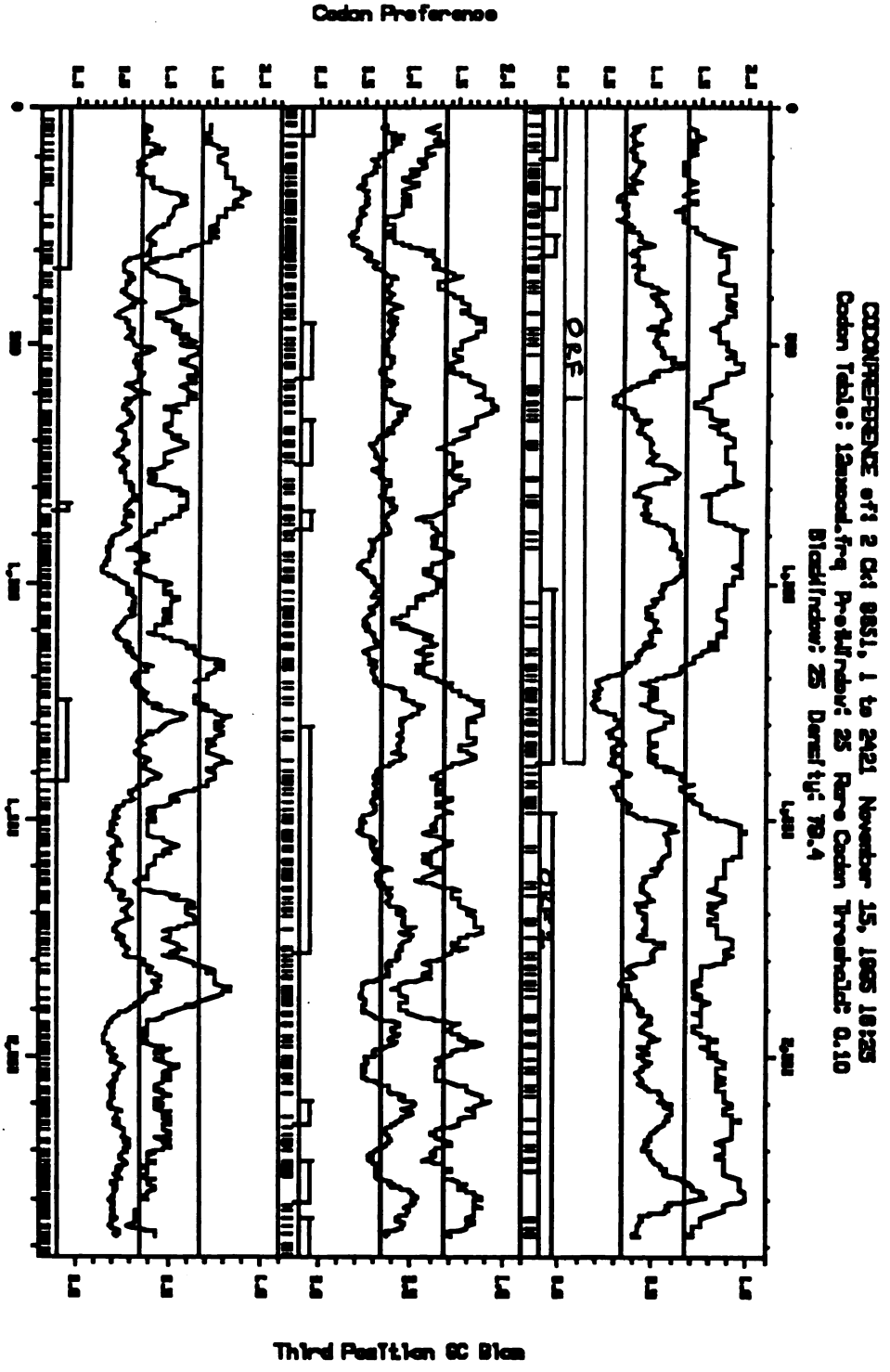
**Figure 3.5.** Nucleotide sequence of 2.1 kb directly upstream of Tn*5* *lac* Ω4499. The two partial open reading frames and their deduced amino acid sequences are indicated. The translation termination codon of ORF1 and the putative translation initiation codon of ORF2 are shown in bold letters. A putative ribosome binding site is shown by the overline.

**ORF1**

```
1     CCGGCCArGCAGGTCCTCCTTGCTGACGGGAACAGGGTCCAGTTCGACCGGTTGCTGATT
      P  A  X  Q  V  L  L  A  D  G  N  R  V  Q  F  D  R  L  L  I

61    GCGACGGGGGCGCGTGCCCGGCCGTGGCCCAACAACGCGGAGGCGTCGCTGGAGGGCGTA
      A  T  G  A  R  A  R  P  W  P  N  N  A  E  A  S  L  E  G  V

121   GAGGTGGTGCGCACCCGCGACGATGCGGCCCAGTTGCAGCGCCGTCTGGCCGCCAGACCC
      E  V  V  R  T  R  D  D  A  A  Q  L  Q  R  R  L  A  A  R  P

181   GGCCGGGTTCTGGTCATCGGCGGCGGGTTCACCGGCTCGGAGGTCGCCTCCGCGTGCCGC
      G  R  V  L  V  I  G  G  G  F  T  G  S  E  V  A  S  A  C  R

241   GAGCTGGGTCTGCCGGTGACCGTCACGGAGCGGGGGGCCACTCCACTGGTGGGGGCGCTC
      E  L  G  L  P  V  T  V  T  E  R  G  A  T  P  L  V  G  A  L

301   GGCGGAGTTGTCGGGGCGGTCGCAGCCGCCTTGCAACGCGACCATGGCGTCGACCTGCGC
      G  G  V  V  G  A  V  A  A  A  L  Q  R  D  H  G  V  D  L  R

361   TGCGGCGTCACCGTCACGCGGCTGGAGGGGGATGCAAACGGGAGGCTGCGGCGGGCCCAC
      C  G  V  T  V  T  R  L  E  G  D  A  N  G  R  L  R  R  A  H

421   TTCTCCGACGGCTCGACACTGAACGTCGACGTGGCCGTGGTCGCGCTAGGGGCGATCCGC
      F  S  D  G  S  T  L  N  V  D  V  A  V  V  A  L  G  A  I  R

481   AACGTGGAGTGCGTGCGCGACTCGGGGCTGGCGGTGACTCGATGGGGAGTGGCCTGTGAT
      N  V  E  C  V  R  D  S  G  L  A  V  T  R  W  G  V  A  C  D

541   GCGGCCTGCCACGCGTTCGACATCAACGGCCTGGTGACCAGGAACATCTTCGTCGCCGGG
      A  A  C  H  A  F  D  I  N  G  L  V  T  R  N  I  F  V  A  G

601   GACGTGGCCCGCGTCCCTCACCCGATCTACGACTACCAGTTCCTGTCCCTCGAGCACTGG
      D  V  A  R  V  P  H  P  I  Y  D  Y  Q  F  L  S  L  E  H  W

661   GGCAACGCGGTCACCCAGGCCAAGATTGCGGCCCACAACATGCTCAGCCCGGAAGCGGAC
      G  N  A  V  T  Q  A  K  I  A  A  H  N  M  L  S  P  E  A  D

721   CGCTGGCCACACCTGACCATCCCCACCTTCTGGTCCACGCAATTCGGCGTCAACATCAAA
      R  W  P  H  L  T  I  P  T  F  W  S  T  Q  F  G  V  N  I  K

781   TCGGTGGGTGTGCCGACCTTCTCGGACGCGCTCGTCATCACCCAGGGCTCACTCGCCGGA
      S  V  G  V  P  T  F  S  D  A  L  V  I  T  Q  G  S  L  A  G

841   AGCGCCGATTCGTCGCCGTCTATGGCTACAAGGGGCGTGTCACCGCCGCGGTCGCTTTCG
      S  A  D  S  S  P  S  M  A  T  R  G  V  S  P  P  R  S  L  S

901   ACAGCCCAAATGGCTGGAGTTCTACCAGGGCTTGATCGATGCAGCCGCGCCGTTCCCGCC
      T  A  Q  M  A  G  V  L  P  G  L  D  R  C  S  R  A  V  P  A

961   CGCCTTCCGCACCGTCCATCAGCCCGCCGAGATGCGGCCCGTCCCTCCCGAATTTCCGAA
      R  L  P  H  R  P  S  A  R  R  D  A  A  R  P  S  R  I  S  E

1021  ACGGGCCGTGTCCGAGCACGAGGCCACCGTCATGGTCACCGGCCGTGAACCCTACGAGCG
      T  G  R  V  R  A  R  G  H  R  H  G  H  R  P  *  T  L  R  A
```

```
1081  GCGGGTTCCGGTGGGTGTACCGGCACCTTTGAGCCGAATGTCGGCAGCCGCGTCACCCCA
      A  G  S  G  G  C  T  G  T  F  E  P  N  V  G  S  R  V  T  P

                                        ORF2
1141  CCCCTCCAACGGACTCAAGGAGACCGCTTCATGGCACAGGCCAGCATCTTCGAGCAGATC
      P  L  Q  R  T  Q  G  D  R  F  M  A  Q  A  S  I  F  E  Q  I

1201  CTCGACCCGGCCAACCGGGCCAACCCCTACCCGCTCTACGCCGAGTTGCGGAAGACTCCC
      L  D  P  A  N  R  A  N  P  Y  P  L  Y  A  E  L  R  K  T  P

1261  GTGGCGCGCGAGGCCGACGGCACGTACATCGTCAGCACCTACGATGAGATTGTCGCGTTG
      V  A  R  E  A  D  G  T  Y  I  V  S  T  Y  D  E  I  V  A  L

1321  CTCCACGACCCACGGGTCAGCTCCGACATACGCAACCTCGTCCGGCAGGCCGGTGCCACG
      L  H  D  P  R  V  S  S  D  I  R  N  L  V  R  Q  A  G  A  T

1381  CCCTCCCCCCAGGAGGGCCCTCCAGGGCTGCCAGAGCCCTTCATCCGGCGCGACCCTCCC
      P  S  P  Q  E  G  P  P  G  L  P  E  P  F  I  R  R  D  P  P

1441  GACCACGACCGGCTCCGGAAGCTGGCGATGAGGCCCTTCGGGCGCACACCCCCGGACGCA
      D  H  D  R  L  R  K  L  A  M  R  P  F  G  R  T  P  P  D  A

1501  TTGATGCCCTTCGTCCCTGGCTGGTCGAGACCACGACGGGCCTGCTTGACGCGCTCGCGG
      L  M  P  F  V  P  G  W  S  R  P  R  R  A  C  L  T  R  S  R

1561  GCAAGAACCAAGTCGACATCGTCGGACGACGTGGCCTACCCATTCCCTGTGACTGTCATC
      A  R  T  K  S  T  S  S  D  D  V  A  Y  P  F  P  V  T  V  I

1621  TGCAAGCTGCTGGGAGTCCCTCGCGAGGACGAAGCGCGCTTTCACGAGTTGGCCGACGCG
      C  K  L  L  G  V  P  R  E  D  E  A  R  F  H  E  L  A  D  A

1681  GGCGTCGAGACCCTTGACCCCACCACAGGGACGATCGAGCAGCGAAAGGCCAAGCGCGAC
      G  V  E  T  L  D  P  T  T  G  T  I  E  Q  R  K  A  K  R  D

1741  CGGACGAAAGCGGAGCTGGGGCAGTACCTCGCTGCGCTCGCCGATGCCCACCTGCGCCAG
      R  T  K  A  E  L  G  Q  Y  L  A  A  L  A  D  A  H  L  R  Q

1801  CCGGGCGGCGACCTGCTCTCCGGGTTCCTGACGGATAACGGCCCCGACGGGCGGATGTCG
      P  G  G  D  L  L  S  G  F  L  T  D  N  G  P  D  G  R  M  S

1861  CGGGAAGAGGTGTTGAGCACCGCCGCGTTGCTGCTCGTCGCTGGCCATGAGACCACCGTC
      R  E  E  V  L  S  T  A  A  L  L  L  V  A  G  H  E  T  T  V

1921  AACCTCATCGCCAATGGCATGCTCACCCTGCTGCGCCACCCCGGCGTGTTCGAGCGGCTG
      N  L  I  A  N  G  M  L  T  L  L  R  H  P  G  V  F  E  R  L

1981  CGCCGCGAGCCCGAGCTGAGCATCCCGCTGGTCGAGGAGCTCGTGCGCTATGAACCACCG
      R  R  E  P  E  L  S  I  P  L  V  E  E  L  V  R  Y  E  P  P

2041  GTCGAGTTCCTGCCTGACCGAGTCACTCTGGCCGACATCGACATTGCCGGCACCACCATT
      V  E  F  L  P  D  R  V  T  L  A  D  I  D  I  A  G  T  T  I

2101  CCCCAGGGC
      P  Q  G
```

**Figure 3.6.** Analysis of the Ω4499 upstream sequence for ORFs likely to be translated. The 2.1 kb DNA sequence presented in Figure 3.5, plus an additional about 300 bp of sequence between 2.4 and 2.1 kb upstream of Ω4499 (based on sequence obtained for only one DNA strand), was analyzed using the codonpreference program of the University of Wisconsin Genetics Computer Group software package, with codon usage data for M. xanthus (Hagen and Shimkets, 1990). The third position GC bias and the codon preference for all three reading frames is plotted. Relative to these plots, the direction of transcription is left to right and Tn*5* *lac* Ω4499 would be located at the right hand end of the plots. Open reading frames identified by the presence of an AUG translation initiation codon and a termination codon are shown as open bars below the plots. The usage of rare codons is indicated by ticks at the bottom of each panel. The numbers along the horizontal axis indicate the number of base pairs from the 5' end of the sequence and the numbers along the vertical axes indicate the third codon-position GC bias (right) and the frequency of using codons preferred by *M. xanthus* (left). The upper and lower plots in each panel reflect the third position GC bias and the codon preference, respectively. Values above the horizontal line associated with each plot indicate favorable sequence for translation in *M. xanthus*.


Codon Preference
CODONPREFERENCE of: 2 Ckt 8851, 1 to 2421 November 15, 1995 10:23
Codon Table: 12bsecd.frq Prob/Index: 25 Rare Codon Threshold: 0.10
Blast/Index: 25 Density: 78.4
ORF 1
Third Position GC Bias

codon of ORF1 may be more than 2.4 kb upstream of the Ω4499 insertion.

Data bank searches using TFASTA and BLASTP showed that there is 32% identity between the deduced 355 amino acid sequence of ORF1 shown in Figure 3.5 and NADH-putidaredoxin reductase, over a 279 amino acid overlap. NADH-putidaredoxin reductase is an enzyme in *Pseudomonas putida*, encoded by the *camA* gene (Koga et al., 1989). *camA* is a component of an operon that consists of three other genes, *camB*, which encodes putidaredoxin, (Koga et al., 1989), *camC*, which encodes a cytochrome P-450 (Koga et al., 1985), and *camD* which encodes 5-exo-hydroxycamphor dehydrogenase (Koga et al., 1986). All four genes in this operon encode enzymes that are involved in the metabolism of camphor.

ORF2 consists of 313 amino acids. The putative translation initiation codon of ORF2 (at 1171 bp) is preceded by a GGAGA sequence seven nucleotides upstream (Figure 3.5), which may serve as a ribosome binding site. The deduced amino acid sequence of ORF2 has 32% identity with two cytochrome P-450 proteins from *Streptomyces griseolus*, $P\text{-}450_{su1}$ and $P\text{-}450_{su2}$. The regions of identity consist of sequences that had previously been shown to be conserved between both P-450 proteins in *S. griseolus* and P-450 in *P. putida* (Omer et al., 1990). These include one of two regions (at positions 1855 to 1957, Figure 3.5) that form α-helices involved in attaching and aligning the heme, and a proline residue (at position 1960) that terminates this helix. Additionally, two of the three conserved amino acids between *P. putida* and *S. griseolus* that were previously shown to be hydrogen bond donors to the proprionate groups of heme in the *Pseudomonas* P-450 are also present at the same positions in ORF2 (the two arginines at positions 1333 and 2059). Finally, the residues proposed to form part of the oxygen-binding site in P-450 are identical between the two P-450 proteins of *S. griseolus* and in the corresponding region of ORF2 (amino acids alanine, glycine, glutamic acid and threonine at positions 1900, 1903, 1909, and 1912, respectively).

The genes that encode P-$450_{su1}$ and P-$450_{su2}$ in *S. griseolus* are components of operons that also encode ferredoxin and an additional less well-characterized ORF (Omer et al., 1990). Both P-450 and ferredoxin are involved in the metabolism of xenobiotics and it has been proposed that the third ORF may also encode an enzyme involved in the same metabolic pathway. By analogy, I speculate that Ω4499 is an insertion into an operon whose products include a reductase and a cytochrome P-450 enzyme involved in the metabolism of alternative carbon source. Sequences that resemble rho-independent transcriptional terminators or known bacterial promoters were not evident between ORF1 and ORF2, supporting the idea that these genes are likely to be present in an operon whose transcriptional initiation site lies more than of 2.4 kb upstream of Ω4499.

## Discussion

*M. xanthus* DNA extending 4.5 kb upstream of Tn*5* *lac* Ω4499 was sufficient to direct proper developmental *lacZ* expression upon integration at Mx8 *att* (Figure 3.1B). Unlike Ω4403 and other fusions that have an absolute requirement for C-signaling (chapter 2), expression of *lacZ* from fusions with Ω4499 upstream DNA at Mx8 *att* was not impaired. Perhaps expression of Ω4499 upstream DNA at an earlier time during development and/or its partial (rather than absolute) C-signal dependency explain the ability of Ω4499 to escape inhibition of expression at Mx8 *att*. In chapter 2, it was proposed that an altered chromosomal state of DNA integrated at Mx8 *att* might affect gene expression. If this model is correct, perhaps activation of Ω4499 transcription occurs early in development and independent of C-signaling prevents the altered chromosomal state from forming. Deletion analyses showed that a promoter element(s) required for developmentally regulated expression of Ω4499 is present more than 2.1 kb upstream

(Figure 3.2). This promoter element or elements could be an upstream activation sequence (UAS) recognized by a DNA-binding activator protein and/or a promoter recognized by RNA polymerase. The latter possibility is supported by the finding that the 3.0 kb region between 1.5 and 4.5 kb upstream of Ω4499 also directed developmental *lacZ* expression, albeit at a later time and at a lower level compared to the Tn*5* *lac* Ω4499 insertion-containing strain, DK4499 (Figure 3.2). Evidence for a UAS located between 4.5 and 3.2 kb upstream of Ω4499 comes from the finding that deleting this region reduced developmental *lacZ* expression by 35% (Figure 3.2). However, 3.2 kb of upstream DNA is sufficient for C-signal dependent expression of Ω4499, since the 3.2 kb segment in wild-type cells directs 65% as much expression as observed for DK4499 (Figure 3.2), but in *csgA* mutant cells it directs only 10-20% as much expression (Figure 3.3).

Northern blot analyses with probes proximal to the Ω4499 insertion site detected a transcript of 3.5 kb that was more abundant in developing cells than in growing cells. If this RNA is the Ω4499 transcript, its presence in growing cells is surprising since expression from Tn*5* *lac* Ω4499 is very low (about the background level) in growing cells. Perhaps the fusion mRNA from the Tn*5* *lac* insertion is less stable than the native Ω4499 mRNA. Probes corresponding to DNA between 2.1 and 1.5 kb upstream of Ω4499 failed to detect the 3.5 kb transcript. This could mean that transcription starts within 1.5 kb of the insertion point; however, two results suggest otherwise. First, DNA between 4.5 and 1.5 kb upstream of Ω4499 did exhibit promoter activity (Figure 3.2). Second, analysis of the upstream DNA sequence identified two ORFs that are likely to be cotranscribed based on analogy with other organisms, and the upstream ORF extends beyond 2.1 kb upstream of Ω4499. If transcription does start more than 2.1 kb upstream of Ω4499, perhaps the transcript is unstable. The 3.5 kb transcript could be a breakdown product of a larger transcript that has lost sequences at its 5' end.

The two ORFs upstream of Ω4499 are separated by 102 bp (Figure 3.5). This

separation is rather large for genes in an operon, but separation of cotranscribed genes by as many as 93 bp has previously been reported (Adams et al., 1988). Transcriptional termination sequences or sequences that resemble known bacterial promoters are not evident within the 102 bp intergenic region, though, of course such sequences could go unrecognized. Perhaps the strongest evidence that Ω4499 is inserted into an operon is the similarity of the deduced amino acid sequences of the two ORFs to proteins that are encoded by genes found in operons in other organisms. ORF1 remains open up to the 5' end of the region for which both DNA strands have been sequenced and this partial ORF would encode a 355 amino acid peptide that is similar to NADH-putidaredoxin reductase from *Pseudomonas putida*. This enzyme is encoded by the *camA* gene, which is present in an operon that also encodes three other enzymes (Koga et al., 1989), putidaredoxin, cytochrome P-450, and 5-exo-hydroxycamphor-dehydrogenase. All four enzymes are involved in the metabolism of D-camphor, which can serve as the sole carbon source for growth of *P. putida*. ORF2 would encode 313 amino acids and remains open downstream to the point of fusion with Ω4499. The deduced amino acid sequence of ORF2 is similar to two cytochrome P-450s from *Streptomyces griseolus*, $P\text{-}450_{su1}$ and $P\text{-}450_{su2}$. The genes that encode both cytochrome P-450s, *suaC* and *subC*, are components of two different operons. These operons also encode ferredoxin (Patel & Omer, 1992). Both cytochrome P-450 and ferredoxin are involved in the metabolism of sulfonylurea herbicide (Patel & Omer, 1992; Omer et al., 1990). A third ORF that is less well-characterized has also been identified in these operons (Patel & Omer, 1992), and is thought to encode enzymes that may be involved in the same metabolic pathway. Interestingly, both cytochrome P-450s from *S. griseolus* have high similarity to the cytochrome P-450 from *P. putida* encoded by *camC*, which is also similar to ORF2.

The initial oxidation of both D-camphor by *P. putida* and sulfonyl urea compounds by *S. griseolus* involves a monooxygenase system that is catalyzed by NADH-

putidaredoxin reductase, putidaredoxin, and cytochrome P-450 in *P. putida* and by ferredoxin and cytochrome P-450 in *S. griseolus*. Monooxygenases catalyze oxygen-dependent reactions that oxidize many organic compounds that are used in growth. This system requires an oxidizable cosubstrate which, in many bacteria, is provided by the reduced form of cytochrome P-450. Typically, enzymes catalyze this reaction according to the equation: monooxygenase

$$\text{substrate-H} + O_2 + AH_2 \text{ ----------> } \text{substrate-OH} + A + H_2O$$

in which $AH_2$ is an oxidizable cosubstrate (Gottschalk, 1986). This system is used in a many different metabolic pathways such as those involved in metabolizing long chain hydrocarbons and ammonia (Gottschalk, 1986), and is present in a diverse group of microorganisms including flavobacteria, mycobacteria, yeast and fungi (Gottschalk, 1986).

By analogy with the operons in *P. putida* and *S. griseolus*, it seems highly likely that Ω4499 is inserted into an operon encoding enzymes that catalyze a monooxygenase reaction(s). The induction of this operon during development suggests that *M. xanthus* may induce alternative metabolic pathways to scavange scarce nutrient sources. For example, compounds released by lysing cells within fruiting bodies may serve as carbon source for the completion of the developmental process in sporulating cells. Since cells containing the Ω4499 insertion showed no obvious developmental defects (Kroos et al., 1986), the putative monooxygenase reaction catalyzed by enzymes encoded by the Ω4499 operon does not appear to be essential for normal development. However, it is possible that other enzymes can substitute for these deficiencies. In this respect, it will be interesting to determine whether the *M. xanthus* genome has other genes that cross-hybridize with genes in the Ω4499 operon. Mutations in the Ω4499 operon could have subtle effects on sporulation efficiency, or effects on spore resistance or germination properties that are not evident upon visual inspection.

# Summary and Perspectives

The aim of the studies presented in this thesis was to characterize the DNA elements involved in regulating the expression of two developmentally regulated transcription units whose expression is modulated by extracellular C-signaling. Ω4403 appears to be inserted into a gene that encodes a serine protease. Sequences between 72 and 80 bp upstream of the transcriptional initiation site were required for the transcription of this gene in a developmentally regulated manner. DNA extending to 80 bp upstream of the transcriptional start site was sufficient for expression in a C-signal-dependent fashion.

The requirement for DNA more than about 50 bp upstream of the Ω4403 transcriptional start site suggests that a protein(s) other than RNA polymerase may activate transcription from the Ω4403 promoter. C-signaling may regulate expression of the Ω4403 gene by modulating (directly or indirectly) either the activity of the putative activator protein or the RNA polymerase that transcribes this gene.

The presence of an activator protein that associates with the -80 to -72 region can be examined by looking for proteins that bind to this region *in vitro*. Analysis of protein extracts from wild-type and *csgA* developing cells can be used to determine whether C-signaling affects the expression of this putative DNA-binding protein. *In vitro* reconstitution experiments would need to be done to demonstrate the role of this protein in activating transcription. An alternative approach is to mutagenize the Ω4403 promoter, then look for suppressor mutations that restore expression.

Expression of the Ω4403 gene is sensitive to its position in the chromosome. At the Mx8 phage attachment (Mx8 *att*) site, expression from the cloned 17 kb upstream DNA is decreased by about 45% compared to expression from the same region at the native site. We hypothesize that the chromosomal structure in the vicinity of the Mx8 *att* site may change as cells progress through development. The chromosome clearly condenses during *B. subtilis* sporulation (Setlow et al., 1991). The effect of chromosomal position on expression of the Ω4403 promoter could be tested further by constructing plasmids

containing Ω4403 upstream DNA fused to *lacZ* that are capable of integrating at different positions in the *M. xanthus* chromosome. Understanding the mechanism of C-signal-dependent inhibition of gene expression is potentially interesting and might be amenable to a genetic approach involving higher expression from the Ω4403 promoter at Mx8 att after mutagenesis.

The analysis of DNA upstream of Tn*5 lac* Ω4499 suggests that it is inserted into a gene that is a component of a developmentally regulated operon. Two ORFs were identified in the 2.1 kb DNA upstream of Ω4499. Since sequence analysis of the region between these ORFs failed to identify transcription termination or initiation signals, it appears that these genes are cotranscribed. Based solely on the sequence analysis, it is not possible to rule out models in which each ORF is transcribed from its own promoter, or the idea that the operon has an internal promoter. This question may be difficult to resolve since the Northern blot data suggest that the Ω4499 transcript may be unstable. An immediate goal is to extend the sequence farther upstream. This will provide information about the location of the likely ORF1 translational start and the presence or absence of additional ORFs that may be part of the operon, and will guide additional deletion analysis of a putative promoter currently thought to be located between 3.2 and 2.1 kb upstream of Ω4499.

It is interesting that both of the developmentally regulated fusions were with genes that may be involved in catabolism. The serine protease encoded by the Ω4403 gene, may be involved in catabolizing proteins and peptides during development, and cytochrome P-450 and NADH-dependent reductase could be involved in metabolizing organic compounds that could serve as carbon sources. The expression of these enzymes during starvation of *M. xanthus* cells may indicate mechanisms used by starving cells to complete spore formation by utilizing carbon from a variety of sources.

The absence of a phenotypic defect in cells harboring insertions in these genes may

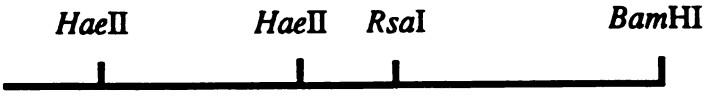
indicate that there are other systems that are turned on during development that compensate for these defects.

# APPENDICES

# Appendix 1

**Figure 1.** Detection of the Ω4403 native message using Northern blot analysis. RNA was isolated from 24 hour developing cells and from vegetatively growing cells as described in Chapter 2. Northern blotting was performed as described previously (Ausubel et al., 1989). RNA was fractionated on a 0.6% agarose gel containing formaldehyde (1.1% v/v), transferred onto a nitrocellulose membrane (Schleicher and Schuell), and baked for two hours at 80°C. Prehybridization was performed at 42°C for 2-3 hours in a solution containing 50% formamide, 6X SSPE, 0.5% SDS, 5X Denhardt's reagent, and 100 μg/ml sheared salmon sperm DNA. Hybridization was performed at 52°C in a fresh solution of prehybridization buffer containing 1-2 x $10^7$ CPM of [$\alpha^{32}$P]dCTP labeled probes synthesized from the *Hae*II-*Hae*II fragment at ~900 to 520 bp upstream of Bam*HI* and from the *Rsa*I-*Bam*HI fragment ( 451 bp upstream of *Bam*HI) isolated from pMF01. Probes were synthesized using a random priming kit from Stratagene according to the manufacturers specifications. The numbers on the left indicate the molecular weight of single stranded size markers (Bethesda Research Laboratories, Inc.).

The *Rsa*I-*Bam*HI fragment detected a transcript of about 1.5 kb in RNA isolated from DK4499 cell that were developed for 24 hours but not in the vegetative RNA. More than two faint transcripts were detected in the RNA isolated from 24 hour developing DK4368 cells. No transcripts were detected with the *Hae*II-*Hae*II fragment in RNA isolated from either strain (data not shown).




Tn*5 lac* Ω4403 Wild-type

0 24 0 24

# Appendix 2

**Figure 2A.** Chromosomal structure of *M. xanthus* containing a pREG1666-derived plasmid at Mx8 *att*. Strains containing a single copy of the integrated plasmid are identified by the 2.0 and 10.0 kb flanking fragments upon digestion with *Eco*RI and *Xho*I, and these fragments are detected by a probe corresponding to the *attP* fragment. Unintegrated plasmid or strains containing multiple copies of integrated plasmid are identified by a 4.0 kb band.

A

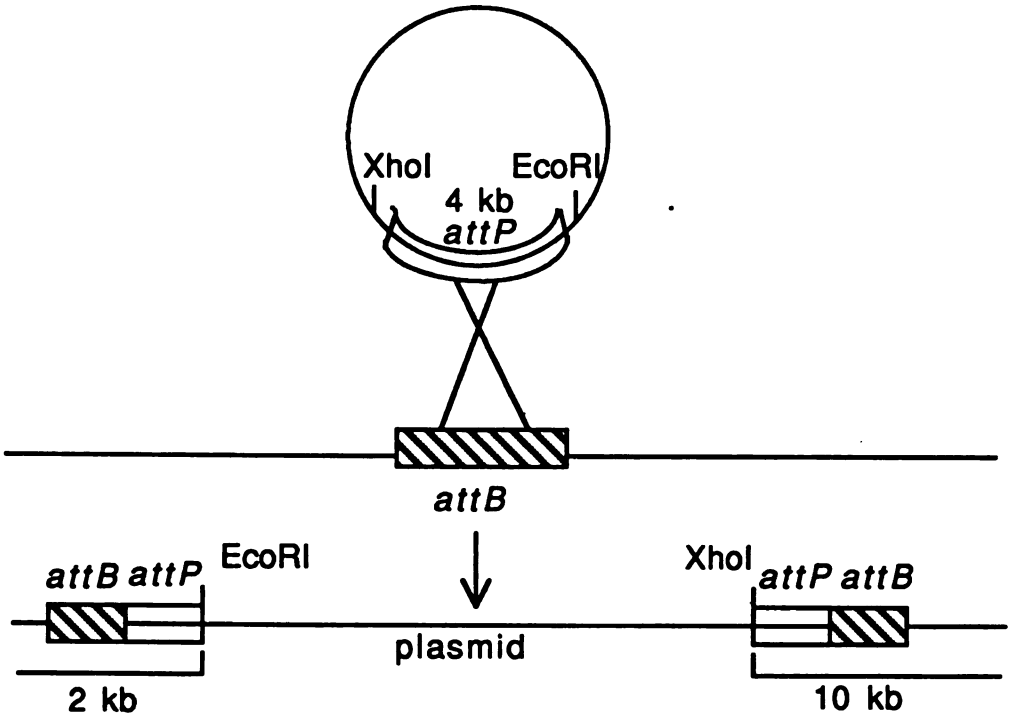
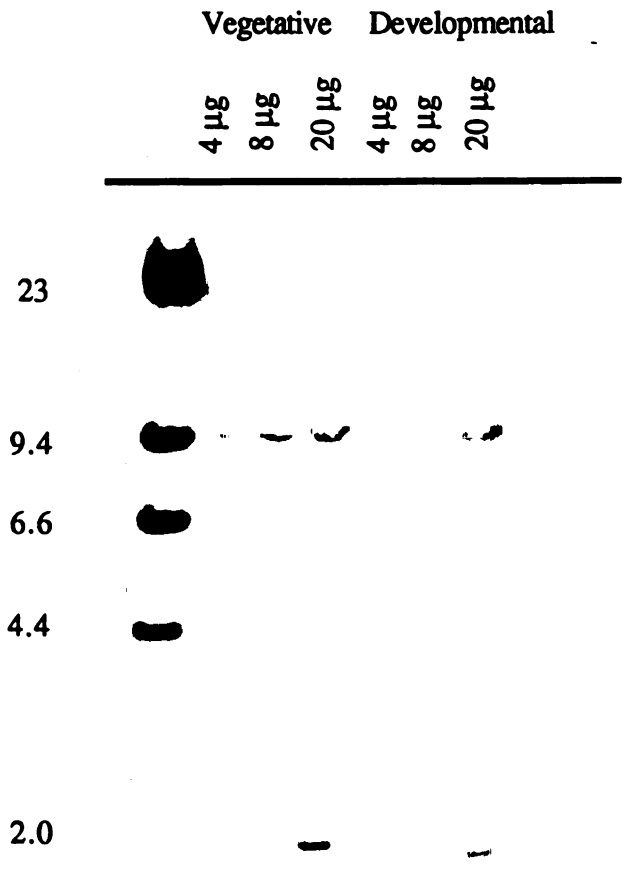

XhoI
EcoRI
4 kb
attP
attB
attB attP
EcoRI
plasmid
XhoI
attP attB
2 kb
10 kb

**Figure 2B.** DNA was isolated as described previously (Avery & Kaiser, 1983) from vegetatively growing MMF31-7 cells and from MMF31-7 cells that had developed for 24 hours on MC7 agarose plates (10 mM morpholine-propanesulfonic acid; pH7.0, 1 mM $CaCl_2$, 1% agarose). 4-20 μg of DNA was digested with *Eco*RI and *Xho*I, fractionated on a 0.5% agarose gel. Transfer of the gel onto a Nylon membrane, pehybridization, and hybridization were carried out as described previously (Avery & Kaiser, 1983). A 3.0 kb *attP* fragment (isolated as described in Chapter 2) was radioactively labeled with [$\alpha^{32}$P]dCTP using a random priming kit from Stratagene, and 1-2 x$10^7$ CPM was used for hybridization.

**B**


Vegetative
Developmental
4 µg
8 µg
20 µg
4 µg
8 µg
20 µg
23
9.4
6.6
4.4
2.0

# Appendix 3

**Table 1.** Sporulation efficiency of *Myxococcus xanthus* strains

| *Myxococcus xanthus* strains | Sporulation efficiency [a] |
|---|---|
| DK1622 | 0.3 |
| DK4368 | 0.1 |
| MMF31-7 | 0.1 |

[a] Spore titers were determined three days after development using the procedure of Kroos and Kaiser (Kroos & Kaiser, 1987), except that cells were spotted onto TPM (10 mM Tris, pH 8.0, 1 mM $KH_2PO_4$, 8 mM $MgSO_4$, final pH 7.6) agar plates (1.5 % Bacto-agar). Sporulation efficiency is presented as a percentage of the number of cells initially plated.

# BIBLIOGRAPHY

# Bibliography

Adams, C. W., Fornwald, J. A., Schmidt, J. F., Rosenberg, M. & Brawner, M. E. (1988). Gene organization and structure of Streptomyces lividans gal operon. *J. Bacteriol.*, **170**, 203-212.

Apelian, D. & Inouye, S. (1990). Development-specific σ-factor essential for late-stage differentiation of *Myxococcus xanthus*. *Genes Dev.*, **4**, 1396-1403.

Apelian, D. & Inouye, S. (1993). A new putative sigma factor of *Myxococcus xanthus*. *J. Bacteriol.*, **175**, 3335-3342.

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. & Struhl, K. (1989). In *Current Protocols in Molecular Biology*. eds), Pages, John Wiley & Sons, Place Published.

Avery, L. & Kaiser, D. (1983). *In situ* transposon replacement and isolation of a spontaneous tandem genetic duplication. *Mol. Gen. Genet.*, **191**, 99-109.

Baker, M. E. (1994). *Myxococcus xanthus* C-factor, a morphogenetic paracrine signal, is similar to *Escherichia coli* 3-oxoacyl-[acyl-carrier-protein] reductase and human 17ß-hydroxysteroid dehydrogenase. *Biochem. J.*, **301**, 311-312.

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

Brandner, J. & Kroos, L. unpublished data.

Brandner, J., White, J. & Kroos, L. unpublished data.

Burton, Z., Gross, C., Watanabe, K. & Burgess, R. (1983). The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in *E. coli* K12. *Cell*, **32**, 335-349.

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

Butler, J. S., Springer, M. & Grunberg-Manago, M. (1987). AUU-to-AUG mutation in the initiator codon of the translation initiation factor IF3 abolishes translational auto control of its own gene (*infC*) in vivo. *Proc. Natl. Acad. Sci. USA*, **84**, 4022-4025.

Cheng, Y. & Kaiser, D. (1989a). *dsg* , a gene required for cell-cell interaction early in *Myxococcus* development. *J. Bacteriol.*, **171**, 3719-3726.

Cheng, Y. & Kaiser, D. (1989b). *dsg* , a gene required for *Myxococcus* development, is necessary for cell viability. *J. Bacteriol.*, **171**, 3727-3731.

Clark, H. R., Leigh, J. A. & Duglas, C. J. (1992). Molecular signals in the interactions between plants and microbes. *Cell*, **71**, 191-199.

Clewell, D. B. & Weaver, K. E. (1989). Sex pheromones and plasmid transfer in *Enterococcus faecalis*. *Plasmid*, **21**, 175-184.

Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.*, **55**, 371-394.

Cumsky, M. G. & Zusman, D. R. (1979). Myxobacterial hemagglutinin: a development-specific lectin of *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. USA*, **76**, 5505-5509.

Davis, J. M., Mayor, J. M. & Plamann, L. S. (1995). *asgC767* is an allele of the major sigma factor gene of *Myxococcus xanthus*. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 16.

Davis, R., Botstein, D. & Roth, J. (1980). In *Advanced bacterial genetics*. eds), Pages, Cold Spring Harbor Laboratory Press, Place Published.

Dillin, A. & Rine, J. (1995). On the origin of a silencer. *Trends Biochem. Sci.*, **20**, 231-235.

Downard, J. & Kroos, L. (1993). In *Transcriptional regulation of developmental gene expression in Myxococcus xanthus*. (Dworkin, M.& Kaiser, D.s, eds), 183-199, American Society for Microbiology, Place Published.

Downard, J., Ramaswamy, S. V. & Kil, K. (1993). Identification of esg, a genetic locus involved in cell-cell signaling during Myxococcus xanthus development. *J. Bacteriol.*, **175**, 7762-7770.

Downard, J. S., Kupfer, D. & Zusman, D. R. (1984). Gene expression during development of *Myxococcus xanthus*: analysis of the genes for protein S. *J. Mol. Biol.*, **175**, 469-492.

Downard, J. S. & Zusman, D. R. (1985). Differential expression of protein S genes during *Myxococcus xanthus* development. *J. Bacteriol.*, **161**, 1146-1155.

Dunny, G. M. (1990). Genetic functions and cell-cell interactions in the pheromone-inducible plasmid transfer system of *Enterococcus faecalis*. *Mol. Microbiol.*, **4**, 689-696.

Dworkin, M. & Kaiser, D. (1993). In *Myxobacteria II*. eds), Pages, American Society for Microbiology, Place Published.

Fawcett, T. & Bartlett, S. (1990). An effective method for eliminating "artifact banding" when sequencing double-stranded DNA templates. *Biotechniques*, **9**, 46-48.

Fink, J. M. & Zissler, J. F. (1989). Defects in motility and development of *Myxococcus xanthus* lipopolysaccharide mutants. *J. Bacteriol.*, **171**, 2042-2048.

Fisseha, M. & Kroos, L. unpublished data.

Frasch, S. C. & Dworkin, M. (1995). Tyrosine kinases with differential activity during development in *Myxococcus xanthus*. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 39.

Gill, R. E. & Bornemann, M. C. (1988b). Identification and characterization of the *Myxococcus xanthus bsg*A gene product. *J. Bacteriol.*, **170**, 5289-5297.

Gill, R. E. & Cull, M. G. (1986). Control of developmental gene expression by cell-to-cell interactions in *Myxococcus xanthus*. *J. Bacteriol.*, **168**, 341-347.

Gill, R. E., Cull, M. G. & Fly, S. (1988a). Genetic identification and cloning of a gene required for developmental cell interactions in *Myxococcus xanthus*. *J. Bacteriol.*, **170**, 5279-5288.

Gill, R. E., Karlok, M. & Benton, D. (1993). *Myxococcus xanthus* encodes an ATP-dependent protease which is required for developmental gene transcription and intercellular signaling. *J. Bacteriol.*, **175**, 4538-4544.

Gilman, M. & Chamberlin, M. J. (1983). Developmental and genetic regulation of *Bacillus subtilis* genes transcribed by $\sigma^{28}$-RNA polymerase. *Cell*, **35**, 285-293.

Gold, L., Stormo, G. & Saunders, R. (1984). Escherichia coli translational initiation factor IF3: a unique case of translational regulation. *Proc. Natl. Acad. Sci. USA*, **81**, 7061-7065.

Gottschalk, G. (1986). In *Bacterial Metabolism.*, 359 Pages, Springer-Verlag, Place Published.

Grossman, A. D. & Losick, R. (1988). Extracellular control of spore formation in *Bacillus subtilis*. *Proc. Nat. Acad. Sci. USA*, **85**, 4369-4373.

Guespin-Michel, J., Letouvet-Pawlak, B. & Petit, F. (1993). In *Protein secretion in myxobacteria*. (Dworkin, M.& Kaiser, D.s, eds), 235-255, American Society for Microbiology, Place Published.

Guo, D., Bowden, G. & Kaplan, H. B. (1995). Analysis of the *Myxococcus xanthus sasA* locus a negative regulator of an A-signal-dependent gene. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 19.

Hagen, D. C., Bretscher, A. P. & Kaiser, D. (1978). Synergism between morphogenetic mutants of *Myxococcus xanthus*. *Dev. Biol.*, **64**, 284-296.

Hagen, T. J. & Shimkets, L. J. (1990). Nucleotide sequence and transcriptional products of the *csg* locus of *Myxococcus xanthus*. *J. Bacteriol.*, **172**, 15-23.

Haldenwang, W. G. (1995). The sigma factors of *Bacillus subtilis*. *Microbiol. Rev.*, **59**, 1-30.

Hao, T. & Kroos, L. unpublished data.

Hartzell, P. & Kaiser, D. (1991). Function of MglA, a 22-kilodalton protein essential for gliding in *Myxococcus xanthus*. *J. Bacteriol.*, **173**, 7615-7624.

Helmann, J. (1995). Compilation and analysis of *Bacillus subtilis* $\sigma^{A}$-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.*, **in press**,

Herskowitz, I. (1989). A regulatory heirarchy for cell specialization in yeast. *Nature*, **342**, 749-757.

Hodgkin, J. & Kaiser, D. (1977). Cell-to-cell stimulation of motility in nonmotile mutants of *M. xanthus*. *Proc. Natl. Acad. Sci. USA*, **74**, 2938-2942.

Hodgkin, J. & Kaiser, D. (1979b). Genetics of Gliding Motility in *Myxococcus xanthus* (Myxobacterales): Two Gene Systems Control Movement. *Mol. Gen. Genet.*, **171**, 177-191.

Igo, M. M. & Losick, R. (1986). Regulation of a promoter region that is utilized by minor forms of RNA polymerase holoenzyme in *Bacillus subtilis*. *J. Mol. Biol.*, **191**, 615-624.

Inouye, M., Inouye, S. & Zusman, D. R. (1979a). Biosynthesis and self-assembly of protein S, a development-specific protein of *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. USA*, **76**, 209-213.

Inouye, M., Inouye, S. & Zusman, D. R. (1979b). Gene expression during development of *Myxococcus xanthus*: pattern of protein synthesis. *Dev. Biol.*, **68**, 579-591.

Inouye, S. (1990). Cloning and DNA sequence of the gene coding for the major sigma factor from *Myxococcus xanthus*. *J. Bacteriol.*, **172**, 80-85.

Inouye, S., Franceschini, T. & Inouye, M. (1983a). Structural similarities between the development-specific protein S from a Gram-negative bacterium, *Myxococcus xanthus*, and calmodulin. *Proc. Natl. Acad. Sci. USA*, **80**, 6829-6833.

Inouye, S., Ike, Y. & Inouye, M. (1983b). Tandem repeat of the genes for protein S, a development-specific protein of *Myxococcus xanthus*. *J. Biol. Chem.*, **258**, 38-40.

Jacobs, M., Eliasson, M., Uhlen, M. & Flock, J. I. (1985). Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*. *Nucleic Acids Res.*, **13**, 8913-8926.

Kaiser, D. (1986). Control of multicellular development: *Dictyostelium* and *Myxococcus*. *Ann. Rev. Genet.*, **20**, 539-566.

Kaiser, D. & Crosby, C. (1983). Cell movement and its coordination in swarms of *Myxococcus xanthus*. *Cell Motil.*, **3**, 227-245.

Kaiser, D. & Kroos, L. (1993). In *Intercellular Signaling*. (Dworkin, M.& Kaiser, D.s, eds), 257-284, American Society for Microbiology, Place Published.

Kaiser, D., Kroos, L. & Kuspa, A. (1985). In *Cell interactions govern the temporal pattern of Myxococcus development*. eds), 823-830, Cold Spring Harbor Laboratory, Place Published.

Kalman, L. V., Cheng, Y. & Kaiser, D. (1994). The *Myxococcus xanthus dsg* gene product performs functions of translation initiation factor IF3 *In vivo*. *J. Bacteriol.*, **176**, 1434-1442.

Kaplan, H. B., Kuspa, A. & Kaiser, D. (1991). Suppressors that permit A-signal-independent development gene expression in *Myxococcus xanthus*. *J. Bacteriol.*, **173**, 1460-1470.

Keseler, I. M. & Kaiser, D. (1995). The role of $\sigma^{54}$ in transcription of the developmentally regulated gene *4521*. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 20.

Kil, K.-S., Brown, G. & Downard, J. (1990). A segment of *Myxococcus xanthus ops* DNA functions as an upstream activation site for *tsp* gene transcription. *J. Bacteriol.*, **172**, 3081-3088.

Kim, S. K. (1991). Intercellular signaling in *Myxococcus* development: The role of C-factor. *Trends in Genet.*, **7**, 361-365.

Kim, S. K. & Kaiser, D. (1990a). Cell alignment required in differentiation of *Myxococcus xanthus*. *Science*, **249**, 926-928.

Kim, S. K. & Kaiser, D. (1990b). Cell motility is required for the transmission of C-factor, an intercellular signal that coordinates fruiting body morphogenesis of *Myxococcus xanthus*. *Genes Dev.*, **4**, 896-905.

Kim, S. K. & Kaiser, D. (1990a). C-factor: a cell-cell signaling protein required for fruiting body morphogenesis of M. xanthus. *Cell*, **61**, 19-26.

Kim, S. K. & Kaiser, D. (1990b). Purification and properties of *myxococcus xanthus* C-factor, an intercellular signaling protein. *Proc. Natl. Acad. Sci. USA*, **87**, 3635-3639.

Kim, S. K. & Kaiser, D. (1991). C-factor has distinct aggregation and sporulation thresholds during *Myxococcus* development. *J. Bacteriol.*, **173**, 1722-1728.

Koga, H., Aramaki, H., Yamaguchi, E., Takeuchi, K., Horiuchi, T. & Gunsalus, I. C. (1986). camR, a negative regulator locus of the cytochrome P-450 $_{cam}$ hydrolyase operon. *J. Bacteriol.*, **166**, 1089-1095.

Koga, H., Rauchfuss, B. & Gunsalus, I. C. (1985). P-450$_{cam}$ gene cloning and expression in *Pseudomonas putida* and *Escherichia coli*. *Biochem. Biophys. Res. Commun.*, **130**, 412-417.

Koga, H., Yamaguchi, E., Matsunaga, K., Aramaki, H. & Horiuchi, T. (1989). Cloning and nucleotide sequence of NADH-putidaredoxin reductase gene (*camA*) and putidaredoxin gene (*camB*) involved in cytochrome P-450cam hydroxylase of *Psedomonas putida*. *J. Biochem.*, **106**, 831-836.

Kroos, L. & Cutting, S. (1994). In *Intercellular and intercompartmental communication during Bacillus subtilis sporulation*. (Piggot, P. J., C.P. Moran, J.& Youngman, P.s, eds), 155-180, Amer. Soc. for Microbiol., Place Published.

Kroos, L., Hartzell, P., Stephens, K. & Kaiser, D. (1988). A link between cell movement and gene expression argues that motility is required for cell-cell signaling during fruiting body development. *Genes Dev.*, **2**, 1677-1685.

Kroos, L. & Kaiser, D. (1984). Construction of Tn*5 lac*, a transposon that fuses *lacZ* expression to exogenous promoters, and its introduction into *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. USA*, **81**, 5816-5820.

Kroos, L. & Kaiser, D. (1987). Expression of many developmentally regulated genes in *Myxococcus* depends on a sequence of cell interactions. *Genes Dev.*, **1**, 840-854.

Kroos, L., Kuspa, A. & Kaiser, D. (1986). A global analysis of developmentally regulated genes in *Myxococcus xanthus*. *Dev. Biol.*, **117**, 252-266.

Kuspa, A. & Kaiser, D. (1989). Genes required for developmental signaling in *Myxococcus xanthus*: three *asg* loci. *J. Bacteriol.*, **171**, 2762-2772.

Kuspa, A., Kroos, L. & Kaiser, D. (1986). Intercellular signaling is required for developmental gene expression in *Myxococcus xanthus*. *Dev. Biol.*, **117**, 267-276.

Kuspa, A., Plamann, L. & Kaiser, D. (1992a). Identification of heat-stable A-factor from *Myxococcus xanthus*. *J. Bacteriol*, **174**, 3319-3326.

Kuspa, A., Plamann, L. & Kaiser, D. (1992b). A-signalling and the cell density requirement for *Myxococcus xanthus* development. *J. Bacteriol.*, **174**, 7360-7369.

LaRossa, R., Kuner, J., Hagen, D., Manoil, C. & Kaiser, D. (1983). Developmental cell interactions of *Myxococcus xanthus*: analysis of mutants. *J. Bacteriol.*, **153**, 1394-1404.

Laue, B. E. & Gill, R. (1994). Use of a phase variation-specific promotor of *Myxococcus xanthus* in a strategy for isolating a phase-locked mutant. *J. Bacteriol.*, **176**, 5341-5349.

Lee, B., Lee, K., Robles, J. & Shimkets, L. J. (1995a). Developmental C-signaling of *Myxococcus xanthus* involves an extracellular NAD(P)-containing protein. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 24.

Lee, B.-U., Lee, K., Robles, J. & Shimkets, L. (1995b). A tactile sensory system of *Myxococcus xanthus* involves an extracellular NAD(P)+-containing protein. *Genes Dev.*, **submitted**,

Lee, K. & Shimkets, L. (1994). Cloning and characterization of the *socA* locus which restores development to *Myxococcus xanthus* C-signaling mutants. *J. Bacteriol.*, **176**, 2200-2209.

Lee, K. & Shimkets, L. (1995). Suppression of a signaling defect during *Myxococcus xanthus* development. *J. Bacteriol.*, **submitted**,

Lee, Y.-C., Miyata, Y., Terada, I., Ohta, T. & Matsuzawa, H. (1991). Involvement of $NH_2$-terminal pro-sequence in the production of active aqualysin I (a thermophilic serine protease) in *Escherichia coli*. *Agric. Biol. Chem.*, **55**, 3027-3032.

Li, S.-F., Lee, B. & Shimkets, L. J. (1992). *csgA* expression entrains *Myxococcus xanthus* development. *Genes Dev.*, **6**, 401-410.

Li, S.-F. & Shimkets, L. J. (1988). Site-specific integration and expression of a

developmental promoter in *Myxococcus xanthus*. *J. Bacteriol.*, **170**, 5552-5556.

Li, S.-F. & Shimkets, L. J. (1993). Effect of *dsp* mutations on the cell-to-cell transmission of CsgA in *Myxococcus xanthus*. *J. Bacteriol.*, **175**, 3648-3652.

Lisser, S. & Margalit, H. (1993). Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.*, **21**, 1507-1516.

MacNeil, S. D., Calara, F. & Hartzell, P. (1994). New clusters of genes required for gliding motility in *Myxococcus xanthus*. *Molec. Microbiol.*, **14**, 61-71.

McCleary, W. R., Esmon, B. & Zusman, D. R. (1991). Myxococcus xanthus protein C is a major spore surface protein. *J. Bacteriol.*, **173**, 2141-2145.

Munoz-Dorado, J., Inouye, S. & Inouye, M. (1991). A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a Gram-negative bacterium. *Cell*, **67**, 995-1006.

Nelson, D. & Zusman, D. (1983). Evidence for long-lived mRNA during fruiting body formation in *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. USA*, **80**, 1467-1471.

O'Connor, K. A. & Zusman, D. R. (1989). Patterns of cellular interactions during fruiting-body formation in *Myxococcus xanthus*. *J. Bacteriol*, **171**, 6013-6042.

Omer, C. A., Lenstra, R., Litle, P. J., Dean, C., Tepperman, J. M., Leto, K. J., Romesser, J. A. & O'Keef, D. P. (1990). Genes for two herbicide-inducible cytochrome P-450 from *Streptomyces griseolus*. *J. Bacteriol.*, **172**, 3335-3345.

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

Parkinson, J. S. (1993). Signal transduction schemes of bacteria. *Cell*, **73**, 857-871.

Patel, N. V. & Omer, C. A. (1992). Phenobarbital and sulfonylurea-inducible operons encoding herbicide metabloizing cytochromes P-450 in *Streptomyces griseolus*. *Gene*, **112**, 67-76.

Persson, B., Krook, M. & Jornvall, H. (1991). Characteristics of short-chain alcohol dehdrogenases and related enzymes. *Eur. J. Biochem.*, **200**, 537-543.

Plamann, L., Davis, J., Cantwell, B. & Mayor, J. (1994). Evidence that *asgB* encodes a DNA-binding protein essential for growth and development of *Myxococcus xanthus*. *J. Bacteriol*, **176**, 2013-2022.

Plamann, L., Kuspa, A. & Kaiser, D. (1992). Proteins that rescue A-signal-defective mutants of *Myxococcus xanthus*. *J. Bacteriol*, **174**, 3311-3318.

Plamann, L., Li, Y., Cantwell, B. & Mayor, J. (1995). The *Myxococcus xanthus asgA* gene encodes a novel signal transduction protein required for multicellular development. *J. Bacteriol.*, **177**, 2014-2020.

Ptashne, M. (1987). In *A genetic switch.*, Pages, Cell press and Blackwell scientific.

Pugsley, A. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.*, **57**, 50-108.

Rhie, H.-G. & Shimkets, L. J. (1991). Low-temperature induction of *Myxococcus xanthus* developmental gene expression in wild-type and *csgA* suppressor cells. *J. Bacteriol.*, **173**, 2206-2211.

Romeo, J. M. & Zusman, D. R. (1991). Transcription of the myxobacterial hemagglutinin gene is mediated by a $\sigma^{54}$-like promoter and a *cis*-acting upstream regulatory region of DNA. *J. Bacteriol.*, **173**, 2969-2976.

Romeo, J. M. & Zusman, D. R. (1992). Determinants of an unusually stable mRNA in the bacterium *Myxococcus xanthus*. *Mol. Microbiol.*, **6**, 2975-2988.

Rosenberg, E., Keller, K. & Dworkin, M. (1977). Cell density-dependent growth of Myxococcus xanthus on casein. *J. Bacteriol.*, **129**, 770-777.

Rosenbluh, A. & Rosenberg, E. (1989). Autocide AMI rescues development in *dsg* mutants of *Myxococcus xanthus*. *J. Bacteriol.*, **171**, 1513-1518.

Rudner, D. Z., LeDeaux, J. R., Ireton, K. & Grossman, A. D. (1991). The *spo0K* locus of *Bacillus subtilis* is homologous to the oligopeptide permease locus and is required for sporulation and competence. *J. Bacteriol.*, **173**, 1388-1398.

Sager, B. & Kaiser, D. (1993a). Spatial restriction of cellular differentiation. *Genes Dev.*, **7**, 1645-1653.

Sager, B. & Kaiser, D. (1993b). Two cell-density domains within the *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. USA*, **90**, 3690-3694.

Sager, B. & Kaiser, D. (1994). Intercellular C-signaling and the traveling waves of *Myxococcus*. *Genes Dev.*, **8**, 2793-2804.

Sambrook, J., Fritsch, E. & Maniatis, T. (1989). In *Molecular Cloning*. eds), Pages, Cold Spring Harbor Laboratory Press, Place Published.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Setlow, B., Magill, P., Febbroriello, P., Nakhimovsky, L., Koppel, D. & Setlow, P. (1991). Condensation of the forespore nucleoid early in sporulation of Bacillus species. *J. Bacteriol.*, **173**, 6270-6278.

Shimkets, L. J. personal communication.

Shimkets, L. J. (1990). Social and developmental biology of the myxobacteria. *Microbiol. Rev.*, **54**, 473-501.

Shimkets, L. J. & Asher, S. J. (1988). Use of recombination techniques to examine the structure of the *csg* locus of *Myxococcus xanthus*. *Mol. Gen. Genet.*, **211**, 63-71.

Shimkets, L. J., Gill, R. E. & Kaiser, D. (1983). Developmental cell interactions in

*Myxococcus xanthus* and the *spoC* locus. *Proc. Natl. Acad. Sci. USA*, **80**, 1406-1410.

Shimkets, L. J. & Kaiser, D. (1982a). Induction of coordinated movement of *Myxococcus xanthus* cells. *J. Bacteriol.*, **152**, 451-461.

Shimkets, L. J. & Kaiser, D. (1982b). Murein components rescue developmental sporulation of *Myxococcus xanthus*. *J. Bacteriol.*, **152**, 462-470.

Shimkets, L. J. & Rafiee, H. (1990). CsgA, an extracellular protein essential for *myxococcus xanthus* development. *J. Bacteriol.*, **172**, 5299-5306.

Simons, R., Houman, F. & Kleckner, N. (1987). Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85-96.

Singer, M. & Kaiser, D. (1995). Ectopic production of guanosine penta-and tetraphosphate can initiate early developmental gene expression in *Myxococcus xanthus*. *Genes and Dev.*, **9**, 1633-1644.

Sodergren, E. & Kaiser, D. (1983). Insertions of Tn*5* near genes that govern stimulatable cell motility in *Myxococcus*. *J. Mol. Biol.*, **167**, 295-310.

Stellwag, E., Fink, J. M. & Zissler, J. (1985). Physical characterizatin of the genome of the *Myxococcus xanthus* bacteriophage MX-8. *Mol. Gen. Genet.*, **199**, 123-132.

Stephens, K., Hartzell, P. & Kaiser, D. (1988). Gliding motility in *Myxococcus xanthus* : *mgl* locus, RNA and predicted protein products. *J. Bacteriol.*, **171**, 819-830.

Stephens, K. & Kaiser, D. (1987). Genetics of gliding motility in *Myxococcus xanthus*: molecular cloning of the *mgl* locus. *Mol. Gen. Genet.*, **207**, 256-266.

Sternberg, P. W. & Horvitz, H. R. (1989). The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. *Cell*, **58**, 679-693.

Stock, J. B., Stock, A. M. & Mottonen, J. M. (1990). Signal transduction in bacteria. *Nature*, **344**, 395-400.

Teintze, M., Thomas, R., Furuichi, T., Inouye, M. & Inouye, S. (1985). Two homologous genes coding for spore-specific proteins are expressed at different times during development of *Myxococcus xanthus*. *J. Bacteriol.*, **163**, 121-125.

Terada, I., Kwon, S., Miyata, Y., Matsuzawa, H. & Ohta, T. (1990). Unique precursor structure of an extracellular protease, aqualysin I, with $NH_2$- and COOH-terminal pro-sequences and its processing in *Escherichia coli*. *J. Biol. Chem.*, **265**, 6576-6581.

Thony-Meyer, L. & Kaiser, D. (1993). *devRS*, an autoregulated and essential genetic locus for fruiting body development in *Myxococcus xanthus*. *J. Bacteriol.*, **175**, 7450-7462.

Toal, D. R., Clifton, S., Roe, B. & Downard, J. (1995). The *esg* locus of *Myxococcus xanthus* encodes the E1$\alpha$ and E1$\beta$ subunits of a branced-chain keto acid dehydrogenase. *Mol. Microbiol.*, **16**, 177-189.

Udo, H., Munoz-Dorado, J., Inouye, M. & Inouye, S. (1995). *Myxococcus xanthus*, a Gram-negative bacterium, contains a transmembrane protein serine/threonine kinase that blocks secretion of ß-lactamase by phosphorylation. *Genes Dev.*, **9**, 972-983.

Weinberg, R. A. & Zusman, D. R. (1990). Alkaline, acid and neutral phosphatase activities are induced during development in *Myxococcus xanthus*. *J. Bacteriol.*, **172**, 2294-2302.

Willey, J., Schwedock, J. & Losick, R. (1993). Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. *Genes and Dev.*, **7**, 895-903.

Yang, C. & Kaplan, H. (1995). The *Myxococcus xanthus* wild-type *sasB7* gene encodes a member of the histidine kinase family. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 53.

Yanisch-Perron, C., Vieria, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences M13 mp18 and pUC19 vectors. *Gene*, **33**, 103-119.

York, K., Kenney, T., Satola, S., Moran, C. J., Poth, H. & Youngman, P. (1992). Spo0A controls the $\sigma^A$-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIE*. *J. Bacteriol.*, **174**, 2648-2658.

Zhang, W., Inouye, M. & Inouye, S. (1995). Reciprocal roles of two protein Ser/Thr kinase genes, *pkn5* and *pkn6* in *Myxococcus xanthus* development that are diversely located on the chromosome with a 46 bp distance. *Abstr. of the 22nd Intl. Conf. on the Biol. of the Myxobacteria*, 41.